Used or discussed? Lake Ontario and Lake America in YouTube comments

Separating name use from name talk in a frozen first snapshot of pre-existing comment sections

Max Sedlmair, independent researcher, Germany · ORCID 0009-0007-2471-5595
Version 1.0 — First Snapshot preprint, 11 October 2026; editorially revised 16 October 2026.
Second observation point fixed in advance for 22 December 2026 and not yet taken.

ABSTRACT

A newly imposed place name can become frequent in public text simply because people are arguing about it. Counting how often a name appears therefore cannot separate adoption from controversy. The question is not only how often a new name occurs, but whether it is being used to refer to the lake or talked about as a name — and when, if at all, a first ordinary use appears. This rapid-response study was designed to capture that distinction at the earliest observable stage after the 27 August 2026 U.S. federal renaming of Lake Ontario as *Lake America*. A contemporary corpus of 121,161 YouTube comments from 36 videos, from which separate development and held-back validation samples were drawn, was used to build and test a coding distinction between ordinary referential use and talk about the name itself. That measure was then applied to a separate, prospectively frozen panel of 35 pre-existing Lake Ontario videos selected entirely on pre-intervention criteria.

In 478 comments before the intervention, exact *Lake America* did not occur. In the first post-intervention window it appeared in 11 of 162 comments across four videos: in ten of them as an object of discussion, correction, reporting, evaluation or humour, and in one as an ordinary reference to the lake. That one case is the first ordinary use of the new name observed anywhere in the panel. It states no position on the renaming, and it occurred in the one video where the name was not otherwise discussed — a comment section in which the new label turned up only as a label. The held-back validation sample showed the same lopsidedness: of 55 exact *Lake America* occurrences, 49 were talk about the name, 2 were ordinary reference, and 4 were unclear. On a frozen subset of that validation material, two coders working independently reached $\kappa = 0.776$ on the occurrence-level distinction; the 35-video panel itself was coded by the author alone.

These first-window counts are small and clustered and do not establish adoption, rejection or the success of the renaming. For now, *Lake America* is more visible as something people talk about than as something they simply call the lake. A single referential case cannot show more than that it has happened once: at this range, a name that will spread and a name that will not look alike. Distinguishing them is what a second fixed observation is for. A second observation of the same 35 videos was fixed in advance for 22 December 2026 and will be published as the next version of this study; it is the first point at which ordinary use spreading to further videos, remaining a single case, or disappearing altogether become distinguishable outcomes.

Keywords: toponymy; renaming; metapragmatics; reference; use–mention distinction; YouTube comments; prospectively frozen observation; corpus linguistics

1. Introduction

Official place-name changes are political acts, but a declaration does not settle how a name will be used. Research on contested toponymy has established that older names persist after formal renaming, that persistence may reflect habit rather than deliberate resistance, and that naming disputes recruit arguments about authority, legitimacy, history, identity and cost. Online spaces add a layer that matters for measurement: a name can appear in text because it is being corrected, reported, ridiculed or disputed, rather than because anybody is using it to refer to a place.

That layer creates a specific methodological risk. If a new name is politically contested, its raw frequency in public text can rise for two different reasons at once — because some people start using it, and because others start arguing about it. A count that does not separate those functions can mix controversy with ordinary use. The distinction itself is old and is not claimed here as a discovery: Pires (2007) already separated an explicitly metalinguistic frame from

straightforward topographic use in naturally occurring online tokens of a popular toponym, and Benton (2026), tracking *Gulf of Mexico* against *Gulf of America* in news media, excluded stories containing both labels on the grounds that they were mostly about the controversy rather than instances of use.

The August 2026 U.S. federal renaming of Lake Ontario as *Lake America* provides a temporally bounded case in which a newly introduced label can be followed close to the intervention. In the frozen legacy corpus used here, no exact *Lake America* occurrence was observed in the pre-window, so the label first becomes observable in these comment spaces after the directive. What it is doing when it does is the question this study asks, and it refuses the broader one. It does not ask whether the renaming was accepted, and it does not treat occurrence of the string *Lake America* as evidence of adoption. It asks: when the new name appears in naturally occurring comments, is it being used to refer to the lake, or is the name itself the topic?

Two comment samples answer it. Search 1 is a contemporary corpus, used to develop the coding scheme and to supply a held-back validation sample. Search 2 is a held-out legacy corpus: videos that existed before the renaming, concerned the lake in an everyday way, and were selected entirely on pre-intervention behaviour, with their comment texts sealed until the coding scheme, retrieval definition, analysis windows and denominators were frozen.

This manuscript is the first of two reports from a single design. It covers one 25-day window that opens on the day the renaming was ordered, and a window that short can show what a new name is doing as it first appears, but not what becomes of it. A second observation of the same 35 videos, under the same retrieval definition and the same coding manual, was fixed before any held-out result was known and is scheduled for **22 December 2026**. That observation is where the question of whether the new label settles into ordinary use can first be addressed with more than a single window, and it will be released as a further version of this record rather than as a separate study. Readers who want the longitudinal picture should look for that version; what follows establishes the baseline against which it will be read.

Two limits on the December observation are worth stating now rather than discovering later. It will still be an observation and not an experiment, and the United States federal general election of 3 November 2026 falls between the two points, so any change will be read under altered political conditions rather than as a clean measure of conventionalisation.

A note on neutrality. This study takes no position on which name is correct, preferable or legitimate; both are treated as strings to be classified. No political preference, nationality, ideology or approval is inferred from anybody's choice of name. Referential use is linguistic behaviour, not endorsement, and this rule was fixed in the coding manual before any outcome was observed.

2. Background and closest prior work

The broad phenomenon is not new, and the literature was searched adversarially in order to establish precisely that. The following works are the closest located, and each is described in terms of what it already does.

Pires (2007) supplemented a questionnaire on the unofficial Birmingham toponym *Pigeon Park* with a search for naturally occurring written tokens, and distinguished tokens embedded in an "explicitly toponymical, and thus metalinguistic frame" from tokens belonging to straightforward topographical discourse. The functional distinction central to the present study is therefore established prior art, not an innovation.

Light and Young (2014) surveyed users of a Bucharest market about socialist-era street names, separating which name respondents actually use from their explanations for persistence and their evaluations of that persistence, and argued that habit rather than ideological resistance better explains continued use of the older name. The principle that use must not be read as attitude is theirs.

Karlander (2017) analysed a contested road-naming process as a temporally extended metapragmatic struggle over authenticity and institutional sanction, attending explicitly to justifications and to expressions of approval and invalidation.

Fabiszak and Rubdy (2021) compared media debates over the renaming of 23 Lutego Street in Poznań and Aurangzeb Road in New Delhi using the Discourse-Historical Approach, drawing on roughly 50 articles and about 1,000 below-the-line reader comments for the Polish case and 22 articles with 147 comments for the Indian one. Online lay argument about renaming, in reader comments, is therefore also established prior art.

Buchstaller, Schneider and Alvanides (2023) built a taxonomy of arguments for and against commemorative street renaming from a century of Leipzig press coverage, with five main categories — historical-preservational, moral-ideological, practical, didactic, and ideological instability — and roughly two hundred coded arguments. Any claim to novelty in constructing reason categories for renaming disputes is pre-empted by this work.

Dieterich, Martínez and Nieswand (2024) documented online renaming conflict in Murcia and, in Johannesburg, a sequence in which a speaker using an old street name is corrected to the new one and responds with ironic distancing. **Dvořáková (2025)** analysed the Koněvova/Hartigova renaming through municipal material, surveys and a Facebook poll attracting more than 900 comments, and closed by noting that it would be interesting to check later how firmly the new name has taken root and how common the old informal form remains — an explicit statement of the longitudinal question, posed but not carried out. **González Milea, Mérida Donoso and Rivero Gracia (2026)** coded 132 narrative productions by 110 preservice teachers responding to the Gulf of Mexico / Gulf of America controversy, yielding 430 coding units across territory and identity, controversy and intentionality, power and ideology, and credibility.

Taken together these works pre-empt as novelty: contested official renaming; persistence of old names; habit as an alternative to resistance; online argument about renaming; correction, irony and ridicule around name choice; authority and jurisdiction as themes; reason taxonomies; temporal tracking of competing old and new names after a political naming act; empirical coding of discourse around a recent politically imposed body-of-water name; and the distinction between referential and metalinguistic use itself. None of these is claimed here.

What is not found among them is a particular combination: a sample of pre-existing comment spaces frozen and sealed before its outcomes were inspected, occurrence-level coding of whether each name token is used or discussed, a separate record of explicitly stated position on the renaming, and a second observation point fixed in advance. The contribution claimed is that combination and the case it yields — a residual empirical and design contribution, not a theoretical one. The claim is bounded to the literature located and re-audited for this study and is not a claim of priority.

3. Data and methods

3.1 Case and intervention

Executive Order 14422, *Honoring the American History of the Great Lakes and Renaming Lake Ontario as Lake America*, was signed on 27 August 2026. It directs the Secretary of the Interior to rename the body of water then named Lake Ontario, instructs the Board on Geographic Names to issue guidance so that federal references and agency maps reflect the change, and requires the Geographic Names Information System to be updated within thirty days. Its directives are addressed to federal officials. The order does not purport to settle state, provincial, private or Canadian usage, and this study accordingly describes the event as a U.S. federal renaming rather than as an unqualified change to the lake's name.

Implementation began early rather than at the end of that thirty-day period. On 29 August 2026, two days after signature, Google reported that the Geographic Names Information System had formally changed the name in the United States and that Maps would accordingly show *Lake America* to users in the United States, *Lake Ontario* to users in Canada, and both names elsewhere, with the display rollout starting at that date. A visible and asymmetric infrastructural cue was therefore in place from the third day of the post window onward, and one of the coded comments turns on exactly that kind of competing platform label.

For this study T0 is 27 August 2026, the date of signature. The dates 25 and 26 August are excluded from the primary analysis — not because the order fell on them, but because reporting and anticipation in the immediate run-up to a signature cannot be assigned cleanly to either side of it. That exclusion was fixed in the design freeze before any outcome was inspected and is retained unchanged. The pre-intervention window is 1 August 2026 00:00 UTC to 25 August 2026 00:00 UTC (exclusive end; 1–24 August), and the first post-intervention window is 27 August 2026 00:00 UTC to 21 September 2026 00:00 UTC (exclusive end; through 20 September). The post window therefore opens at the start of the calendar day on which the order was signed, so comments written on 27 August before signature would fall on the post side of the boundary; in the observed material the earliest new-name comment is dated 28 August, so this boundary does not affect the result reported below.

Temporal placement uses the YouTube Data API field `publishedAt` only; `updatedAt` is retained as edit metadata and never substituted. All times are UTC.

3.2 Search 1 — contemporary corpus

Search 1 was collected on 11 September 2026 under frozen contemporary queries: 36 videos, 121,161 retrieved comments, of which 9,086 contained an exact occurrence of either target name (Ontario only 5,517; America only 2,777; both 792). Two videos yielded no exact-name candidate and remain in the video corpus. From the candidate frame, a development sample of 267 comments was drawn for building the coding scheme, and a separate held-back sample of 134 comments was reserved for validation, with an 80-comment subset fixed in advance for reliability.

The three samples have distinct and non-interchangeable jobs. The development sample was read freely while the coding scheme was still being built and revised, and nothing drawn from it is reported as a result. The validation sample was sealed during that process and opened only once the manual was frozen, so it tests whether the frozen categories can be applied to material that did not shape them. The reliability subset was drawn from the validation sample before any of it was coded, which is what allows agreement to be computed on genuinely independent decisions rather than on cases selected after the fact.

Search 1 supplies measurement validation only. Its proportions are never generalised to the 9,086 candidates, and they are not comparable to the Search-2 figures as prevalence: the contemporary queries that produced it were chosen because they surface discussion of the renaming, so the corpus over-represents metalinguistic material by construction. That bias is harmless for the use to which the sample is put and fatal for any other.

3.3 Search 2 — held-out legacy corpus

Search 2 was designed to observe comment spaces established before the renaming and selected without reference to any post-intervention name outcome. Only the established name was queried, with frozen ordinary-use query families (fishing; boating and water sports; weather and water; travel and leisure; local and community; nature and shoreline). A video was eligible if it was uploaded before 25 August 2026 00:00 UTC, carried the exact established name in its current title or description, carried no new-name variant and no renaming wording in current metadata, and had at least 50 top-level comments in the twelve months before the upload cutoff. No post-intervention count, name count or code influenced selection. The resulting sample is 35 videos across 30 channels (27 channels contribute one video, two contribute two, one contributes four).

Because YouTube does not expose title and description edit history, the screen establishes only that no new-name cue is present in current metadata. Full transcripts, audio and pinned creator comments could not be verified systematically for all 35 videos. The corpus is therefore described throughout as **pre-existing legacy videos without a *Lake America* cue in the current title or description**, and is never described as host-uncued. Following the controlling method amendment, it is described as metadata-uncued: the frozen title and description metadata do not frame these videos as being about the 2026 renaming.

The primary unit is the top-level comment. Replies were not retrieved for the post-period and are excluded from the present analysis by explicit decision rather than by omission; a pre-period reply archive exists and is preserved as provenance.

3.4 Retrieval

Primary retrieval is raw exact-string matching, case-insensitive, on the stored comment text: `\blake\s+ontario\b` and `\blake\s+america\b`. This matches the generator of the Search-1 candidate frame exactly, so that cross-corpus statements compare identically defined sets.

A separate recall audit, frozen before inspection, applied a broader detector inheriting the Search-1 span-detector family and covering transparent variants (`LakeAmerica`, `#LakeAmerica`, `Lake of America`, `Lake Amerika`, `Lake Murica` / `Lake 'Murica`, and the symmetric Ontario forms), together with zero-width normalisation and separation of glued @-mentions. The difference set was adjudicated field-blind: rows were ordered by an implementation-independent hash key, `publishedAt` and window membership were withheld, and temporal fields were rejoined only after the adjudication export had been sealed and hashed. Confirmed variants form a declared recall stratum reported alongside the primary counts and are never pooled into them.

3.5 Coding

Two human decisions are made, and no others.

Per name occurrence, `REFERENCE_FUNCTION` is `REFERENTIAL` (the name is used to refer to the lake), `METALINGUISTIC` (the name itself is quoted, discussed, proposed, rejected, corrected, compared or reported) or `UNCLEAR`. The operational test is substitution: replace the name with “the lake”; if the sentence still works, the occurrence is referential. *The lake looked beautiful last night* preserves the original; *it is not the lake any more, it is the lake* does not, because the name is what the sentence is about. Each occurrence is decided independently, and one comment may contain both functions — a correction that supplies a new label while describing where the writer lives will do so.

`UNCLEAR` is used where both readings remain live after the substitution test, and it is a recorded outcome rather than a missing value. Suppressing it would push genuinely ambiguous occurrences into one of the two substantive categories and inflate whichever of them the coder happened to favour.

Per comment, and only where at least one occurrence is metalinguistic, `EXPLICIT_RENAMING_STANCE` is `SUPPORT`, `OPPOSE`, `NONE` or `UNCLEAR`. “Explicit” means that the position can be pointed at in the words of the comment. Tone, irony, emoji, flags and the mere choice of one name over the other are not evidence of a position. `NONE` covers reports, questions, jokes without a position, statements about one’s own usage, and insults aimed at a person rather than at the renaming.

Name identity (`ONTARIO` / `AMERICA`) and form (`EXACT` / `VARIANT`) are derived from the string and are not human decisions. `UNCLEAR` is a valid code in both human variables, not a failure state.

This scheme replaced a finer metapragmatic scheme during development, before any validation or held-out coding, because several of the finer categories could not be applied without interpretive reconstruction. That replacement is recorded as a formal post-freeze amendment to the parent design, which the amendment supersedes in the sections

concerned.

3.6 Freeze logic and thresholds

The design freeze, the coding manual, the retrieval definition, the analysis windows and the denominators were sealed and hashed before any held-out outcome was inspected. The coding interface displayed the comment text and the highlighted name span and nothing else: no video identifier, no date, no window. A hidden map linking coding identifiers to video, date and window was opened only after every coding decision had been written and the coded file hashed.

The parent freeze contained a prespecified feasibility condition for the secondary video-level analysis: inferential interpretation would be attempted only if at least ten eligible videos contained both a metapragmatic case and referential attestations of both competing names. If that condition was not met, the relationship was to remain descriptive. It was not met in the first Search-2 snapshot.

Additional restrictive guardrails were recorded on 11 October 2026 after the first reveal: at least 30 post-T0 new-name candidate comments across at least five videos for a central quantitative analysis; 10–29 candidates, or concentration in fewer than five videos, for exploratory and descriptive treatment only; fewer than ten candidates for no quantitative diffusion claim; and at least 20 clearly referential new-name uses across at least five videos before referential-use analysis could be central. No pre-reveal source for these later numerical guardrails has been located. They are therefore treated here as post-reveal anti-escalation rules, not as preregistered criteria. The descriptive status of the first snapshot does not depend on them.

Denominators were counted before outcomes were inspected: 478 top-level comments in the pre-window with 23 of 35 videos active, and 162 in the post-window with 21 of 35 active; 17 videos were active in both, six in the pre-window only, four in the post-window only, and eight in neither. Counting comments by date reveals no name information, and no denominator-triggered change of window was permitted.

For a zero finding, an upper bound is reported at video level using the one-sided exact binomial form `1 − 0.05^(1/n)`; no comment-level bound is presented as though comments were independent. With 35 videos and 30 channels the corresponding bounds for a complete null are 8.2% and 9.5% respectively. The minimum detectable correlation for the design's secondary video-level question, at $\alpha = .05$ two-sided and 80% power with 35 videos, is $|r| \approx 0.46$; the design can therefore detect only large associations, and this was computed before any coded relationship was inspected.

3.7 Independent coding and reliability

The author coded the full 134-comment held-back validation set. A frozen 80-comment subset, drawn before any outcome was seen, was coded independently by a second coder working from the same manual in the same blinded interface and without access to the author's decisions. Agreement was computed on the independent pre-adjudication decisions. Eleven comments contained at least one disagreement; all were subsequently adjudicated by the author against the frozen manual. Both original coder files are preserved unchanged and the reported reliability figures remain the pre-adjudication values.

As a process audit rather than a reliability measure, the same 80 comments were also coded by a large language model. Its agreement with each human coder was substantially lower than human–human agreement ($\kappa = 0.598$ and 0.594 for reference function, 0.509 and 0.538 for the metalinguistic gate, 0.623 and 0.601 for conditional stance). This is reported because it is the reason the reliability claim rests on two people.

3.8 Ethics and privacy

Public comments are not treated as consequence-free text. No commenter is named and no channel identifier or user handle is reported. Comment text is quoted only where the wording is analytically necessary; otherwise it is paraphrased. Video identifiers are withheld in the manuscript body and referred to as video A–D. No ethics approval was sought or is claimed; the study uses publicly accessible material collected through the platform's own API.

The public replication deposit is deliberately privacy-minimising; its contents and the material withheld from it are specified under *Data and code availability* below.

4. Results

4.1 Before the renaming

In the pre-intervention window, 23 of the 35 videos received top-level comments at all, 478 in total. Seventeen of those comments contain an exact target name, and all seventeen are ordinary referential uses of *Lake Ontario*. Exact *Lake America* occurs in **none** of the 478 comments and in **none** of the 23 active videos. No pre-window comment discusses either name: the metalinguistic count is zero.

A zero is robust against the composition problem that affects ratios: it does not depend on which videos happened to be busy. Reported at video level, zero cases in 23 active videos places the one-sided 95% upper bound on video-level incidence at 12.2%.

The 17 pre-window candidates come from three videos (nine, six and two comments).

4.2 After the renaming

In the first post-intervention window, 21 videos were active with 162 top-level comments. Sixteen contain an exact target name, distributed across six videos. Five comments contain referential *Lake Ontario*; ten contain at least one metalinguistic occurrence; one contains referential *Lake America*.

Eleven comments contain exact *Lake America*. Of these, **ten are metalinguistic and one is referential**. Five of the ten metalinguistic comments also contain *Lake Ontario*. Across both windows the coded material comprises 41 exact occurrences in 33 comments (29 *Lake Ontario*, 12 *Lake America*).

The single referential case uses the new name straightforwardly to describe the lake, without commenting on the naming act.

The metalinguistic comments are corrections, status reports, appeals to rules, competing platform labels, and jokes — for example a bare asterisk correction supplying the new form, an assertion that the old name is now wrong, a report that one map provider has changed the label while another has not, and a one-word dismissal attached to the name.

Table 1. Comments containing exact *Lake America*, first post-intervention window, by video.

Video	PRE top-level comments	POST top-level comments	comments with the new name	name discussed	name used referentially
A	66	19	6	6	0
B	65	39	3	3	0
C	20	10	1	1	0
D	47	5	1	0	1
Total	**198**	**73**	**11**	**10**	**1**

All four videos in which exact *Lake America* appeared during the post-window were already active in the pre-window. Together they contributed 198 pre-window top-level comments, none containing exact *Lake America*, and 73 post-window comments, eleven containing the exact new name. This four-video comparison is descriptive and is not treated as an independent rate estimate.

4.3 Dispersion and robustness

Read per video rather than pooled, the pattern is this: **in three of the four videos where the new name appears, every occurrence is talk about the name; in the fourth there is a single ordinary use and no metalinguistic case.** This statement does not depend on pooling and therefore does not depend on the weight of any one comment section.

The concentration still deserves a number. Six of the ten metalinguistic comments occur in video A. Removing that video leaves four metalinguistic against one referential across three videos. The direction survives; the ratio, with a single case in one of its cells, cannot be stable under removal — removing video D instead leaves ten against zero. Five remaining cases fall below the threshold for any quantitative reading. The leave-one-out pass is therefore reported as a robustness note, not as a second estimate.

4.4 Explicit stance

Among the ten comments that discuss the name, the sealed coding records four stating explicit support for the renaming, one explicit opposition, two with no position stated, and three in which a position seems intended but cannot be read off the words.

Two qualifications belong beside these numbers rather than after them. First, **all four supporting comments come from video A** and the single opposing comment from video C; at video level this is one comment section containing a small pro-renaming exchange and one comment elsewhere objecting, not a four-to-one balance of opinion in a corpus. Second, one of the four turns on a borderline case: a comment consisting only of an asterisk and the new name. It was sealed as `SUPPORT` before the hidden map was opened; a later strict re-reading finds no position stated in the words, and under that reading the distribution is three support, one opposition, three no position and three unreadable. Both readings are reported. Neither changes the result in § 4.2, because that comment is metalinguistic under either reading.

With ten cases in three comment sections, no stance figure is drawn and no proportion is claimed.

4.5 Recall boundary

The frozen recall audit returned three loose-only occurrences in the post-window, all adjudicated field-blind as true target variants: one *Lake of America* and two *Murica* forms, all three in video A and all derived as *AMERICA*. Their identity as true target variants was therefore established before window membership was revealed. A later and explicitly post-reveal conservative functional sensitivity classified the isolated *Lake of America* form as `UNCLEAR` and the two *Murica* forms as `METALINGUISTIC`. These cases remain supplementary and are not pooled into the exact-string primary count.

Under that post-reveal sensitivity, the recall margin adds no clearly referential new-name use. Had the three been pooled into the primary count, the post-window figure would read fourteen comments rather than eleven and the share that is ordinary reference would have fallen rather than risen. Reporting the two sets separately keeps that visible instead of absorbing it into a single number.

No loose-only occurrence was found in the pre-window, and the extended observation window to 29 September 2026 contained no further exact-name candidate — so the eleven comments are not an artefact of where the primary window happens to end.

4.6 Measurement validation from the contemporary corpus

In the held-back 134-comment validation sample from Search 1, the 55 occurrences of *Lake America* divide into 2 referential, 49 metalinguistic and 4 unclear; the 100 occurrences of *Lake Ontario* into 19 referential, 79 metalinguistic and 2 unclear. At comment level, 20 of the 134 comments contain at least one referential target occurrence and 112 contain at least one metalinguistic occurrence, with 13 comments containing metalinguistic occurrences of both names.

These proportions are **not** an estimate for the 9,086 Search-1 candidates: the 134 were drawn for validation, not prevalence, from a corpus assembled under contemporary queries that over-represent renaming talk by construction. Their role is narrower and sufficient — in an independently collected sample, hand coding again finds that the presence of the string and ordinary use of the name are not the same measurement.

The pattern is not specific to the newly imposed label within this validation material. *Lake Ontario* was also predominantly metalinguistic: 79 of 100 occurrences were coded as metalinguistic, compared with 19 referential and 2 unclear. Because Search 1 was deliberately assembled under renaming-salient queries, this is not a population comparison between the two names. It reinforces only the narrower measurement point that, in this controversy-oriented validation corpus, exact name occurrence cannot be equated with ordinary reference for either label.

Table 2. Pre-adjudication agreement between two independent coders (80 comments, 91 occurrences).

Decision	raw agreement	Cohen's κ
Reference function, per occurrence	85/91 = 93.4%	0.776
Comment-level metalinguistic gate	75/80 = 93.8%	0.763
Stance, where both coders reached the question	60/65 = 92.3%	0.808
Stance, all 80 comments including `NOT_ASKED`	70/80 = 87.5%	0.780

5. Discussion

5.1 What the case shows

In this corpus and this first window, the new name became visible in text more readily than it became a resource for ordinary reference. That is the whole finding. It is not evidence that the renaming failed, and a single 25-day observation window is far too short a span for any such claim. Nor is the one ordinary use evidence that the renaming succeeded.

The concrete value of the distinction is visible in the arithmetic. A string-only analysis of the same material would have reported eleven post-intervention comments containing exact *Lake America*, against none in the frozen pre-window. Functional coding classifies ten of those eleven as metalinguistic and one as ordinary reference. The other ten are instances in which the name itself is being discussed, corrected, reported, evaluated or otherwise made salient — counted, under the string-only approach, as though they were instances of the name in use.

The gap between those two readings is not a rounding difference. One describes a label entering circulation; the other describes a label becoming a subject of argument. Both are real events and both are worth counting, but they answer different questions, and only the second is what the raw frequency actually measured here.

5.2 How far that generalises

Not far, and the limits should be stated rather than implied. Eleven comments cannot establish that string counts generally overestimate adoption. Pires (2007) already showed functional heterogeneity in online toponym tokens, and Benton (2026) already excluded both-name stories from a name-choice ratio for essentially this reason. What this case adds is a worked instance with the functional judgement made explicit, per occurrence, under a frozen manual, in material collected before the outcome was known. For research on contested renaming the practical implication is modest and specific: where a name is likely to become an object of public argument, a frequency series should be accompanied by evidence about what the occurrences are doing.

The Search-1 validation sample also cautions against treating metalinguistic predominance as a property of *Lake America* alone: the established name was predominantly metalinguistic there as well, under a sampling design deliberately rich in renaming talk.

5.3 Stance and use are separate things

The design keeps explicit stance apart from referential use, and the first window illustrates why. The comments that state support for the renaming are corrections and appeals to rules — they are talk about the name, not use of it. The one comment that uses the new name plainly states no position at all. Two non-equivalences follow for this material: explicit approval is not linguistic adoption, and linguistic adoption does not express approval. This is not offered as a theory of stance; it is a reason not to build political interpretation into the classification of name use, which is what the coding manual required from the outset.

5.4 Local clustering

The material is concentrated. Six of ten metalinguistic comments, all four supporting comments and all three loose-recall variants come from one video. Whatever that comment section was doing in early September, it is one comment section. The per-video reading in § 4.3 is offered precisely because it survives this concentration, and nothing in this paper should be read as a rate for YouTube commenters in general.

5.5 The second snapshot

A second observation of the same 35 videos, with the same retrieval definition and the same manual, was fixed before the first held-out result was known and remains scheduled for 22 December 2026. This is the part of the design that the present report cannot deliver, and it is worth being explicit about why. A single window of 25 days, opening on the day of signature, can show how a new label behaves at the moment it becomes available. It cannot distinguish a label that is being argued about on its way into use from one that is being argued about and then dropped, because both look identical at this range. Only a second, equally frozen observation separates them.

December is also a later state of the naming environment, not merely a later date. By the December observation that environment will have had substantially more time to develop, including continued exposure through federal naming infrastructure and commercial map products. Because those exposures had already begun during the first window, the second snapshot is a later observational state rather than a clean before-and-after contrast on implementation.

The project holds no directional hypothesis. December may show the first-window configuration persisting, clearly referential uses becoming more visible, the new-name signal decaying, or a mixed picture. None of these is privileged here, and setting them out in advance is what prevents whichever arrives from being presented afterwards as the expected one. The counts will again be small, and the same per-video discipline will apply: a second observation widens the design in time, not in sample size.

The United States federal general election of 3 November 2026 falls between the two observation points. December therefore cannot be read as the isolated effect of elapsed time since the renaming: political salience, media attention and who comments at all may all differ. The second snapshot is a second frozen observation under changed

conditions, not a controlled follow-up. Stating this before the data exist is the only way to keep it from being discovered afterwards.

6. Limitations

Small and clustered counts. Eleven comments, four videos, one referential case. The prespecified video-level feasibility condition for inferential analysis was not met, and the first snapshot is therefore treated descriptively. Later numerical guardrails were added only as post-reveal anti-escalation rules and are not presented as preregistered.

Reliability is established on Search 1, not on Search 2. The held-out Search-2 material was coded by the author alone, under the frozen manual and in the blinded interface, with the hidden map opened only afterwards. The agreement figures in Table 2 come from the Search-1 reliability subset and show that the categories can be applied consistently by two independent coders on comparable material; they do not directly establish agreement on the 33 Search-2 comments. A second coder on the held-out material would have strengthened that link and was not available for this snapshot.

No prevalence estimate. Neither corpus supports a statement about how often anybody uses either name. Search 1 was assembled under renaming-salient contemporary queries and over-represents metalinguistic material by construction; Search 2 is 35 purposively selected videos.

Host cues only partly observable. The videos were screened for absence of a new-name cue in current title and description. Transcripts, audio and pinned creator comments could not be checked systematically across all 35, and unobserved host-side exposure cannot be excluded.

Retrieval boundary. The primary count is exact strings only. Transparent variants were audited separately under a procedure frozen before inspection and are reported as a recall stratum, never pooled into the primary count. A post-outcome count-only diagnostic on the full Search-1 archive found 369 comments that a broader detector would have retrieved beyond the 9,086-comment exact frame; those 369 are a mechanical, unverified upper bound on the Search-1 recall margin and were never coded or added to any analysis.

Window asymmetry. The pre-window is 24 days and the post-window 25; the post-window is also the smaller sample, which is why counts rather than rates are reported throughout.

One case per cell in places. Several cells in Table 1 contain a single comment. They are shown because hiding them would misrepresent the dispersion, not because one comment supports anything.

7. Conclusion

In 35 pre-existing comment spaces about Lake Ontario, observed in a frozen window before and a frozen window after the 27 August 2026 U.S. federal renaming directive, the exact new name was absent from the pre-window and appeared afterwards in eleven comments across four videos. Ten of those eleven were about the name; one used it. In three of the four videos containing the new name every occurrence was metalinguistic. An independently collected validation sample points the same way.

The contribution is not a theory of naming, toponymic resistance, or linguistic uptake; each of those belongs to prior work cited above. It is a transparent demonstration, in material sealed before its outcomes were seen, that the visibility of a contested name and its use as a name can be separated in natural comment data — together with a second observation of the same spaces, already scheduled, that will say something about what happened next.

Data and code availability

The preprint is deposited at https://doi.org/10.5281/zenodo.23303764. The accompanying replication and provenance dataset is deposited at https://doi.org/10.5281/zenodo.23304624.

Public replication materials for Version 1 / First Snapshot include the frozen design documents, the final minimal codebook, the post-freeze method amendment, the loose-recall freeze and audit documentation, analysis and audit scripts, aggregate result tables, reliability outputs, pseudonymised coding decisions, and checksum/provenance records, together with a verification script that recomputes every reported aggregate from the deposited files.

Raw YouTube comment text, commenter identifiers, comment identifiers, and the lookup table linking manuscript video labels to public video IDs will not be redistributed in the public deposit, in order to reduce the traceability of individual commenters. The unredacted source files and lookup tables are retained privately by the author with recorded SHA-256 hashes.

The project record preserves the blinded coding files, the sealed pre-reveal coded file, the hidden map, the superseded intermediate state of the borderline stance case, the recall-audit ledger, the reliability files of both human coders, the retrieval and sensitivity scripts, and a checksum chain linking controlling artefacts. Recorded controlling hashes include the parent design freeze `b7852855…`, coding manual `2bef695d…`, method amendment `2f1c555b…`, loose-recall freeze `015f8a67…`, sealed Search-2 coding `b13df94d…`, hidden map `6afcd53d…`, and superseded provisional coding `3e60bd75…`.

Note on sources and scope

Pires (2007), Light and Young (2014), Karlander (2017), Dieterich et al. (2024), Dvořáková (2025), Benton (2026) and González Milea et al. (2026) were re-checked in primary full text or publisher-hosted originals during the project's literature re-audit of 11 October 2026. Buchstaller, Schneider and Alvanides (2023) and Fabiszak and Rubdy (2021) were verified against publisher records on the same date. The negative literature finding is bounded by a documented search and is qualified accordingly: coverage of non-English dissertations and local onomastic series is incomplete, indexing lag affects 2025–2026 work, some full texts were inaccessible, and platform-internal datasets are unpublished. No claim of priority is made. The word *first* is used only to identify the first of two prospectively fixed observation points, not to claim priority over earlier research.

The review was deliberately adversarial but necessarily bounded. Pointers to relevant work that may have escaped it are welcome and will be acknowledged in the next version.

Provenance and AI disclosure

This is a human-directed, AI-assisted independent research project, and the disclosure below is deliberately specific rather than formulaic.

The author was responsible for the research question and case selection; the decision to run adversarial literature kill tests and the decision to continue rather than abandon the project after they returned a pivot; the design of both searches; the sampling logic and all freeze decisions; the decision to replace the original fine-grained coding scheme with the minimal scheme when it proved to require interpretive guessing; the decision to separate referential use from metalinguistic discussion and to record stance only where it is explicit; the coding of the validation material; the adjudication of coder disagreements; the final decision on each difficult coding case, including the borderline case reported in § 4.4; the decision not to widen samples or windows when the held-out result proved sparse; the decision to keep exact and variant retrieval separate; the retention of the 22 December observation point; and all substantive claims and limitations in this manuscript.

AI systems were used extensively, as tools: exploratory source research and adversarial literature review; comparison of candidate designs; generation of the coding interface, API scripts and data-processing code; technical and provenance audits including checksum verification and recomputation of all reported statistics from the raw files; independent re-derivation of the reliability coefficients and the dispersion analysis; methodological stress-testing that changed the framing of several results in this paper, including the shift from pooled to per-video reporting in § 4.3 and § 4.4; and manuscript outlining, drafting and English-language revision. AI-generated suggestions were treated as proposals and were accepted, modified or rejected by the author. Some suggestions materially changed the workflow or the framing after author review; final substantive decisions remained with the author. No empirical claim in this paper rests on an AI judgement that was not checked against the sealed project record.

Independent coding. The author coded the full held-back validation set. A frozen subset was coded independently by a second coder using a blinded interface and without access to the author's decisions. Reliability was computed on the independent pre-adjudication decisions; disagreements were adjudicated afterwards by the author, and both original coder files are retained unchanged. The second coder is referred to throughout as an independent second coder and is not named.

Provenance is reported with its imperfections. The project record contains a superseded intermediate state of the stance decision reported in § 4.4, written under the same filename as the final sealed file and subsequently renamed; the method amendment formalising the coding simplification was written after the simplification itself and says so; several sensitivity analyses are post-reveal and are labelled as such; and the later numerical guardrails in § 3.6 could not be traced to a pre-reveal source and are therefore reported as post-reveal anti-escalation rules rather than preregistered criteria. These imperfections are retained rather than tidied, because the point of the record is to let a later reader contradict the paper rather than merely reproduce its totals.

Version history

v1.0, 11 October 2026 — Version 1 / First Snapshot. Reports the first of two prospectively fixed observation points. Search-1 measurement is closed; the first Search-2 exact-name reveal is complete. The prespecified video-level feasibility condition for inferential analysis was not met, so the first snapshot is reported descriptively. Later numerical guardrails are documented as post-reveal anti-escalation rules rather than preregistered criteria. The second observation remains scheduled for 22 December 2026 and has not yet been taken. Version 1 is intended to remain permanently available and not to be overwritten by the December update.

Editorial revision of 16 October 2026, within Version 1.0. The deposited file was replaced during the period in which the record still permits file replacement. The abstract was rewritten for clarity: the measurement problem is stated in the opening rather than implied, the single referential case is identified explicitly as the first ordinary use of the new name observed in the panel, and the limits of a single case are stated alongside it. The Zenodo deposit identifiers were inserted, and a note inviting pointers to missed literature was added. **No number, coding decision, window, threshold or interpretation was changed.** The results, methods, limitations and provenance sections are identical to the file of 11 October 2026.

Planned v2.0, after 22 December 2026 — Second Snapshot. The second observation of the same 35 videos will be released as a new version of this record under the same design, retrieval definition and coding manual, with Version 1 preserved unchanged and an explicit account of any method or coding clarification introduced after it. Readers seeking the two-point longitudinal picture should consult that version; the present version is cited for the first observation window and for the design.

References

Benton, J. (2026). Trump can't get the media to use "Gulf of America." Will "Lake America" be any different? *Nieman Journalism Lab*, 3 September 2026.

Buchstaller, I., Schneider, C., & Alvanides, S. (2023). Towards a taxonomy of arguments for and against street renaming: Exploring the discursive embedding of street name changes in the Leipzig cityscape. *Linguistic Landscape*, 9(1), 5–35. https://doi.org/10.1075/ll.21016.buc

Executive Order 14422 of 27 August 2026. *Honoring the American History of the Great Lakes and Renaming Lake Ontario as Lake America*. The White House. https://www.whitehouse.gov/presidential-actions/2026/08/honoring-the-american-history-of-the-great-lakes-and-renaming-lake-ontario-as-lake-america-3a36/

Google. (2026, 29 August). *How the GNIS Lake Ontario/Lake America name change in the U.S. will appear in Maps.* Google Blog. https://blog.google/products-and-platforms/products/maps/gnis-lake-ontario-lake-america-name-change/

Dieterich, M., Martínez, D. O., & Nieswand, B. (2024). Re-name the streets. Toponymic struggles and civic belonging in Murcia (Spain) and Johannesburg (South Africa). *Ethnography*, 27(3). https://doi.org/10.1177/14661381241269638

Dvořáková, Ž. (2025). Koněvova Street: A case study on the decommunisation and derussification of Czech urbanonymy. *Onoma*, 60, 69–96. https://doi.org/10.34158/ONOMA.60/2025/5

Fabiszak, M., & Rubdy, R. (2021). Media debates over the renaming of the cityscape. *Linguistics Vanguard*, 7(s5). https://doi.org/10.1515/lingvan-2020-0089

González Milea, A., Mérida Donoso, J. A., & Rivero Gracia, P. (2026). Gulf of Mexico/Gulf of America: Critical literacy and mediatized discourses in teacher education. *Social Sciences*, 15(9), 584. https://doi.org/10.3390/socsci15090584

Karlander, D. (2017). Roads to regimentation: Place, authenticity and the metapragmatics of naming. *Language & Communication*, 53, 11–24. https://doi.org/10.1016/j.langcom.2016.11.002

Light, D., & Young, C. (2014). Habit, memory, and the persistence of socialist-era street names in postsocialist Bucharest, Romania. *Annals of the Association of American Geographers*, 104(3), 668–685. https://doi.org/10.1080/00045608.2014.892377

Pires, M. (2007). Investigating non-universal popular urban toponyms: Birmingham's Pigeon Park. *Onoma*, 42, 131–154. https://doi.org/10.2143/ONO.42.0.2047078