The OSCAR PUBLISHING Services
American Journal Of Social Sciences And Humanity Research

VOLUME Vol.06 Issue04 2026
PAGE NO. 198-207
DOI 10.37547/ajsshr/Volume06Issue04-47

Codification of Language Norms and Its Linguistic Criteria: A Comparative Study of Uzbek And English

D.Sh. Egamnazarova
Associate Professor, Head of the Department of English Philology, Andijan Branch of Kokand University, Uzbekistan

Received: 18 March 2026; **Accepted:** 06 April 2026; **Published:** 30 April 2026

Abstract: Codification is commonly thought of as the point at which a language acquires fixed forms, recorded in dictionaries, grammars, spelling rules and official documents. The forms it fixes, however, continue to depend on usage, on the communicative needs of speakers and on whoever has the authority to choose between competing variants. The article examines this tension by comparing Uzbek and English, two languages whose routes to a standard could hardly be more different. Uzbek has acquired an explicit legal and administrative framework for codification, most visibly in matters of script, spelling, terminology and the public use of the state language. English has never had an academy, in Britain or in the United States; its norms were fixed piecemeal by grammarians, lexicographers, schools, publishers and editors and, more recently, by corpus-based description. Drawing on language-planning theory, recent research on standardization, studies of both languages and Uzbek legislative texts, the article separates codification from standardization in the wider sense and asks what makes a codificatory decision linguistically justified. Neither the force of law nor frequency of use proves sufficient in either language. Durable norms emerge where several conditions are met at once: compatibility with the language system, attestation, functional adequacy, semantic precision, acceptance by users, continuity with earlier practice, ease of learning, digital interoperability and openness to revision. These conditions are grouped into a four-layer model of legitimacy (structural, empirical, functional-social and implementation legitimacy). Codification, on this account, regulates variation and does not abolish it. The two languages turn out to require broadly the same kinds of evidence, although they weight them differently.

Keywords: Codification; language norm; standardization; Standard Uzbek; Standard English; orthography; language planning; linguistic criteria.

Introduction: The term codification looks straightforward enough. It seems to name a technical procedure: a spelling system is approved, a grammar is written, a dictionary marks certain forms as preferred, a terminology commission recommends an equivalent for a foreign term. Each of these acts, though, rests on an earlier and harder decision. Speakers use several forms; one of them has to be chosen as the public reference form, and someone must decide on what evidence, for which domains and how binding the choice is to be. That decision is taken by institutions, yet its substance is linguistic, sociolinguistic and, in the end, epistemological.

Early language-planning research treated codification as one component of the process through which a variety is developed into a standard. Haugen (1966) distinguished selection, codification, elaboration and acceptance and regarded them as interdependent dimensions; he did not present them as a strict sequence of stages. Cooper (1989), Kaplan and Baldauf (1997) and Spolsky (2004) shifted the emphasis to institutions, ideologies and the people who actually carry out planning, through whom norms come to carry authority. Recent work has moved further still from the top-down picture. Standardization is now seen as historically variable, with codifiers, linguists, official bodies, model writers and ordinary speakers sometimes pulling in the same direction and sometimes not (Ammon, 2003; Ayres-Bennett & Bellamy, 2021). Seen from this angle, codification does not create correctness. What it does is stabilize

selected forms in public reference works that can be taught, consulted, applied and, when necessary, revised.

METHODS

The Uzbek documents examined are the Law "On the State Language" in its revised post-independence version; the 1993 law introducing a Latin-based Uzbek alphabet; the 1995 law amending that alphabet; Cabinet of Ministers Resolution No. 339 of 24 August 1995, which approved the main rules of Uzbek orthography; Presidential Decree PF-6084 of 20 October 2020 on the further development of Uzbek and the improvement of language policy; Presidential Resolution PQ-308 of 20 October 2025 on additional measures for developing the state language; and Cabinet of Ministers Resolution No. 437 of 7 August 2026 on the effective organization of official work in the state language. These texts are read as evidence of institutional codification and its implementation, not as linguistic studies.

For English, the sources are historical and present-day studies of the codification of Standard English. Particular attention is paid to the grammars and dictionaries of the eighteenth century, to the tradition of the Oxford English Dictionary, to the absence of a language academy in Britain and the United States and to the emergence of several codified national varieties (Nevalainen & Tieken-Boon van Ostade, 2006; Ross, 2004; Hickey, 2012a; Tieken-Boon van Ostade, 2012, 2017). Huddleston and Pullum (2002) represent modern descriptive grammatical reference, which carries no prescriptive force in law.

Three domains are compared: (1) writing and orthography, (2) grammar and (3) lexicon and terminology. Each source was examined with two sets of questions in mind. The first concerns the codification ecology: who holds authority, through which instrument, over which domain and how far that authority can be enforced. The second concerns linguistic justification, namely the explicit or implicit criteria for preferring one form, variant or written representation over another. The criteria that recur are then grouped into four layers of legitimacy: structural, empirical, functional-social and implementation legitimacy.

The analysis is interpretive, not statistical. No new frequency data are presented, and the examples serve as illustrations without any claim to exhaust either language. Where usage is discussed, the argument rests on published research or on the documented status of codified variants. All materials used are publicly available.

In the modern history of Uzbek, orthographic codification has been unusually conspicuous. Fierman (1991) showed that language planning in twentieth-century Uzbekistan was inseparable from script policy, from lexical planning and from the shifting political functions assigned to the language. Independence added a further layer. The 1993 law introduced a Latin-based alphabet, and the 1995 amendment altered it substantially: the special letters of the 1993 version gave way to digraphs and to letters marked with an apostrophe-like sign (sh, ch, ng, o‘, g‘). Cabinet of Ministers Resolution No. 339 of the same year then approved the main orthographic rules for the new script. The sequence is significant because within a short period it brought together three levels that in many languages are handled by different bodies: the design of the alphabet, the regulation of spelling and implementation by the state.

Legislating an alphabet does not by itself settle questions of spelling. An alphabet supplies grapheme; an orthography governs how they represent words and morphemes and how capitals, compounds, affixes and line breaks are handled. Resolution No. 339 accordingly goes well beyond an inventory of letters and contains rules on vowels and consonants, on the spelling of stems and suffixes, on words written solid, hyphenated or separately, on capitalization and on word division. Uzbek is thus a useful case for distinguishing graphization from orthographic stabilization in the full sense. A law can establish the graphic system; the norm begins to work only when the spelling conventions are consistent with one another and simple enough to be taught and applied routinely.

Socially, the transition has been slow. Uzman (2010) places the romanization of the 1990s within the longer history of script change in the region, while Narmatova and Abdurakhmanova (2022) draw attention to the persistence of two scripts and to the practical difficulties of implementation. The case illustrates a general point: official selection and actual normalization move at different speeds. A script chosen by law can live alongside an older one for decades, the older script being kept alive by the stock of printed books, by what readers are able to read, by archives and by habit. Whether a rule exists is therefore only part of the question. Its uptake, the cost of the transition and its capacity to bridge generations and domains have to be weighed as well.

The structure of Uzbek sets limits on orthographic codification. As an agglutinative Turkic language, Uzbek builds word forms largely through suffixation, and its spelling rules must keep the relation between stem and affix predictable without moving written forms too far from present-day pronunciation. A workable rule balances morphological constancy against

phonological transparency, and the current rules do so in a recognizable way. Kitob keeps its final b in writing although it is pronounced with [p], so that the stem stays constant in kitobi and kitobim; the alternation in yurak – yuragi or qishloq – qishlog‘i, by contrast, is written as it is pronounced. Spelling rules of this kind preserve the distinctions readers need, treat recurrent morphological patterns consistently and avoid exceptions that serve no purpose. The considerations involved are linguistic, and they remain so when the rule acquires legal force.

Frequent changes of script also give continuity unusual weight. Where orthography has been stable for a long time, continuity can be treated as a background condition; in Uzbek it is an active variable of policy. A spelling may be perfectly regular and still impose heavy conversion costs on textbooks, signage, databases and older printed material. Too much deference to inherited forms, on the other hand, can slow down the consolidation of the new script. The codifier has to weigh the formal merits of a solution together with the cost of transition, established literacy practices and long-term consistency.

For most users, codification takes concrete form in reference works far more often than in legislation. Dictionaries mediate between expert analysis and the everyday decisions people make about words. The multi-volume O‘zbek tilining izohli lug‘ati is the most important resource of this kind for Uzbek. It does much more than list words: its definitions, labels, sense divisions, phraseology and usage notes together determine what counts as established vocabulary. A dictionary does not need legal force to have normative influence. Its authority comes from the institution behind it, the expertise of its editors, its coverage and the fact that people keep consulting it.

Lexical codification is harder than orthographic codification. Vocabulary grows quickly, and competing forms can remain semantically viable side by side for a long time. New terms arrive through borrowing, calquing, derivation from native resources, extension of meaning or deliberate coinage, and each route raises its own problems. A borrowing may already be transparent to specialists across languages and still be awkward in Uzbek phonology or morphology. A new Uzbek coinage may be well formed and yet unfamiliar to users. A calque may look transparent while importing semantic interference from the source language. Recent Uzbek history offers instructive cases: of the purist coinages proposed in the early 1990s, bekat ‘stop, station’ took root, whereas tayyoragoh ‘airport’ never displaced aeroport in general use. Such alternatives cannot responsibly be decided by a blanket preference for either native or international vocabulary. Each has to be tested along several dimensions.

In terminology, at least five questions decide the matter. Does the candidate term denote the intended concept without unwanted ambiguity? Does it conform to Uzbek phonotactics and derivational morphology? Can further terms be derived from it productively? Is it already established among specialists, or can it be learned at reasonable cost? Will it remain stable across teaching, administration and professional documentation? These questions explain why, in specialized domains, semantic precision and morphological productivity may count for more than raw frequency.

Recent policy has made the infrastructure of terminology more visible. Decree PF-6084 (2020) set long-term goals for the development of Uzbek, among them the compilation of linguistic and terminological dictionaries, integration with information technology and wider use of the state language in specialized fields. Resolution PQ-308 (2025) puts the emphasis on digital mechanisms, including electronic resources and an electronic register of newly approved terms. Whatever institutional shape these platforms finally take, their linguistic significance is considerable. A register of terms can turn codification from occasional publication into a reference that is versioned, searchable and continuously updated. The revision history of a norm then becomes part of its quality: users need to know which form is currently recommended and also whether, and why, an earlier recommendation was changed.

Digital codification also raises the value of machine-readable lexicography. A printed dictionary can live with inconsistencies in alphabetical ordering or character encoding that a search engine cannot. Uzbek supplies a telling example. The letters o‘ and g‘ are written in practice with several different characters (the modifier letter turned comma U+02BB, the right single quotation mark, the ASCII apostrophe and even the grave accent), and a search for a word spelled with one of them may fail to retrieve the same word spelled with another. In digital environments, spelling decisions affect tokenization, search equivalence, spell-checking, morphological analysis and the recognition of proper names. Interoperability has ceased to be an external convenience; without it a codified norm cannot work at scale.

A norm that exists only in a reference book is, in sociolinguistic terms, only partly codified. Implementation gives it an institutional life. PF-6084 explicitly links the development of the language to education, professional domains, lexicography and

digital technology, and in doing so treats standardization as continuing infrastructure, not as a historical event already completed.

The measures of 2025 and 2026 continue in the same direction. PQ-308 introduced further mechanisms for developing the state language, and Resolution No. 437 of 7 August 2026 tightened the organization of work in the state language in public institutions, building on existing provisions for the linguistic examination of draft legal acts. The connection with codification lies in the feedback between the norm and institutional texts. Legal and administrative writing is a high-stakes domain, where ambiguous terms, inconsistent naming or unstable spelling can have consequences that go well beyond style. Linguistic examination works as a mechanism of implementation, since it checks whether codified norms are reproduced in the very texts that go on to shape public usage.

This institutionalization brings advantages and risks. Its main advantage is coordination: spelling, official terminology and administrative language can be aligned across agencies instead of being left to local conventions. The risk is overreach. If every variant is treated as an error simply because an institution prefers another form, codification loses sight of register, of innovation and of the evidence of usage. The greater the formal authority of a codifier, the more it needs transparent linguistic criteria and regular review.

For contemporary Uzbek the most defensible criteria are thus consistency with the phonological and morphological system; orthographic regularity; continuity with established literate practice, wherever continuity spares unnecessary cost; broad attestation appropriate to the domain; semantic precision; derivational productivity; teachability; compatibility with digital text processing; and clarity about the scope and status of each recommendation. These criteria often pull in different directions, and it is in resolving such conflicts that codification turns into linguistic work instead of clerical work.

English shows that strong standard norms can arise without a state academy. Neither Britain nor the United States has ever had an official language academy, which sets English apart from languages with a central body formally empowered to regulate usage; proposals for one, Swift's in 1712 and John Adams's in 1780 among them, came to nothing (Ross, 2004). Yet English is anything but uncodified. Its spelling, grammatical conventions and lexicographic tradition are among the most thoroughly documented anywhere, and the difference lies in how authority is distributed.

Historians of English place a major phase of explicit codification in the eighteenth century. Johnson's A Dictionary of the English Language (1755) became a lasting point of reference, not least because it illustrated meanings with quotations from literary authors. Lowth's A Short Introduction to English Grammar (1762) represents a different impulse, the attempt to define and enforce grammatical correctness, frequently by criticizing the usage of esteemed writers. His best-known remark concerns the preposition placed at the end of a clause, which he accepted as natural in conversation and the familiar style but judged less graceful in elevated writing; later handbooks hardened this stylistic advice into a prohibition. Tieken-Boon van Ostade (2012, 2017) argues that the prescriptive tradition has to be understood within the social history of standardization and should not be caricatured as mere error-hunting. Grammars answered a demand for linguistic guidance created by rising literacy, social mobility and the prestige of educated usage.

Lexicographic documentation and grammatical prescription did not remain separate for long. Dictionaries influenced spelling and vocabulary, grammars influenced schooling and editorial practice, and publishing reinforced both. Some early prescriptions lost their authority over time, while others became deeply entrenched. Webster's An American Dictionary of the English Language (1828), with spellings such as color and center, shows how a single reference work could consolidate a national orthographic tradition. Codification in English was cumulative and selective: not every rule proposed by a prescriptive grammarian entered Standard English, just as not every attested variant found its way into the reference works.

The Oxford English Dictionary gave lexical description a historical foundation. Published in fascicles from 1884 and completed in 1928, it established a model in which the record of use over time was itself a source of authority (Gilliver, 2016). A terminology commission selecting a preferred technical term works on quite another principle. The authority of the OED rests on historical evidence, even though users often treat its decisions on inclusion as normative.

Today English is standardized through many interacting institutions: the major dictionaries and grammars, school systems, academic and trade publishers, professional style guides, examination bodies, national media and digital platforms. None has universal jurisdiction. One and the same form may be preferred by one publisher, standard in American English, less common in British English and entirely acceptable in international academic prose. Codification here is

plural in two senses, institutional and geographical.

Hickey's edited volume Standards of English (2012a) treats Standard English as a family of codified varieties with common foundations and national or regional differences, rather than as a single undifferentiated global code. Once several standards exist, the logic of codification changes. If both color and colour are codified, the rule cannot be that one is right and the other wrong; it is that each belongs to a coherent orthographic convention and that a text should be consistent within the convention it adopts. The object of evaluation becomes the system, and the isolated form loses its importance.

The principle extends beyond spelling. National standards differ in vocabulary, in punctuation and in certain grammatical preferences: British English readily treats collective nouns such as government or team as plural, for instance, while edited American prose overwhelmingly treats them as singular. Variation has limits, however. A standard variety depends on wide intelligibility, stable written conventions and a shared grammatical core. Pluricentricity therefore calls for bounded tolerance, meaning that codification must specify which alternatives are legitimate standard variants and which are restricted, nonstandard or marked for context.

Modern descriptive grammar has complicated the older prescriptive model, and usefully so. Huddleston and Pullum (2002) separate grammatical description from the traditional prohibitions of usage guides and examine systematically the contrasts between formal and informal constructions, the competing analyses and the patterns actually found. Their treatment of the split infinitive, which they show to be fully grammatical and often the clearest option available, is typical. A grammar of this kind will not tell a government office which construction to require, but it gives codifiers, editors and teachers a more reliable account of the language system. Descriptive scholarship in this way acts as an evidential check on prescription.

Corpus methods have made that check considerably stronger. Usage can now be examined by genre, region, period and degree of formality instead of being inferred from individual intuition. Peters (2019) shows how much norms matter within the diverse ecology of World Englishes, where corpus evidence brings to light what is recurrent in particular communities. Frequency is still not the same thing as standardness, but distributional data allow codifiers to separate stable cross-register patterns from local or short-lived innovations.

Because no single body makes the rules for English, its criteria of codification are mostly implicit in lexicographic and editorial practice. Several of them recur nonetheless. One is attestation: a dictionary normally requires evidence that a word or sense is genuinely in use. Distribution is another, since a form found only in a narrow subculture is described differently from one established across formal prose. Historical and national convention matters as well; centre and center are not judged by phonological transparency alone, because each belongs to a stable spelling tradition. Functional clarity counts, especially in edited prose. Finally there is register: a construction can be fully grammatical and still unsuitable for a particular formal context.

English also shows the limits of a purely systemic criterion. Its spelling contains a great many inherited irregularities (knight, colonel, though), and codifying only the most phonographically regular spellings would break continuity on an enormous scale. In English orthography, historical stability therefore outweighs surface regularity. Grammatical codification can rely more heavily on structural and distributional analysis, while lexical codification depends largely on attestation, semantic differentiation and chronological evidence.

The resulting system is decentralized but by no means free of criteria. Authority arises from convergence: when the major dictionaries, the descriptive grammars, educational practice and edited usage agree, a norm can be robust without any legislation behind it. When they disagree, the appropriate outcome is often a usage note specifying region, register, frequency or editorial preference, and seldom a categorical prohibition. For comparative work on codification the lesson is an important one. In some cases the most accurate statement of the norm is a set of labeled alternatives.

The first layer asks whether a candidate norm is compatible with the productive organization of the language. The criterion is strongest where rules can be generalized. Because Uzbek is agglutinative, the codification of stems, suffixes and new terms has to be assessed against morphological productivity. A lexical item that does not readily take ordinary Uzbek inflectional or derivational suffixes imposes recurrent costs in processing and teaching. Orthographic rules, likewise, gain in strength when they represent recurrent relations between stem and affix predictably.

English is a more mixed case. Structural compatibility is decisive in syntax and morphology, but it cannot by itself justify a spelling. Many conventional spellings do not transparently represent present-day pronunciation, and a structurally simpler spelling such as nite or thru is not for that reason a better codified

one, since it clashes with the wider conventional system and carries high costs in continuity. The weight of systemicity thus depends on the domain: phonographic regularity can count for a great deal in one orthographic history and for much less in another.

Structural legitimacy also protects against a common mistake in terminology planning, that of judging a proposed term only in its base form. Terms enter paradigms, and a candidate should be tested in the plural, in case forms, in derivation, in compounds, in abbreviation and within phrases. A term that looks acceptable in isolation may yield clumsy derivatives. The test matters especially for Uzbek, where derivational morphology is a main resource for building specialized vocabulary, but it applies equally to English technical word families.

A codifier who ignores usage is likely to produce norms that speakers perceive as artificial. Still, “actual usage” is too vague to serve as a criterion until its distribution has been specified: who uses the form, in which genres, how often, over what period and in competition with which alternatives. Where the target norm is administrative or technical, a million informal tokens from social media do not necessarily outweigh stable use in legal or scientific writing.

English codification has gained a great deal from large corpora and searchable historical databases, which make it possible to separate national, regional, diachronic and register-based patterns of distribution. The methodological lesson can be transferred to Uzbek. Corpus evidence should enter the review of spelling, grammar and terminology, and its job there is to map the social and functional distribution of forms, which is a good deal more than counting them. For a language still consolidating its script, corpus design must also take script into account; otherwise searches will systematically underrepresent older usage and usage in the parallel script.

Empirical legitimacy, then, is not a matter of majority rule. A less frequent form may be the better choice if it is semantically clearer, structurally more productive or already established in the relevant specialist community. Conversely, an officially preferred term may need revising if long-term evidence shows that it has not been taken up and that a rival form now dominates in the professional domains concerned. Codification should respond to evidence without being driven mechanically by frequency.

Language norms exist because recurrent communication benefits from shared expectations, which makes functional adequacy a primary criterion. In orthography it serves quick recognition and unambiguous search; in grammar, predictable interpretation; in terminology, conceptual precision. The higher the stakes of a domain, the more a stable reference is worth. Law, medicine, engineering and education demand far greater terminological consistency than casual conversation does.

Social acceptance belongs to the same layer, since communicative efficiency depends in part on recognizability. A technically flawless form that speakers reject may cause more friction than the problem it was meant to solve. Classic language-planning theory gave implementation and acceptance a central place for precisely this reason. Publishing a rule completes only the expert part of codification; the rule takes effect once users can recognize it, learn it and reproduce it.

Acceptance is built through different institutions in the two languages. Uzbek can draw on state education, public administration and official terminology bodies. English relies more on school conventions, publishing, dictionaries, professional editing and the prestige of established written practice. The underlying requirement is the same in both: a successful reference norm must win enough recognition that compliance no longer depends on constant pressure.

Functional-social legitimacy also calls for sensitivity to variation. National variants of English are often fully standard, and suppressing one in favor of another would reduce communicative adequacy; the appropriate rule is consistency within the convention chosen. Uzbek faces a different but related problem in the coexistence of scripts and inherited forms, where some variants may need to be labeled as transitional before they can be excluded. Codification can accordingly work with a graded scale (preferred, permitted, transitional, regional, informal) in place of a binary division into correct and incorrect.

The fourth layer is often underestimated because it looks extralinguistic. Teachability, editorial consistency and the behavior of technology nevertheless feed back into linguistic practice. A spelling rule that users cannot apply reliably will generate variants. A character sequence that is encoded inconsistently will fragment search results. A terminological decision that is not available in accessible databases will be displaced by local alternatives.

The current policy environment in Uzbekistan makes this especially clear. PF-6084 ties the development of the language to digital technology, and PQ-308 extends the same approach to digital lexicographic and terminological resources. The point of such measures goes beyond putting dictionaries online. Digital publication makes possible a different model of codification, with versioning, structured metadata,

cross-references, machine-readable entries, quick correction and traceable revision. These possibilities increase transparency, provided they are used to show evidence and scope and not merely to reproduce a printed list on screen.

Technological interoperability ought to be tested before a spelling or a term is institutionalized on a large scale. For orthography this means reliable keyboard input, correct Unicode representation, search normalization, alphabetical sorting and compatibility with common text-processing software. The o‘/g‘ problem described above would probably have been far smaller had a single character been specified and supported on standard keyboards from the outset. For lexicography the test covers lemmatization, morphological search and stable identifiers for senses and terms. For grammar, the question is whether digital writing tools can represent a rule without flattening legitimate variation.

English has a far more stable script, but digital implementation affects the circulation of its norms as well. Spell-checkers, predictive text, grammar assistants and search engines increasingly act as everyday codifiers. Their suggestions may reinforce one national convention, mislabel a regional standard or turn a stylistic preference into an apparent grammatical prohibition. Earlier models of standardization did not anticipate this kind of norm authority. Algorithmic mediation makes transparency all the more pressing, since users need to be able to tell a grammatical error from a house-style preference and a national variant from a nonstandard form.

The criteria discussed above should not be added up into a score, as if the candidate with the highest total automatically prevailed. Some of them work as thresholds. A technical term that is semantically wrong should not be codified merely because it is frequent, and a spelling that standard digital systems cannot represent reliably should not be preferred merely because it has historical appeal. Others are matters of balance: continuity may be traded against regularity, international recognizability against morphological integration, strict uniformity against pluricentric legitimacy.

In practice a decision can be approached through four questions asked in order. The first is whether the candidate is structurally viable; if it is not, the burden of justification becomes heavy. Next comes credible evidence of use in the relevant domain, without which a form should normally remain experimental or merely proposed. A form that passes these two tests must then improve, or at least preserve, functional clarity and social usability, or there is little communicative reason to codify it. Last, the decision must be capable of being implemented, searched, taught and revised; a norm that fails here is likely to exist only on paper.

Such a procedure lets languages weight evidence differently while remaining comparable. Uzbek may at present give more weight to script consistency, morphological integration and coordinated implementation; English, to national convention, historical continuity, corpus distribution and register labeling. In both, durable codification needs all four kinds of legitimacy.

DISCUSSION

The comparison calls two common assumptions into question. According to the first, centralized codification is inherently prescriptive and decentralized codification inherently descriptive. Uzbek certainly has stronger formal mechanisms than English for setting alphabetic, orthographic and administrative norms. These mechanisms, however, depend more and more on scholarly lexicography, linguistic examination and digital evidence. English has no academy in Britain or the United States, yet editorial and educational institutions can be strongly prescriptive within their own domains. The real difference lies in how authority is distributed and how far it can be enforced.

The second assumption holds that variation disappears in a mature standard. English plainly refutes it, with several national forms that are codified, stable and socially recognized. Uzbek shows, for its part, that variation may persist through a historical transition even after a public reference norm has been selected. Standardization is better modeled as the differentiation of variation. Some alternatives are raised to standard status, some are marked for region or register, some are kept provisionally, and some are rejected because they disturb structural or communicative predictability.

For Uzbek codification this perspective has direct consequences. The country’s extensive legal and institutional infrastructure gives it the chance to make codification unusually transparent. Future decisions on spelling and terminology could be accompanied by brief public notes setting out the alternatives considered, their distribution in corpora, the structural arguments, the domain of application and the date of the next review. Documentation of this kind would not weaken authority. It would separate linguistic justification from institutional preference and make later revisions more accountable.

A second consequence concerns the architecture of dictionaries. The value of a digital dictionary lies less in speed of access than in its ability to link spelling, morphology, meaning, terminology and evidence of

use. For Uzbek, a unified digital environment could connect each lemma with its inflectional behavior, its approved spelling, variant forms, domain labels, citations, terminological status and machine-readable metadata. The orientation of policy toward digital integration since 2020, together with the 2025 measures on digital resources, makes such a project timely both technically and institutionally. It would also reduce the risk of different agencies maintaining divergent term lists with no shared identifiers or revision histories.

A third consequence concerns corpus design. Codification needs data from the domains for which norms are intended. A general web corpus is likely to overrepresent informal journalism, duplicate pages and faulty encoding. A decision on legal terminology needs legal and administrative texts, an academic term needs disciplinary writing, and a decision on public spelling needs broad, balanced evidence. For Uzbek, materials in both scripts should be searchable through normalization layers, so that Latin and Cyrillic evidence can be compared without obscuring the historical fact of script choice. The method is familiar from corpus linguistics; building it into the institutional routine of codification would mark a real shift from intuition toward evidence that can be audited.

English offers a complementary lesson about plural standards: the codifier does not always have to pick a winner. Where two forms are stable, widely recognized and functionally equivalent within separate standard traditions, labeling describes the language more accurately than prohibition. Uzbek terminology and orthography are not identical cases, but the principle carries over. Costly variation needs to be distinguished from structured variation, and a variant that tracks region, genre, generation or specialist field may call for metadata, not eradication.

The model also clarifies the role of historical continuity. English spelling shows that formal irregularity can be outweighed by stability across an immense body of text. The twentieth-century history of Uzbek script shows the other side of the same principle, namely that continuity can itself be plural when successive generations inherit different graphic conventions. A reader schooled in Cyrillic and one schooled in the Latin script do not share the same written past. The codifier must therefore ask with what, for whom and in which domain continuity is to be maintained, and cannot invoke tradition as if it were a single, uniform good.

The two languages converge most clearly in the digital domain, where writing tools increasingly stand between codified resources and ordinary users. Spell-checkers, grammar checkers, predictive keyboards, search engines and large language models issue recommendations on a scale no printed grammar ever reached. Trained on noisy or unbalanced data, they may amplify unstable variants; built on rigid rule sets, they may stigmatize legitimate alternatives. Codification in the twenty-first century accordingly needs a machine-facing layer, with structured data, explicit variant labels and transparent update policies.

The relationship between codifiers and users changes as a result. Printed codification moved in one direction, from the institution to the user. Digital systems produce continuous feedback: searches reveal retrieval failures, error logs reveal recurrent spelling problems, corpora show whether terms are being taken up, and user queries point to ambiguous dictionary entries. Codification can thus become cyclical, with selection followed by publication, implementation, monitoring and review. The cycle resembles the classic planning models, though it draws on far richer evidence.

The study has its limitations. It compares ecologies of codification and offers no matched quantitative corpus analysis. English is far more widely dispersed than Uzbek, and “Standard English” covers several national traditions that would each repay separate study. Uzbek, too, has regional and stylistic variation that cannot be reduced to state language policy. The proposed criteria are best regarded as an analytical framework still to be tested on specific cases, such as competing Uzbek technical terms, contested spellings or grammatical variants across the national standards of English. Parallel corpora, expert judgments and tests of user comprehension would be the natural means of doing so.

CONCLUSION

Codification is sometimes described as the moment at which linguistic variability is turned into rule. The Uzbek–English comparison points to a more exact formulation: codification is the process by which selected forms become publicly referential, that is, teachable, searchable, citable, reproducible and, where necessary, enforceable by institutions. The rule is the most visible product of this process, and by no means its starting point.

The two languages arrive at that product by different institutional routes. In Uzbek, a relatively explicit legal and administrative framework can connect script, orthography, the use of the state language, terminology and linguistic examination through formal policy. English developed strong norms through a dispersed ecology of dictionaries, grammars, schools, publishers, national usage traditions and descriptive research, with no academy in either Britain or America. These differences shift the weight of particular criteria,

above all in orthography and implementation, while leaving a common evidential core intact.

In both languages durable codification rests on four layers of legitimacy. Structural legitimacy concerns whether a form fits the productive organization of the language. Empirical legitimacy concerns whether it is genuinely attested and how it is distributed across communities and registers. Functional-social legitimacy concerns whether it improves clarity, precision and shared communicative expectations without suppressing meaningful variation. Implementation legitimacy concerns whether the norm can be learned, published, processed digitally, maintained and revised. The specific criteria follow from these layers: systemic compatibility, attested use, distribution, functional adequacy, semantic precision, continuity, learnability, technological interoperability and transparent institutional revision.

The main theoretical conclusion is that codification should not be equated with maximal uniformity. A strong standard distinguishes between kinds of variation and regulates each according to communicative need. In English the principle can be seen in the coexistence of national standard variants; in Uzbek it bears on the transition between scripts, on the development of terminology and on the relation between official preference and established usage. Codification does not freeze a language. It provides linguistic change with a stable public interface.

For Uzbek in particular, current work on digital dictionaries, terminological infrastructure and the administration of the state language opens the way to codification grounded in evidence, in which every major norm would be backed by structural analysis, balanced usage data and explicit procedures for revision. English shows the long-term value of extensive documentation and labeled variation, Uzbek the coordinating power of institutional planning. Taken together, the two cases suggest that a language norm carries the most lasting authority when institutional legitimacy and linguistic evidence support each other.

REFERENCES

1. Ammon, U. (2003). On the social forces that determine what is standard in a language and on conditions of successful implementation. Sociolinguistica, 17(1), 1–10. https://doi.org/10.1515/9783110245226.1
2. Ayres-Bennett, W., & Bellamy, J. (Eds.). (2021). The Cambridge handbook of language standardization. Cambridge University Press. https://doi.org/10.1017/9781108559249
3. Cooper, R. L. (1989). Language planning and social change. Cambridge University Press.
4. Egamnazarova D. S. Gap semantikasi masalasi //Scientific Bulletin. Series: Humanitarian Research. – 2021. – №. 1. – С. 120-127.
5. Fierman, W. (1991). Language planning and national development: The Uzbek experience. De Gruyter Mouton. https://doi.org/10.1515/9783110853384
6. Gilliver, P. (2016). The making of the Oxford English Dictionary. Oxford University Press. https://doi.org/10.1093/acprof:oso/97801992836 20.001.0001
7. Haugen, E. (1966). Language conflict and language planning: The case of modern Norwegian. Harvard University Press.
8. Hickey, R. (Ed.). (2012a). Standards of English: Codified varieties around the world. Cambridge University Press. https://doi.org/10.1017/CBO9781139023832
9. Hickey, R. (2012b). Standard English and standards of English. In R. Hickey (Ed.), Standards of English: Codified varieties around the world (pp. 1–33). Cambridge University Press. https://doi.org/10.1017/CBO9781139023832.002
10. Huddleston, R., & Pullum, G. K. (2002). The Cambridge grammar of the English language. Cambridge University Press. https://doi.org/10.1017/9781316423530
11. Johnson, S. (1755). A dictionary of the English language. W. Strahan for J. and P. Knapton, T. and T. Longman, et al.
12. Kaplan, R. B., & Baldauf, R. B., Jr. (1997). Language planning: From practice to theory. Multilingual Matters.
13. Lowth, R. (1762). A short introduction to English grammar, with critical notes. J. Hughs for A. Millar and R. and J. Dodsley.
14. Milroy, J., & Milroy, L. (2012). Authority in language: Investigating Standard English. Routledge.
15. Narmatova, S., & Abdurakhmanova, M. (2022). Language policy in Uzbekistan: The adoption of Uzbek language and the Latin alphabet. InContext, 2(3), 144–155. https://doi.org/10.54754/incontext.v2i3.34
16. Nevalainen, T., & Tieken-Boon van Ostade, I. (2006). Standardisation. In R. Hogg & D. Denison (Eds.), A history of the English language (pp. 271–311). Cambridge University Press.
17. Peters, P. (2019). Norms and standards in World

Englishes. In D. Schreier, M. Hundt, & E. W. Schneider (Eds.), The Cambridge handbook of World Englishes. Cambridge University Press.

18. Ross, N. J. (2004). Academies and attitudes. English Today, 20(3), 23–28. https://doi.org/10.1017/S0266078404003049

19. Rutten, G., & Vosters, R. (2021). Language standardization 'from above'. In W. Ayres-Bennett & J. Bellamy (Eds.), The Cambridge handbook of language standardization (pp. 65–92). Cambridge University Press. https://doi.org/10.1017/9781108559249.003

20. Schlyter, B. N. (1998). New language laws in Uzbekistan. Language Problems and Language Planning, 22(2), 143–181. https://doi.org/10.1075/lplp.22.2.03sch

21. Spolsky, B. (2004). Language policy. Cambridge University Press. https://doi.org/10.1017/CBO9780511615245

22. Tieken-Boon van Ostade, I. (2012). The codification of English in England. In R. Hickey (Ed.), Standards of English: Codified varieties around the world (pp. 34–54). Cambridge University Press. https://doi.org/10.1017/CBO9781139023832.003

23. Tieken-Boon van Ostade, I. (2017). Perspectives on standardization: From codification to prescriptivism. In L. J. Brinton (Ed.), English historical linguistics: Approaches and perspectives (pp. 276–302). Cambridge University Press. https://doi.org/10.1017/9781316286562.011

24. Uzman, M. (2010). Romanisation in Uzbekistan past and present. Journal of the Royal Asiatic Society, 20(1), 49–60. https://doi.org/10.1017/S1356186309990307

25. Webster, N. (1828). An American dictionary of the English language (Vols. 1–2). S. Converse.

26. O'zbekiston Respublikasi. (1993). Lotin yozuviga asoslangan o'zbek alifbosini joriy etish to'g'risida [Law No. 931-XII, 2 September 1993]. LexUZ.

27. O'zbekiston Respublikasi. (1995a). Lotin yozuviga asoslangan o'zbek alifbosini joriy etish to'g'risidagi qonunga o'zgartishlar kiritish haqida [Law No. 71-I, 6 May 1995]. LexUZ.

28. O'zbekiston Respublikasi. (1995b). Davlat tili haqida [Law of the Republic of Uzbekistan, revised edition, 21 December 1995]. LexUZ.

29. O'zbekiston Respublikasi Vazirlar Mahkamasi. (1995). O'zbek tilining asosiy imlo qoidalarini tasdiqlash haqida [Resolution No. 339, 24 August 1995].

30. O'zbekiston Respublikasi Prezidenti. (2020). Mamlakatimizda o'zbek tilini yanada rivojlantirish va til siyosatini takomillashtirish chora-tadbirlari to'g'risida [Decree PF-6084, 20 October 2020]. LexUZ.

31. O'zbekiston Respublikasi Prezidenti. (2025). Davlat tilini yanada rivojlantirishga oid qo'shimcha chora-tadbirlar to'g'risida [Resolution PQ-308, 20 October 2025]. LexUZ.

32. O'zbekiston Respublikasi Vazirlar Mahkamasi. (2026). Respublikada davlat tilida ish yuritishni samarali tashkil etishning qo'shimcha chora-tadbirlari to'g'risida [Resolution No. 437, 7 August 2026]. LexUZ.