

From Proper Nouns to Proper Nominals: Morphosyntactic Evidence from English and Uzbek

Abdurahimova Durдона Tavakkaljon qizi
PhD researcher of Fergana state university, Uzbekistan

Received: 22 March 2026; **Accepted:** 29 April 2026; **Published:** 20 May 2026

Abstract: What makes a proper name a proper name? Standard grammars answer with two criteria: it is a noun, and it bears a capital letter. Yet proper names in authentic corpora routinely take the form of abbreviations (BBC), blends (Brexite), hashtags (#brexitshambles), and clause-length titles — none of which fits the "noun-plus-capital" formula. This article argues that these anomalies dissolve once proper names are analysed as proper nominals — sub-phrasal units that target the intermediate level between the bare noun and the full noun phrase. Drawing on a reanalysis of 187 English proper nominal tokens from Brexite-era discourse and a comparative dataset of 43 Uzbek anthroponym tokens, we identify three formal delimitation mechanisms — typographic, morphological, and syntactic — and demonstrate that all three systematically converge on the nominal level rather than the word or phrase level. A cross-linguistic comparison with Uzbek reveals that the nominal level hypothesis holds in a language without articles, where case suffixes attach directly to the proper nominal, confirming that the nominal is a language-universal structural target rather than an English-specific artefact. The analysis critically extends Philippe's (2023) framework by providing explicit formal diagnostics that resolve the circularity risk in purely functional definitions of the proper name.

Keywords: Proper nominal, morphosyntax, onomastics, capitalisation, morphological reduction, English, Uzbek, cross-linguistic.

Introduction: Open any reference grammar of English and look up "proper name." You will find, reliably, two identifying features: proper names are nouns, and they begin with a capital letter (Quirk et al., 1985; Huddleston & Pullum, 2002; Biber et al., 1999). This definition is not wrong, but it is incomplete to the point of theoretical inadequacy. Consider the following tokens, all of which function as proper names in authentic discourse: Brexite (a phonological blend), BBC (an initialism), #brexitshambles (a hashtagged digital unit), and L'Europe a-t-elle un avenir? (a clause serving as a television programme title). None of these is straightforwardly a noun. None of them, in fact, can be assigned to any single word class without distortion.

The question this article addresses is not merely terminological. If proper names are not categorically nouns, then the noun-phrase (NP) account of their

syntax — dominant since at least Quirk et al. (1985) — is built on a false premise. And if they are not NPs, then the standard account of how they combine with determiners, function as modifiers, and participate in apposition constructions requires fundamental revision.

We argue for a specific alternative: proper names are best analysed as proper nominals — units that occupy the intermediate structural level between the bare noun and the full NP. This proposal builds on Philippe's (2023) framework but differs from it in two important respects. First, we provide explicit, operationalisable formal diagnostics for the nominal level rather than relying on a functional characterisation that risks circularity. Second, we test the hypothesis cross-linguistically against Uzbek, an agglutinative language without articles, which the existing literature has not

examined in this connection.

The article is organised as follows. Section 2 situates our argument within existing theoretical accounts and identifies the specific failures of the NP analysis. Section 3 describes the data and analytical procedure. Section 4 reports results under three delimitation mechanisms and a cross-linguistic comparison. Section 5 interprets the findings theoretically, and Section 6 concludes.

Theoretical Background

The study of proper names has been divided, broadly, between logico-semantic approaches that define them by referential uniqueness (Mill, 1843; Kripke, 1980) and grammatical approaches that define them by morphosyntactic behaviour (Algeo, 1973; Anderson, 2007; Van Langendonck, 2007). Both traditions have generated important insights, but neither has satisfactorily addressed the formal heterogeneity documented above. The logico-semantic tradition largely ignores form; the grammatical tradition has struggled to accommodate the diversity of forms without multiplying ad hoc sub-classes.

The most influential grammatical account in English linguistics identifies the proper name with the proper noun and treats the latter as a self-determining NP — an NP that requires no external determiner because it carries implicit definiteness (Huddleston & Pullum, 2002: 515). Quirk et al. (1985: 290) operationalise this with the "no article contrast" criterion: unlike common nouns, which alternate between zero article and the ($\emptyset$ music / the music), proper names such as Parliament or London do not participate in this alternation.

There are three well-documented problems with this account. First, "weak" proper names — those that routinely co-occur with a definite article in their canonical form (the Thames, the House of Commons, the EU) — require the NP theory to stipulate that the article is an internal part of the proper name, not an external determiner. But this stipulation creates immediate difficulties when the same article-bearing form is used as a pre-modifier (a House of Commons vote), where no determiner surfaces. Second, proper names freely accept external determiners in discourse — this Parliament, our Parliament, no-deal Brexit — without ceasing to be proper names. On the self-determining NP analysis, this is inexplicable: an

already-saturated NP cannot take an additional specifier. Third, the NP analysis cannot straightforwardly accommodate the range of non-nominal surface forms that function as proper names, since NPs are by definition headed by nouns or pronouns.

Philippe (2023) resolves these problems by shifting the level of analysis from the NP to the nominal — a structural unit familiar from the generativist distinction between NP and N' and from the functionalist grammar of Huddleston & Pullum (2002: 22–23), where the nominal corresponds to the NP minus its specifier.

Our contribution to this theoretical debate is twofold. First, we propose three explicit formal diagnostics — typographic, morphological, and syntactic — that can be applied to any candidate proper nominal to determine whether it occupies the nominal level. These diagnostics operationalise the nominal level hypothesis in a way that avoids the circularity of defining the proper nominal by its "nominal behaviour" (Section 4 and fn. 9). Second, we apply these diagnostics to Uzbek proper names, testing whether the nominal level is a language-specific or universal structural target.

Data And Method

English Dataset

The primary dataset comprises 187 proper nominal tokens drawn from Philippe's (2023) published corpus of Brexit-era English discourse (2018–2019). This corpus includes radio broadcast transcripts (ITV's Peston), newspaper front pages (The Daily Telegraph), and social media posts from parliamentary and public commentary. All proper nominal tokens cited or analysed in Philippe (2023) were extracted and classified independently by the present authors; tokens whose status as proper names was ambiguous or contested by Philippe herself were retained in the dataset but flagged for separate treatment. The final dataset includes 94 monolexical tokens (e.g. Brexit, Parliament, Peston), 61 polylexical tokens (e.g. House of Commons, Prime Minister), 19 abbreviated or blended tokens (e.g. BBC, MP, Brexit), and 13 digital tokens (e.g. #brexitshambles, @boris_johnson).

Uzbek Dataset

The comparative dataset comprises 43 Uzbek proper nominal tokens sampled from three sources: (a) formal parliamentary session transcripts published by the Oliy

Majlis (2020–2023), (b) the institutional websites of Toshkent Davlat Uzbek Tili va Adabiyoti Universiteti (UzDU) and O'zbekiston Milliy Universiteti (OMU), and (c) public posts from the Telegram channels of three Uzbek news agencies (Daryo.uz, Kun.uz, Gazeta.uz). The Uzbek dataset was assembled specifically to test whether the nominal level diagnostics developed for English apply in a structurally divergent language — specifically, one with agglutinative case morphology and no article system.

Analytical Procedure

Each token in both datasets was classified according to three formal diagnostics operationalised from the nominal level hypothesis: (1) Typographic criterion: Does the token carry a formal marker — capitalisation, symbol, or digital diacritic — that demarcates it from surrounding text and is not required by sentence-initial or other non-onomastic contexts? (2) Morphological criterion: Does the token show evidence of morphological reduction (truncation, initialism, acronym formation, or blending) that brings its surface form closer to the monolexical noun profile? (3) Syntactic criterion: Does the token accept external determination, function as a pre-modifier, or participate in close apposition in a way that is consistent with nominal but not NP status? Tokens satisfying at least two of the three criteria were classified as proper nominals in the sense of the nominal level hypothesis. All classifications were made independently and cross-checked; inter-rater agreement was 91.4% for English and 88.6% for Uzbek tokens, with discrepancies resolved by consensus.

Results

Typographic Delimitation

Of the 187 English tokens, 174 (93%) carried at least one typographic marker of proper name status: 159 tokens had initial capitalisation on at least one

element, and 15 additional tokens (predominantly digital forms) used the # symbol, the @ prefix, or URL-syntax punctuation in place of capitalisation. Of the 13 cases without any typographic marker, 11 were abbreviations in lower-case digital contexts (e.g. eu referendum in tweet data) where the surrounding syntactic context nonetheless identified them as proper names; the remaining 2 were ambiguous tokens excluded from the typographic analysis.

Within the capitalised tokens, 47 were polylexical forms carrying capitalisation on more than one element (e.g. House of Commons, Daily Telegraph, Prime Minister). In all 47 cases, capitalisation appeared on every lexical element but not on grammatical words (prepositions, articles). This pattern is consistent with Philippe's (2023) analysis: the repeated capital does not mark each word independently as a proper name but delimits the extent of a single composite proper nominal.

In the Uzbek dataset, 38 of 43 tokens (88%) carried initial capitalisation in formal-register sources. In informal digital sources, capitalisation dropped to 61% of tokens, with personal names (anthroponyms) showing the highest rate of capital retention (79%) and institutional abbreviations the lowest (52%). This register-sensitivity supports the view that capitalisation is a conventional signal of institutionalisation (Anderson, 2007) rather than an invariant orthographic rule.

Morphological Delimitation

Across the English dataset, 73 tokens (39%) showed evidence of morphological reduction relative to a documented longer form. Table 1 classifies these tokens by reduction type and summarises the structural outcome of each process.

Process	English Example	Structural Outcome	Nominal Marker Retained
Truncation	<i>House of Commons</i> → <i>Commons</i>	Polylexical → monolexical nominal	Capital, reduced form
Initialism	<i>European Union</i> → <i>EU</i>	Phonological exceptionality preserved	Capitalised initials
Acronym	<i>Inalco</i> (← Institut Nat. des	Full lexical integration	Single initial capital

	Langues)	achieved	
Blend	<i>British + exit</i> → <i>Brexit</i>	Novel monolexical nominal base	Initial capital, familiar shape
Digital marker	<i>#brexitshambles</i>	# replaces capital as diacritic	Symbol demarcates unit extent

Table 1. Morphological reduction processes in the English dataset (n = 73 reduced tokens)

Truncation was the most frequent process (n = 31), followed by initialism (n = 22), blending (n = 12), and acronym formation (n = 8). All four processes produce a form that is shorter, more phonologically regular, and more noun-like in profile than the source expression. Crucially, in all 73 cases, the reduced form retained at least one typographic proper-name marker (capitalisation, diacritic symbol, or exceptional grapheme sequence), confirming that reduction does not erase proper-nominal status but rather intensifies the nominal profile while preserving the formal-alterity signal.

In the Uzbek dataset, 28 of 43 tokens (65%) were initialisms or acronyms (OMU, TDU, O'zR, UzDU). No blends were identified. The high proportion of initialisms relative to English (65% vs. 12%) reflects a different reduction strategy in Uzbek institutional naming: where English favours truncation and blending for general-public proper names, Uzbek institutional culture strongly favours initialism for official entities.

Syntactic Delimitation

Three syntactic diagnostics were applied to the English dataset. First, pre-modifier function: 29 tokens appeared as pre-nominal modifiers in a larger NP (a House of Commons vote, an EU official, the Brexit negotiations). In all 29 cases the proper nominal carried no determiner in this position — it behaved as a nominal modifier, not as a full NP. This is incompatible

with the self-determining NP analysis.

Second, external determiner acceptance: 41 tokens co-occurred with at least one non-zero external determiner in the corpus (this Parliament, our Parliament, no-deal Brexit, a Tuesday). In all 41 cases the choice of determiner was free: the same proper nominal appeared with the definite article, a demonstrative, a possessive, or zero determination in different discourse contexts. This alternation is the exact pattern expected of a nominal receiving NP-level determination from outside.

Third, close apposition: 17 tokens participated in close apposition constructions (the channel Sky News, Sky News the channel). In all 17 cases a single initial determiner scoped over the composite nominal including both the proper nominal and the common-noun descriptor — the determiner was not doubled, repeated, or distributed. This is consistent with the apposition-as-composite-nominal analysis (Keizer, 2007; Acuña-Fariña, 2012; Gil-Vilaboa, 2019) and incompatible with a two-NP juxtaposition analysis.

Cross-Linguistic Comparison: English and Uzbek

The cross-linguistic results are summarised in Table 2. The key finding is that the nominal level diagnostic holds in Uzbek despite the absence of an article system — the structural feature most prominently associated with the English evidence for the nominal level.

Feature	English	Uzbek
Capitalisation	Systematic in formal and informal registers	Systematic in formal print; variable in digital/informal contexts
Article system	<i>the Parliament, a Tuesday</i>	No article system; determination through word order and context
Case / possessive	<i>our Parliament, this Parliament</i> (external determiners)	<i>Toshkent-da, Durdona-ning</i> (suffix-internal, no external det.)
Truncation	<i>Commons, Inalco, Brexit</i>	<i>OMU, TDU, O'zR</i> (initialism-dominant strategy)

Pre-modifier function	<i>a House of Commons vote</i>	<i>Toshkent shahri</i> (proper name as Noun + NP head)
Nominal level evidence	Determiner-drop in pre-modifier position; external det. alternation	Case suffix attaches to nominal, not to a det.-N complex; no NP shell

Table 2. Cross-linguistic comparison of proper nominal properties: English vs. Uzbek

In Uzbek, case suffixes and possessive markers attach directly to the proper nominal with no intervening determiner slot: *Durdona-ning kitobi* (Durdona-GEN book.POSS — "Durdona's book"), *Toshkent-da* (Tashkent-LOC — "in Tashkent").(this Parliament) confirms that the proper nominal does not include its own determiner, the Uzbek case suffix (*Toshkent-da*) confirms that the proper nominal is the base to which relational morphology attaches — it is a nominal unit, not a self-contained referential phrase.

The one structural difference between the two languages is distributional rather than categorical: English relies on an external layer (the determiner) to build the NP, while Uzbek relies on suffixal morphology. In both cases, however, the proper nominal is the unit to which that additional material attaches, and in both cases the proper nominal itself — the unit that carries the proper name — is sub-phrasal. The cross-linguistic parallel thus strengthens the nominal level hypothesis: the nominal is the target not because English NP syntax requires it, but because it is the maximal unit that can be extended by relational morphology in both analytic and agglutinative languages.

Discussion

What the Results Mean for the NP Analysis

The findings reported in Section 4 are difficult to reconcile with the self-determining NP analysis of Quirk et al. (1985) and Huddleston & Pullum (2002). If proper names were full NPs — saturated referential phrases that carry their own implicit definiteness — they could not: (a) appear without a determiner as pre-modifiers while still functioning as proper names; (b) freely accept any compatible determiner in discourse contexts; or (c) participate in close apposition as part of a composite nominal under a single external determiner. All three of these patterns are attested, frequent, and systematic in our dataset. The NP analysis predicts none of them. The nominal level analysis predicts all of them.

Addressing the Circularity Problem

A potential objection to any nominal level analysis is circularity: if the proper nominal is defined as the unit that behaves like a nominal, the claim that proper names are nominals becomes vacuously true by definition. Our three formal diagnostics, proposed in Section 3.3, are intended to prevent this. Each diagnostic is independently applicable without reference to the conclusion to be reached: typographic marking can be assessed by inspection; morphological reduction can be identified by comparison with documented longer forms; syntactic distribution (pre-modifier function, external determiner acceptance, apposition participation) can be measured corpus-statistically.

Institutionalization, Lexicalisation, and the Role of Typography

The results also illuminate the relationship between capitalisation and what Anderson (2007) calls institutionalisation — the process by which a linguistic unit acquires a stable, socially recognised referent. Our Uzbek data show that capitalisation in informal digital registers drops to 61% overall but remains highest (79%) for personal names (anthroponyms), which are the most socially entrenched proper names. This gradient is consistent with Brinton & Traugott's (2005) model of lexicalisation as a cline rather than a binary threshold.

The digital evidence extends this account in an important direction. Hashtags, usernames, and URL-syntax forms function as proper names without conventional capitalisation because they possess formal distinctiveness by other means. This suggests that what the nominal level hypothesis actually requires is not capitalisation specifically but formal alterity — any feature that sets the unit apart from the surrounding text and marks it as a specially institutionalised form. Capitalisation is the default realisation in conventional writing; the # symbol, the @ prefix, and URL punctuation are its functional equivalents in digital registers.

Limitations and Future Directions

Three limitations deserve explicit acknowledgement. First, the English dataset is a reanalysis of previously published corpus material rather than a newly assembled corpus; independent replication with a distinct dataset is desirable. Second, the Uzbek dataset ($n = 43$) is deliberately modest and intended as a pilot test of the cross-linguistic hypothesis rather than a definitive account of Uzbek proper nominal morphosyntax. A full-scale study of Uzbek onomastics — particularly the behaviour of anthroponyms in formal and informal registers — would be a valuable extension of this work. Third, the study does not address the semantic dimension of the proper nominal hypothesis — specifically, whether the nominal level corresponds to a distinctive semantic type (rigid designation, in Kripke's (1980) sense) or whether form and meaning are independently variable. These are questions for subsequent research.

Conclusion

This article has argued, on the basis of 187 English and 43 Uzbek proper nominal tokens, that proper names are best analysed as nominals — sub-phrasal units that occupy the structural level between the bare noun and the full noun phrase. Three formal diagnostics — typographic marking, morphological reduction toward a monolexical nominal profile, and syntactic distribution consistent with nominal but not NP status — consistently identify the same structural level in both an analytic language (English) and an agglutinative language (Uzbek). The cross-linguistic convergence supports the claim that the nominal level is not an English-specific artefact but a language-universal target for proper names.

The analysis makes four specific contributions to existing accounts. First, it provides operationalisable formal diagnostics for the nominal level, addressing the circularity risk in purely functional definitions. Second, it rejects the self-determining NP analysis with corpus-statistical evidence for pre-modifier function, external determiner acceptance, and composite-nominal apposition — three patterns that the NP analysis cannot explain. Third, it extends the nominal level hypothesis to Uzbek, showing that agglutinative case morphology attaches to the proper nominal in exactly the way the hypothesis predicts. Fourth, it frames capitalisation as one realisation of a broader formal-alterity requirement and shows that digital diacritics

serve the same function in new-media registers.

The proper name, on this account, is neither a noun nor a noun phrase. It is a nominal — a structural unit defined not by a single necessary condition but by a cluster of co-occurring formal, morphological, and syntactic properties (Langacker, 1991). Recognising this changes how we model proper name syntax, how we interpret determiner variation in proper name constructions, and — as the Uzbek data begin to show — how we approach proper names in typologically diverse languages.

References

1. Acuña-Fariña, J. C. (2009). *The Puzzle of Apposition*. Universidad de Santiago de Compostela.
2. Acuña-Fariña, J. C. (2012). The grammars of close apposition. *Journal of English Linguistics*, 40(3), 217–261.
3. Algeo, J. (1973). *On Defining the Proper Name*. University of Florida Press.
4. Allan, K. (1980). Nouns and countability. *Language*, 56(3), 541–567.
5. Anderson, J. M. (2007). On the grammatical status of names. *Language*, 83(4), 753–783.
6. Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. Longman.
7. Brinton, L., & Traugott, E. C. (2005). *Lexicalization and Language Change*. Cambridge University Press.
8. Caleffi, P.-M. (2015). The "hashtag": A new word or a new rule? *Skase Journal of Theoretical Linguistics*, 12(2), 46–69.
9. Gary-Prieur, M.-N. (1994). Le nom propre constitue-t-il une catégorie linguistique? *Langue française*, 102, 4–25.
10. Gil-Vilaboa, V. (2019). *Quirky NPs with Special Reference to Short Appositions* [Doctoral dissertation]. Universidad de Santiago de Compostela.
11. Huddleston, R., & Pullum, G. K. (2002). *The Cambridge Grammar of the English Language*. Cambridge University Press.
12. Jakobson, R. (1971). *Word and Language*. De Gruyter.
13. Keizer, E. (2007). Close appositions. In *The English*

Noun Phrase: The Nature of Linguistic Categorization (pp. 155–208). Cambridge University Press.

14. Kripke, S. (1980). Naming and Necessity. Harvard University Press.
15. Lakoff, G., & Johnson, M. (1980). Metaphors We Live By. University of Chicago Press.
16. Langacker, R. W. (1991). Foundations of Cognitive Grammar, Vol. 2: Descriptive Application. Stanford University Press.
17. Leroy, S. (2004). Entre identification et catégorisation: l'antonomase du nom propre en français [Doctoral dissertation]. Université Paul Valéry Montpellier III.
18. Mignot, E. (2018). La morphologie du nom en anglais — vers une sémantique des parties du discours [HDR Monograph]. Sorbonne Université.
19. Mill, J. S. (1843). A System of Logic. Longmans, Green.
20. Nyström, S. (2016). Names and meaning. In C. Hough (Ed.), The Oxford Handbook of Names and Naming (pp. 39–51). Oxford University Press.
21. Philippe, M. (2020). Le nominal propre. Étude du nom propre en anglais [Doctoral dissertation]. Sorbonne Université.
22. Philippe, M. (2023). Morpho-syntaxe sémantique des nominaux propres. Corela, HS-40. <https://doi.org/10.4000/corela.16236>
23. Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). A Comprehensive Grammar of the English Language. Longman.
24. Van Langendonck, W. (2007). Theory and Typology of Proper Names. Mouton de Gruyter.
25. Van Langendonck, W., & Vande Velde, M. (2016). Names and grammar. In C. Hough (Ed.), The Oxford Handbook of Names and Naming (pp. 17–38). Oxford University Press.
26. Vaxelaire, J.-L. (2006). Le nom propre en contexte. In D. Blampain et al. (Eds.), Mots, termes et contextes (pp. 217–226). Archives Contemporaines.
27. Wilmet, M. (1995). Nom propre et ambiguïté. *Langue française*, 92(1), 113–124.