Names | A Journal of Onomastics

Monroe Lake or *Lake Monroe*? Prosodic, Grammatical, and Semantic Influences on Word Order in US Place Names

Michael H. Kelly
Independent Scholar

ans-names.pitt.edu
ISSN: 0027-7738 (print) 1756-2279 (web) Vol. 73 No. 4, Winter 2025 DOI 10.5195/names.2025.2824

Abstract

This paper describes prosodic, grammatical, semantic, and other lexical characteristics that influence the choice between "X Lake" and "Lake X" in US place names. Prosodically, the incidence of "Lake X" rises relative to "X Lake" when "X" is a disyllabic word with iambic rather than trochaic stress (e.g., "guitar" versus "glacier"). This pattern is consistent with a preference for rhythmic alternation in English, which is maintained when trochaic "X" appears in "X Lake" and iambic "X" in "Lake X". Grammatically, "X Lake" increases in likelihood with the tendency for "X" to be used as an adjective, maintaining the standard adjective-noun word order in English noun phrases. Semantically, "Lake X" is more likely when "X" is a proper name or expresses positive sentiment, suggesting that a notion of psychological "prominence" affects the choice between "X Lake" and "Lake X", with the latter, rarer word order associated with heightened psychological prominence. Finally, "Lake X" likelihood increases with the number of syllables and frequency of use of "X". Implications and directions for further research are discussed, such as how to better characterize naming patterns that are attributed to the "euphony" they create and what are avenues for refining our understanding of psychological "prominence".

Keywords: toponym, hydronym, word order, rhythmic alternation, proper name, word sentiment

General Introduction

Most lake names in the United States have the structure "X Lake" as in *Eagle Lake*, *Breezy Lake,* and *Mosquito Lake*. However, the word order in a significant minority is "Lake X" as in *Lake Hamilton*, *Lake Serene*, and *Lake Providence*. This paper will test whether specific prosodic, grammatical, semantic, and other lexical characteristics influence the choice between these word orders. These potential influences will be described in turn.

Prosody: The Preference for Rhythmic Alternation

A preference for alternations between stressed and unstressed syllables has been argued to be a fundamental property of English prosody (e.g., Nespor & Vogel 1986; Selkirk 1984). This preference for rhythmic alternation is expressed in several ways:

- Stress patterns in polysyllabic words like "testimony" and "participate" alternate between strong and weak beats (Halle & Vergnaud 1987; Selkirk 1984; Liberman & Prince 1977).
- Changes in stress patterns when derivational affixes are applied have been attributed to rhythmic alternation (Schane 1979). For example, primary stress on "solid" shifts from the first to the second syllable in "solidity", creating a more optimal pattern of rhythmic alternation.
- Over 90% of formal English poetry follows iambic or trochaic meter (i.e., weak-strong or strong-weak units) (Shapiro & Beum 1965); such rhythmic alternation also dominates the structure of oral "counting-out" poetry used and maintained by English-speaking children (Kelly & Rubin 1988).
- Stress patterns on words change over time so that they maintain rhythmic alternation in the contexts in which they typically appear. For example, English verbs are more likely than nouns to appear in contexts that pressure verbs towards iambic stress, leading to homographs like "permit" and "record" in which the noun has trochaic and the verb iambic stress (Hoffman 2020; Kelly 1989, 1988; Kelly & Bock 1988).
- Word order in conjunctive phrases like "antique and doll" tend to be produced such that rhythmic alternation is preserved (e.g., "doll" is more likely ordered second when conjoined with "antique" than "attic"; Benor & Levy 2006; McDonald et al. 1993).

ISSN: 0027-7738 (print) 1756-2279 (web) Vol. 73 No. 4, Winter 2025 DOI 10.5195/names.2025.2824

With rhythmic alternation in mind, consider the following names of US lakes:

(1) *Glacier Lake*
(2) *Guitar Lake**
(3) *Lake Louise*
(4) *Lake Ellen**

Names like (1) and (3) preserve alternating stress whereas (2) and (4) create a clash between two consecutive stressed syllables. We would therefore expect weak-strong iambic words like "Louise" and "Guitar" to be more common in names beginning than ending in "Lake".

Grammatical Structure: Preference for Modifier + Modified Word Order

In English noun phrases, adjectives typically precede the nouns that they modify. To the extent that place names preserve grammatical structures of English phrases, we would expect "Adjective Lake" word order to be more common than the overall high base rate of "X Lake". Similarly, present participle/gerund verb formations like "whistling" and "rolling" should also appear more often than expected in "X Lake" structures.

Semantic Influences: Prominence

Beisner & Carey (2016) found that the likelihood of "Lake X" word order increased with the surface area of the lake. The names of the Great Lakes—the largest bodies of fresh water on Earth by surface area—provide the most notable example of this pattern. This correlation between word order and lake surface area suggests that the more general concept of *prominence* may affect the choice between "Lake X" and "X Lake", with higher levels of prominence driving greater preference for the "Lake X" word order.

Several sources of evidence suggest that the names of people are psychologically prominent. For example, person names are routinely capitalized in English and related languages (Allerton 1987). In reading text, referents are more accessible in memory if they are denoted via proper names like "Mr. Bloggs" rather than definite descriptions such as "The manager". Specifically, readers more quickly connect an anaphoric pronoun such as "he" to the appropriate referent if previously introduced by a proper name than definite description, indicating that the proper name enhances prominence in memory (Sandford et al. 1988). Other evidence also indicates that readers assume that the central character in a story is more likely denoted by a proper name than common noun, leading to greater accessibility in memory (e.g., McDonald & Shaibe 2002). Given this psychological prominence, we would predict that proper nouns—which generally denote people in place names—should increase the probability of "Lake X" word order.

Going further, to the extent that "X" is more likely to signal prominent semantics in "Lake X" than "X Lake", one would predict that "X" should be more likely to express positive sentiment in the former word order. This prediction is consistent with the distribution of positive and negative sentiment words in US place names. While positive words appear more often than negative words in US place names generally, they especially dominate in names that denote human habitations and associated human-made structures such as churches and bridges (Kelly 2000). This linkage between positive sentiment and human prominence should carry over to higher incidence of positive terms in "Lake X" word order.

Lexical Accessibility

Words with few syllables or high in frequency tend to appear first in conjunctive phrases such as "salt and pepper" and "bread and butter" (e.g., Benor & Levy 2006; Kelly 1986; Cooper & Ross 1975). This phenomenon seems to be due to their greater accessibility from memory (Bock 1982), though it may be driven more by advantages of such words in language comprehension than production (McDonald et al. 1993; Pinker & Birdsong 1979). Specifically, listeners/readers tend to prefer such phrases with higher frequency and shorter words in the first position even though little if any bias is observed in language production. Over time, though, the comprehension bias will tend to differentially preserve structures like "salt and pepper" more than "pepper and salt". If such preferences extend into naming patterns, then the "Lake X" word order should be correlated with the number of syllables in "X" and inversely correlated with the word frequency of "X".

DOI 10.5195/names.2025.2824 Vol. 73 No. 4, Winter 2025 ISSN: 0027-7738 (print) 1756-2279 (web)

Methodology

The United States Geological Survey's (USGS) database of US place names was the basis for the analysis (https://www.usgs.gov/us-board-on-geographic-names/domestic-names). The specific data file, "NationalFile_20210825.txt", downloaded on November 23, 2021, consisted of 2.26 million place names in the 50 US states plus District of Columbia. From this starting point, all place names beginning or ending with the word "lake" were extracted (82,245 records). Note that the sites so named were not restricted to lakes or reservoirs, though 88.45% were in one of these two categories. Using the ntoken function in the Quanteda package in R (version 4.3.0), these records were filtered down to 63,769 that contained two words (i.e., "X Lake" or "Lake X"). The other word in each place name besides "Lake" (henceforth word "X") was assigned a number of syllables using the nsyllable function in the Quanteda package. Thirty-five names for which word "X" could not be assigned syllable number were deleted. In these cases, word "X" contained non-word letter strings (e.g., *Lake JDM*), numbers (e.g., *Lake 30-7*), or combinations (e.g., *Lake H-4*). Stress patterns for disyllabic "X" words were determined from the Carnegie Mellon University Pronouncing Dictionary (data file "cmudict-07b.txt" downloaded on April 19, 2021, from http://www.speech.cs.cmu.edu/cgi-bin/cmudict). Words that matched to the CMU dictionary were classified as "trochaic" if they had "12" or "10" stress and as iambic if they had "21", "01", or "11" stress, where "1" corresponds with primary stress, "2" as secondary stress, and "0" as unstressed.[1] 26,834 records matched to the CMU dictionary. This figure represented 85.22% of the total 31,489 disyllabic words. Disyllabic words without matching stress patterns were retained for analysis given values on other predictive variables.

Word frequency and grammatical uses of each "X" word were based on Francis & Kucera (1982). The cumulative frequency of each word across its grammatical uses was the measure of overall frequency. The proportions of proper noun, adjective, and present participle/gerund uses were calculated as (# occurrences in specific grammatical use)/(overall frequency). Place names were excluded from the analysis if word "X" did not appear in Francis & Kucera (1982), providing no estimates for overall frequency or any of the grammatical usage variables. 24,810 place names were dropped based on this criterion (38.93%) leaving 38,924 place names in the final database.[2] The "X" words were then matched to the Hu & Lui (2004) sentiment lexicon, which consists of 6,786 words coded for as having positive or negative sentiment. 2,127 words matched to the lake name database (5.46% of the total), 1137 with positive sentiment (e.g., "love", "victory", "bright") and 990 with negative sentiment (e.g., "lonesome", "dead", "poison"). Words with positive sentiment were coded as 1, negative sentiment as -1, and all other words as 0.

The primary analysis consisted of a mixed-effects logistic regression conducted using the glmer function in the lme4 R package (Bates et al. 2018; version 1.1-37). The regression predicted whether the place name order was "X Lake" (coded as 0) or "Lake X" (coded as 1) from the following eight parameters:

- Whether word X had trochaic stress (coded as "1" if the word was disyllabic with trochaic stress, otherwise 0)
- Whether word X had iambic stress (coded as "1" if the word was disyllabic with iambic stress, otherwise 0)
- The proportion of occasions in Francis & Kucera (1982) in which word X was used as an adjective
- The proportion of occasions in Francis & Kucera (1982) in which word X was used as a present participle or gerund
- The proportion of occasions in Francis & Kucera (1982) in which word X was used as a proper noun
- Sentiment of word X (coded as -1 if negative, 1 if positive, and 0 if the word did not match to the sentiment database)
- Number of syllables in word X
- Natural log word frequency of word X

Random intercepts for individual words were included in the regression to ensure that significant fixed effects generalize across words. Additionally, variance inflation factors were calculated to assess for risk of collinearity among the predictor variables (see Fox & Monette 1992; specific calculations conducted using the vif function in Fox & Weinberg's [2019] car package in R, version 3.1-3).

ISSN: 0027-7738 (print) 1756-2279 (web) Vol. 73 No. 4, Winter 2025 DOI 10.5195/names.2025.2824

Results

Overall, 89.46% of place names containing "Lake" plus a disyllabic word have the order "X Lake". However, as table 1 shows, the likelihood of the rarer "Lake X" structure rises or falls further depending on prosodic, grammatical, semantic, and other lexical characteristics of word X. All of the predictors were statistically significant except for the proportion of present participle/gerund uses of word X. Positive coefficients in the regression analysis indicate predictors associated with higher uses of the "Lake X" word order while negative coefficients indicate predictors associated with lower uses of "Lake X". Variance inflation factors are less than 2 for all variables, indicating that multicollinearity is not an issue in interpreting results.

Table 1: Logistic Regression Results Predicting Likelihood of "Lake X" versus "X Lake" Word Order in US Place Names

Parameter	Coefficient	Max Est. Change in Prob "Lake X" Order[3]	St. Error	z-Value	p-Value
Intercept	-5.11	NA	0.18	-29.03	< .00001
Trochaic Stress on Disyllabic Words	-0.18	-0.05	0.09	-2.05	< .05
Iambic Stress on Disyllabic Words	0.97	0.24	0.23	4.24	< .00003
Proportion Adjective Use	-0.69	-0.17	0.22	-3.07	< .005
Proportion Participle/ Gerund Use	-1.20	-0.30	0.90	-1.33	NS
Proportion Proper Noun Use	1.47	0.37	0.10	14.25	< .00001
Sentiment	0.71	0.18	0.18	3.99	< .0001
Number of Syllables	0.57	0.14	0.06	9.62	< .00001
Word Frequency	0.15	0.04	0.03	4.76	< .00001

Since the logistic regression function is non-linear, the estimated change in the likelihood that a word X with particular characteristics appears in a "Lake X" rather than "X Lake" structure is not constant. However, the maximum change in probability can be estimated as coefficient/4, corresponding to the steepest part of the curve (Gelman et al. 2021). For example, a disyllabic word with iambic stress is estimated to increase the probability of "Lake X" word order by at most 24.34%. In contrast, a disyllabic word with trochaic stress is estimated to decrease the probability of "Lake X" word order by at most 4.60%.

Impact of Lexical Prosody

As predicted from the preference for rhythmic alternation, the "Lake X" order is significantly more likely when word X has iambic rather than trochaic stress. Specifically, the incidence of "Lake X" rises from 10.64% when X has trochaic stress to 34.06% when X has iambic stress, corresponding to a substantial 3.2x increase (table 2). As shown in the regression results, iambic stress significantly increases the incidence of "Lake X" word order above the 11.47% baseline for disyllabic words whereas trochaic stress significantly lowers the incidence, albeit less dramatically.

Table 2: Percent of US Place Names with "X Lake" or "Lake X" Depending on Stress of Disyllabic Word "X"

Word Order	Trochaic Stress	Iambic Stress
X Lake	89.36	65.94
Lake X	10.64	34.06

DOI 10.5195/names.2025.2824 Vol. 73 No. 4, Winter 2025 ISSN: 0027-7738 (print) 1756-2279 (web)

Impact of Grammatical Uses of Word X

Table 3 shows the relationship between grammatical uses of a word and the incidence of "Lake X" word order. Specifically, the percent of "Lake X" order is shown depending on whether usage in a particular grammatical class is less than 50% or at least 50%. Consistent with predictions, the more often word X is used as an adjective, the less likely is the "Lake X" order. Participle/gerund uses of X are also directionally associated with lower incidence of "Lake X" as predicted, but the pattern is not significant. However, the lack of significance is likely due to the small sample size of names containing present participles/gerunds. Only 98 names contained a word X with any present participle/gerund use. In contrast, 6,582 names contained a word with adjective uses.

Table 3: Percent of "Lake X" Word Order Depending on Whether the Percentage of Uses of Word X in a Particular Grammatical Category is Less than 50% or 50%+

Grammatical Usage	Percentage Use in Grammatical Class < 50%	Percentage Use in Grammatical Class 50% or Higher
Adjective	11.43	3.21
Present Participle/Gerund	10.53	3.26
Proper Noun	6.67	16.42

Impact of Semantics

Consistent with predictions, the more often word X is used as a proper noun, the more likely is the "Lake X" order. As shown in table 2, as the incidence of Word X as a proper noun rises from less than 50% to 50% or more, the incidence of "Lake X" word order more than doubles, from 6.67% to 16.42%. This effect is not due to a more general characteristic of nouns. If anything, common nouns show the opposite pattern: the more often a word is used as a common noun relative to other grammatical categories, the less likely it is to appear in "Lake X" place names. For instance, the incidence of "Lake X" is 12.32% among words used less than 50% of the time as common nouns, but only 8.02% among words used as common nouns 50% of the time or higher.[4]

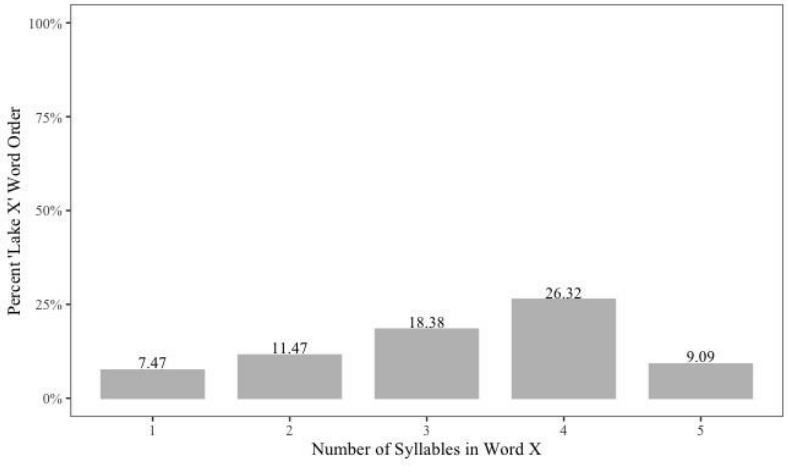
Consistent with predictions, positive sentiment is more strongly associated with "Lake X" word order (11.96%) than negative sentiment (9.19%). Though small in absolute terms, the change represents a 30.14% increase in "Lake X" word order among "X" words with positive sentiment compared to those with negative sentiment. Words that did not match the Hu & Liu sentiment dictionary appeared in "Lake X" structures 10.51% of the team, or intermediate between words with negative and those with positive sentiment.

Other Lexical Characteristics: Number of Syllables and Word Frequency

As predicted, the number of syllables in word X is directly associated with "Lake X" word order. Specifically, figure 1 shows that as the number of syllables increases from one to four, the incidence of "Lake X" rises from 7.47% to 26.32%, or almost a fourfold increase. The "Lake X" percentage drops sharply when word X contains five syllables, but only 22 five-syllable words appeared in the database. In contrast, four-syllable words—the category with the lowest sample size after five-syllable words—consisted of 456 cases, or more than 20 times the sample size of five-syllable words. Results for the latter should therefore be treated cautiously, while the patterns for words with one through four syllables shows a strong relationship with "Lake X" incidence, as confirmed by the significant effect of syllable number in the regression analysis.

ISSN: 0027-7738 (print) 1756-2279 (web) Vol. 73 No. 4, Winter 2025 DOI 10.5195/names.2025.2824

Figure 1. Percent of US Place Names with 'Lake X' Rather than 'X Lake' Word Order Depending on Number of Syllables in Word X



Word frequency is also significantly associated with "Lake X" word order, but in the opposite direction from that predicted. As the frequency of X increases, so does the likelihood of "Lake X" word order. This pattern conflicts with the lexical accessibility hypothesis in that the greater accessibility of higher frequency words should lead them to be favored in name-initial position but instead they are more likely in name-final position. The Discussion will consider this pattern further.

Discussion

This investigation identified several factors that influence the choice between "X Lake" and "Lake X" word order among US place names. Prosodically, the choice between these alternatives tends to preserve an alternation between strongly and weakly stressed syllables. Grammatically, the more often a word is used as an adjective in American English, the more often "X Lake" word order occurs, preserving the standard adjective-noun word order in English noun phrases. Semantically, proper names and words with positive sentiment are favored in "Lake X" word order, suggesting that "Lake X" is associated with higher psychological prominence. Finally, words with more syllables and higher frequency also predict higher "Lake X" usage. Each of these findings, except for the final one between higher word frequency and greater incidence of "Lake X", aligned with predictions. The use of a regression analysis in evaluating the hypotheses provides the ability to examine each potential predictor independently controlling for the others.

The remainder of this discussion will examine three topics: how to characterize the "euphony" of place names, how to account for the unanticipated relationship between higher word frequency and greater likelihood of "Lake X" word order, and what are further considerations of the concept of psychological "prominence".

Characterizing Name "Euphony"

One general factor proposed to influence place name creation is a desire for "euphony" as in the following examples quoted from Stewart's seminal *American Place-Names*:

> **Indianola.** "A repeated habitation name coined from 'Indian' and a Latin or Latin-like ending, chiefly used for its euphony" (1970, 220).
>
> **Linden**. "A habitation name in as many as 20 states [. . .] the euphonious sound has helped to popularize it" (1970, 258).
>
> **Ladoga.** "From the largest European lake; used in several states, probably aided by its euphony" (1970, 247).

DOI 10.5195/names.2025.2824 Vol. 73 No. 4, Winter 2025 ISSN: 0027-7738 (print) 1756-2279 (web)

These comments, however, leave unanswered the specific factors that cause a name to sound more or less "euphonious". Such factors need to be described in sufficient detail so that we can devise tests to discern whether they in fact influence the structure of place names or other names.

This paper takes a step in this direction by testing the impact of one potential factor in driving name euphony: a preference for alternations between strong and weak syllables. The fact that "Lake X" word order is more likely when "X" is a disyllabic iambic than disyllabic trochaic word is consistent with this preference as names like *Lake Louise* and *Lucy Lake* conform with an alternating rhythm, whereas their reversals, *Louise Lake* and *Lake Lucy,* do not.

Several linguistic characteristics of English and psychological characteristics of language use suggest other factors besides rhythmic alternation that could affect name euphony and, ultimately, the frequency of different name structures. I describe two of these factors here and their hypothetical effects to illustrate further avenues of investigation.

First, a wide variety of evidence indicates that speakers learn statistical properties of words in lexical categories, expressing this implicit knowledge in experimental and naturalistic contexts (e.g., Treiman et al. 2020; Fitneva et al. 2009; Kelly 1992). For example, name gender is associated with various phonological characteristics in English, such as female names being more likely to have iambic stress (e.g., *Danielle* versus *Daniel*) and end in vowels (e.g., *Alexandra* vs. *Alexander*) than male names. English speakers classify names more quickly and accurately if they have gender-typical phonological properties. And, over time in US culture, unisex names are more likely to shift toward solely female or solely male usage to the extent that they sound consistent with the respective category (Cassidy et al. 1999).

With such findings in mind, consider the place name *Penn Yan, NY*. This name was formed by blending "Pennsylvania" and "Yankee" (Stewart 1970). There are two components to the process of creating such name blends: (1) determining which portions of each name to include, such as "Penn" and "Yan" in this case, and (2) selecting the order in which the portions appear. Focusing on this second component, "Penn" and "Yan" could be arranged as "Penn Yan" like the actual name or "Yan Penn". Can we predict which order is likely to be favored? Analyses of the USGS place name database show that the orthographic sequence "penn" is 3.2 times more likely to appear in word initial than word final position. "Yan", on the other hand, is 2.8 times more likely to appear in word final than initial position. If English speakers in the US have learned these patterns, then blends like "Penn Yan" should have higher euphony than those like "Yan Penn" since they sound more like other place names (or English words more generally). Indeed, in a corpus of 227 place name blends culled from Stewart (1970), I found that 70% are structured like "Penn Yan" such that the first part of the blend is more likely than the second part to appear in word initial position.

Second, perceived euphony may depend on context, such as the type of place being named. Consider lakes versus rivers. Both refer to geographical features pertaining to water, but lakes are more circumscribed whereas rivers are often long and meandering. When coupled with *sound symbolism*—or correlations between phonological and semantic properties of words (e.g., Sidhu & Pexman 2018; Hinton et al. 1994)—these physical differences between lakes and rivers suggest pressures toward phonological differences in their names. Given their greater extent, rivers should tend to have longer names than lakes, such as the *Mississippi* and *Allegheny Rivers* versus *Crater* and *Seneca Lakes*.[5] Initial tests of this prediction support this hypothesis, as river names contain more characters (6.6 versus 6.1, $t(4679.8)=13.54$, $p < .0001$) and syllables (2.2 versus 1.9, $t(4616.2)=19.78$, $p < .0001$) on average than lake names (excluding the words "River" and "Lake" from the names of course).[6]

Along with the impact of rhythmic alternation on the choice between "X Lake" versus "Lake X" word order, these examples demonstrate how several linguistic and cognitive factors may contribute to name "euphony", leading to testable predictions about the statistical distribution of place name characteristics.

Explaining the Effects of Number of Syllables and Word Frequency

Similar to word order patterns in conjunctive phrases like "milk and honey" and "kit and kaboodle" in which the word with fewer syllables appears before the word with more syllables, "Lake X" word order increases with the number of syllables in "X". This pattern is consistent with the greater accessibility in memory of words with fewer syllables, leading to increased likelihood of appearing before rather than after "Lake".

Higher frequency words are also more accessible from memory than lower frequency words, so we would expect to see a similar relationship between word order and frequency in the statistical distribution of the "X Lake" and "Lake X" alternatives. However, the opposite turns out to be the case: As word frequency increases, so does the likelihood of "Lake X" word order. Ease of retrieval from memory cannot account for this pattern.

A potential alternative explanation could draw on the well-established positive relationship between frequency of exposure and attraction (see Montoya et al. 2017;Bornstein 1989 for summaries and meta-analyses). In a series of seminal studies, Zajonc and colleagues manipulated how often observers encountered

ISSN: 0027-7738 (print) 1756-2279 (web) Vol. 73 No. 4, Winter 2025 DOI 10.5195/names.2025.2824

different stimuli (e.g., words, people) and assessed the impact of exposure frequency on subsequent attraction. For example, the more often participants in an experiment see a particular pseudoword or symbol, the more likely they are to subsequently rate the pseudoword or symbol as having a positive meaning (Zajonc 1968). And the more often a person "accidentally" encounters another person as part of an experiment (with the frequency of encounters surreptitiously manipulated by the experimenters), the more positively they rate that person later (Saegert et al. 1973).

When extended to the English lexicon, this "mere exposure effect" suggests that higher frequency words will tend to be viewed more positively than lower frequency words independently of their meaning. We know from the relationship between word sentiment and word order that "Lake X" is more likely when "X" has intrinsically positive (e.g., "pleasant") rather than negative (e.g., "poison") meaning. When extended to the mere exposure effect, this pattern would predict that higher frequency words will also tend to appear in the "Lake X" structure more than lower frequency words due to the positive aura they would have accrued due to their higher frequency. Of course, further research will be needed to test this possibility.

Further Considerations of Psychological Prominence

Several patterns suggest that the notion of psychological "prominence" impacts the choice between "X Lake" and "Lake X" word order. First, "Lake X" is more common as lake surface area increases (Beisner & Carey 2016). Second, the research presented here shows that "Lake X" is more common when "X" is a proper name. Third, this research also finds that "Lake X" is more common when "X" has positive rather than negative sentiment.

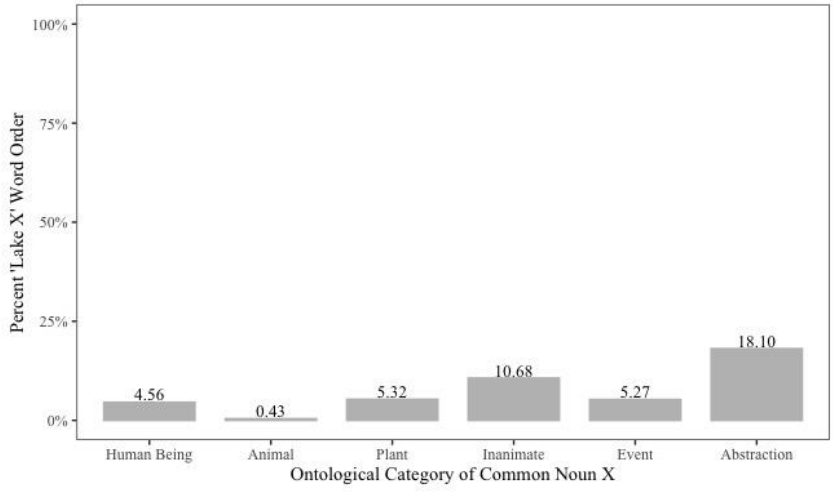
Despite these findings, the specific characteristics that contribute to psychological prominence remain unclear. The patterns above suggest broadly that physical size, humanness, and sentiment valence (positive vs. negative) are associated with prominence. However, further research will be needed to establish that these themes impact place names broadly and discern limitations that may help to define psychological "prominence" more precisely. For example, in terms of breadth of application, consider the choice between "X Mountain" and "Mount X". The elevation of the geographical feature "summit" in the USGS database is significantly higher for "Mount X" (mean: 1,743 meters) versus "X Mountain" (mean: 1,277 meters) place names
($t(3039.9) = 18.06$, $p < .0001$). This pattern indicates that physical size—defined as surface area for lakes and elevation for mountains—impacts the structure of US place names at least somewhat broadly.

In terms of defining psychological prominence more precisely, consider the finding that "Lake X" structures are more common when "X" is a proper name like *George* or *Wallace*. This pattern suggests that human names are associated with higher psychological prominence. But is this relatively specific conclusion warranted? Perhaps common nouns that denote human beings like "queen", "pilgrim", or "pioneer" are also more likely than other common nouns to be associated with "Lake X" word order. Such a finding would suggest that "humanness" is prominent broadly and not specific to proper names of individual persons.

To test this possibility, the X words in "Lake X" or "X Lake" place names were filtered to those used at least 75% of the time as nouns in Francis & Kucera (1982) and that never appeared as proper nouns. These were classified into the following broad categories: human beings, animals (e.g., coyote, eagle), plants (e.g., blueberry, rosebush), inanimate objects (e.g., candle, glacier), events (e.g., earthquake, sunrise), and abstractions (e.g., peace, solitude), corresponding to high-level "ontological" categories investigated by philosophers (e.g., Sommers 1959, 1965) and psychologists (e.g., Keil 1979, 1981). Figure 2 shows the percent of "Lake X" (as opposed to "X Lake") US placenames for each ontological category. Unlike proper names, common nouns denoting human beings do not show especially high rates of appearing in "Lake X" structures, suggesting that proper names are specifically prominent psychologically, as other psycholinguistic research cited earlier suggests, rather than nouns denoting human beings broadly.

DOI 10.5195/names.2025.2824 Vol. 73 No. 4, Winter 2025 ISSN: 0027-7738 (print) 1756-2279 (web)

Figure 2. Percent of US Place Names with 'Lake X' Rather than 'X Lake' Word Order Depending on Ontological Category of Common Noun X



Two patterns in figure 2 are notable, however. First, common nouns denoting abstractions are most likely to appear in "Lake X" structure, suggesting that idealizations represented in words such as "comfort", "defiance", "omen", and "refuge" are psychologically prominent. Second, nouns denoting animals are almost never included in "Lake X" structures. This pattern is consistent with the anthropological literature describing a psychological ambivalence toward animals. On the one hand, they have long been an integral part of our lives, first as prey or predators, and later, after domestication, as a source of food, clothing, work, transportation, and companionship. This close relationship can elevate the prominence of animals, as shown in the totemic role of animals in many cultures (e.g., Frazer 1951) including our own where sports teams are commonly named after animals. On the other hand, "their" obvious similarity to "us" creates conflict and denial,[8] arousing an urge to suppress the psychological prominence of animals lest our "humanness" be contaminated (e.g., Rozin & Nemeroff 1990; see Pinker 2012, for summary). The rare appearance of animal references in "Lake X" place names may be a manifestation of such suppression.

Patterns like these suggest that further investigations of the psychological concept of prominence and its impact on place name patterns will be fruitful. For example, the greater association of "Lake X" and, perhaps, "Mount X" with psychological prominence may lead to increased preference for these forms over "X Lake" and "X Mountain" for sites that are correspondingly prominent, such as the names of places closely linked to human activities—towns, bridges, churches—compared with names of natural features.[9] Similarly, an urge to separate human from animal may drive lower use of words denoting animals in names for sites closely associated with people. Zelinsky (1994), for example, found that names of cemeteries in the United States rarely mention animals. But we may expect relatively few references to animals in sites associated with people more generally.

Conclusion

The choice between "X Lake" and "Lake X" word order in US place names is far from random, as several prosodic, grammatical, semantic, and other lexical characteristics combine to influence the choice. The findings suggest several lines of further research to address limitations of the current study and extend its scope. For example, this study focused on a single alternation in one geographic setting: the choice between "X Lake" and "Lake X" in US place names. To evaluate the generality of the findings, one could test whether the same predictors influence word order in, for instance, Canadian lake names or shed light on other place name choices, such as that between "X Mountain" and "Mount X". For example, the preference for rhythmic alternation would favor "Mount X" for disyllabic "X" words with iambic stress. Additional onomastic research—and psychological research more generally—is also needed to better refine the concept of psychological "prominence" and its impact on naming patterns.

ISSN: 0027-7738 (print) 1756-2279 (web) Vol. 73 No. 4, Winter 2025 DOI 10.5195/names.2025.2824

Notes

[1] Note that to facilitate processing stress pattern as a character variable, the "0" code for an unstressed syllable in the original CMU database was replaced with "x" for the analyses reported in this article. Additionally, only 71 disyllabic words have "11" stress in the CMU dictionary, but these were classified as iambic based on manual review (e.g., fourteen, pawnee, halfway). However, results are unchanged if they are not classified as iambic or trochaic.

[2] A regression model that included all words , forcing the total frequency and grammatical class values to 0 for words not listed in Francis & Kucera (1982), failed to converge due to the large number of 0 values. This situation reinforces the decision to drop such words from the analysis.

[3] See text for explanation of this measure.

[4] Common noun and proper noun usage of words are highly negatively correlated (-.70). Consequently, proportion common and proper noun usage cannot both be included in the regression analysis. Variable inflation factors for these variables are greater than four if both are included, indicating unacceptably high levels of multicollinearity.

[5] Walt Whitman captured this hypothesis poetically when he wrote: "Mississippi, the word winds with chutes—it rolls a stream three thousand miles long".

[6] This analysis was restricted to lake and river names that contained only one word in addition to "lake" or "river" thereby excluding potential artifacts from names such as *West Fork Choctawhatchee River*. River names also contain more characters and syllables if all names are included in the analysis.

[7] A regression controlling for number of syllables, word frequency, stress, and sentiment confirmed that abstract nouns are significantly more associated with "Lake X" word order, animal nouns significantly less associated, with human nouns showing no impact on the likelihood of "Lake X" word order.

[8] For example, Descartes, who, as described by Pinker "wrote that animals were clockwork, so there was no one home to feel pain or pleasure. What sounds to us like cries of distress were merely the output of a noisemaker, like a warning buzzer on a machine" (2012, 682).

[9] "Lake X" occurs more often than "X Lake" as a name for populated places than other sites (35.91% vs. 9.83%). Similarly, but even more dramatically, "Mount X" occurs more often than "X Mountain" for populated places than other sites (82.39% vs. 12.61%).

AI Disclosure Statement

No AI tools or technology were used to conduct the research or write this article.

References

Allerton, David J. 1987. "The Linguistic and Sociolinguistic Status of Proper Names". *Journal of Pragmatics* 11, no. 1: 61–92.

Bates, Douglas, Martin Maechler, Ben Bolker, and Steve Walker. 2018. "Fitting Linear Mixed-Effects Models Using lme4". *Journal of Statistical Software* 67, no. 1: 1–48.

Benor, Sarah, and Roger Levy. 2006. "The Chicken or the Egg? A Probabilistic Analysis of English Binomials". *Language* 82, no. 2: 233–278.

Beisner, Beatrix E., and Cayelan C. Carey. 2016. "Lake Name or Name Lake? The Etymology of Lake Nomenclature in the United States". *Freshwater Biology* 61, no. 9: 1601–1609.

Bock, J. Kathryn. 1982. "Toward a Cognitive Psychology of Syntax: Information Processing Contributions to Sentence Formulation". *Psychological Review* 89, no. 1: 1–47.

Cassidy, Kimberly Wright, Michael H. Kelly, and Lee'at J. Sharoni. 1999. "Inferring Gender from Name Phonology". *Journal of Experimental Psychology: General* 128, no. 3: 362–381.

Cooper, William E., and John Robert Ross. 1975. "World Order". *Papers from the Parasession on Functionalism*. Edited by Robin E. Grossman, L. James San, and Timothy J. Vance, 63–111. Chicago: University of Chicago Press.

Fitneva, Stanka A., Morten H. Christiansen, and Padraic Monaghan. 2009. "From Sound to Syntax: Phonological Constraints on Children's Lexical Categorization of New Words". *Journal of Child Language* 36, no. 5: 967–997.

DOI 10.5195/names.2025.2824 Vol. 73 No. 4, Winter 2025 ISSN: 0027-7738 (print) 1756-2279 (web)

Fox, John, and Georges Monette. 1992. "Generalized Collinearity Diagnostics". *Journal of the American Statistical Association* 87, no. 417: 178–183.

Fox, John, and Sanford Weisberg. 2019. *An R Companion to Applied Regression*, 4th ed. Thousand Oaks CA: Sage.

Francis, W. Nelson, and Henry Kucera. 1982. *Frequency Analysis of English Usage: Lexicon and Grammar*. Boston: Houghton Mifflin.

Frazer, James George. 1951. *The Golden Bough: A Study in Magic and Religion*. New York: Macmillan.

Gelman, Andrew, Jennifer Hill, and Aki Vehtari. 2021. *Regression and Other Stories*. Cambridge: Cambridge University Press.

Halle, Morris, and Jean-Roger Vergnaud. 1987. *An Essay on Stress*. Cambridge, MA: MIT Press.

Hinton, Leanne, Joanna Nichols, and John Ohala. 1994. "Sound-Symbolic Processes". *Sound Symbolism*. Edited by Leanne Hinton, Joanna Nichols, and John Ohala, 1–14. Cambridge, UK: Cambridge University Press.

Hofmann, Klaus. 2020. "Stress in Real Time: The Noun–Verb Stress Contrast and the Rhythmic Context Hypothesis in the History of English". *Journal of Historical Linguistics* 10, no. 3: 452–486.

Hu, Minqing, and Bing Liu. 2004. "Mining and Summarizing Customer Reviews". *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 168–177. n.p.

Keil, Frank C. 1979. *Semantic and Conceptual Development: An Ontological Perspective*. Cambridge, MA: Harvard University Press.

Keil, Frank C. 1981. "Constraints on Knowledge and Cognitive Development". *Psychological Review* 88, no. 3: 197–227.

Kelly, M. H. 1986. "On the Selection of Linguistic Options". Unpublished doctoral dissertation, Ithaca, NY: Cornell University.

Kelly, Michael H. 1988. "Rhythmic Alternation and Lexical Stress Differences in English". *Cognition* 30, no. 2: 107–137.

Kelly, Michael H. 1989. "Rhythm and Language Change in English". *Journal of Memory and Language* 28, no. 6: 690–710.

Kelly, Michael H. 1992. "Using Sound to Solve Syntactic Problems: The Role of Phonology in Grammatical Category Assignments". *Psychological Review* 99, no. 2: 349–364.

Kelly, Michael H. 2000. "Naming on the Bright Side of Life". *Names* 48, no. 1: 3–26.

Kelly, Michael H., and J. Kathyrn Bock. 1988. "Stress in Time". *Journal of Experimental Psychology: Human Perception and Performance* 14, no. 3: 389–403.

Kelly, Michael H., and David C. Rubin. 1988. "Natural Rhythmic Patterns in English Verse: Evidence from Child Counting-out Rhymes". *Journal of Memory and Language* 27, no. 6: 718–740.

Liberman, Mark, and Alan Prince. 1977. "On Stress and Linguistic Rhythm". *Linguistic Inquiry* 8, no. 2: 249–336.

McDonald, Janet L., Kathryn Bock, and Michael H. Kelly. 1993. "Word and World Order: Semantic, Phonological, and Metrical Determinants of Serial Position". *Cognitive Psychology* 25, no. 2: 188–230.

McDonald, Janet L., and Deborah M. Shaibe. 2002. "The Accessibility of Characters in Single Sentences: Proper Names, Common Nouns, and First Mention". *Psychonomic Bulletin & Review* 9, no. 2: 356–361.

Montoya, R. Matthew, Robert S. Horton, Jack L. Vevea, Martyna Citkowicz, and Elissa A. Lauber. 2017. "A Re-Examination of the Mere Exposure Effect: The Influence of Repeated Exposure on Recognition, Familiarity, and Liking". *Psychological Bulletin* 143, no. 5: 459–498.

Nespor, Marina, and Irene Vogel. 1986. *Prosodic Phonology*. Dordrecht: Foris Publications.

Pinker, Steven. 2012. *The Better Angels of Our Nature*. New York: Penguin Books.

Pinker, Steven, and David Birdsong. 1979. "Speakers' Sensitivity to Rules of Frozen Word Order". *Journal of Verbal Learning and Verbal Behavior* 18, no. 4: 497–508.

Rozin, Paul, and Carol Nemeroff. 1990. "The Laws of Sympathetic Magic: A Psychological Analysis of Similarity and Contagion". *Cultural Psychology: Essays on Comparative Human Development*. Edited by James W. Stigler, Richard A. Schweder, and Gilbert Herdt, 205–232. Cambridge: Cambridge University Press.

Sanford, Anthony J., Kanwalraj Moar, and Simon C. Garrod. 1988. "Proper Names as Controllers of Discourse Focus". *Language and Speech* 31, no. 1: 43–56.

ISSN: 0027-7738 (print) 1756-2279 (web) Vol. 73 No. 4, Winter 2025 DOI 10.5195/names.2025.2824

Schane, Sanford. A. 1979. "The Rhythmic Nature of English Word Accentuation". *Language* 55, no. 3: 559–602.

Selkirk, Elizabeth O. 1984. *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, MA: MIT Press.

Shapiro, Karl, and Robert Beum. 1965. *A Prosody Handbook*. New York: Harper and Row.

Sidhu, David M., and Penny M. Pexman. 2018. "Five Mechanisms of Sound Symbolic Association". *Psychonomic Bulletin & Review* 25, no. 5: 1619–1643.

Sommers, Fred. 1959. "The Ordinary Language Tree". *Mind* 68, no. 270: 160–185.

Sommers, Fred. 1965. "Predictability". *Philosophy in America*. Edited by Max Black, 262–281. Ithaca, NY: Cornell University Press.

Stewart, G. R. 1970. *American Place-Names*. New York: Oxford University Press.

Treiman, Rebecca, Nicole Rosales, Lauren Cusner, and Brett Kessler. 2020. "Cues to Stress in English Spelling". *Journal of Memory and Language* 112: 104089.

Zajonc, Robert B. 1968. "Attitudinal Effects of Mere Exposure." *Journal of Personality and Social Psychology Monograph Supplement*, no. 9: 1-27.

Zelinsky, Wilbur. 1994. *Exploring the Beloved Country: Geographic Forays into American Society and Culture*. Iowa City: University of Iowa Press.

Notes on the Contributor:

Michael Kelly, PhD, recently retired from NAXION, a marketing research and business consulting firm based in Philadelphia, PA, USA. As Vice President at NAXION, he led research design and data analytics teams to help clients in diverse industries make informed business decisions. Before joining NAXION, he was Associate Professor of Psychology at the University of Pennsylvania where he conducted research in linguistics and psycholinguistics (e.g., phonology, word formation, language acquisition, language change, creative uses of language). In onomastics specifically, he has worked on cultural, semantic, and phonological influences on naming patterns.

Correspondence to: Michael Kelly. Email: kellymh@gmail.com

DOI 10.5195/names.2025.2824 Vol. 73 No. 4, Winter 2025 ISSN: 0027-7738 (print) 1756-2279 (web)