



(REVIEW ARTICLE)



Skibidi in the Cambridge Dictionary: When does slang become legit?

Dinesh Deckker^{1,*}, Subhashini Sumanasekara² and Sree Lakshmi Ammanamanchi³

¹ Faculty of Arts, Science and Technology, Wrexham University, United Kingdom.

² Faculty of Computing and Social Sciences, University of Gloucestershire, United Kingdom.

³ Faculty Member at the University of Technology and Applied Sciences - Al Mussanah, Sultanate of Oman.

International Journal of Science and Research Archive, 2025, 16(03), 551–563

Publication history: Received on 01 August 2025; revised on 07 September 2025; accepted on 10 September 2025

Article DOI: <https://doi.org/10.30574/ijrsra.2025.16.3.2582>

Abstract

This article examines how digital slang gains legitimacy through its inclusion in authoritative dictionaries, focusing on Cambridge's 2025 additions of *skibidi*, *delulu*, and *tradwife*. Drawing on a narrative review of sociolinguistic theory, lexicographic practice, and recent computational corpus studies (2022–2025), the paper reframes slang legitimation as a filtering process rather than a linear trajectory. Classical frameworks (Saussure, Labov, Milroy, Hockett) are revisited in light of accelerated digital diffusion, where algorithmic amplification and weak-tie networks enable slang to spread globally within weeks. Findings highlight three core criteria for codification: sustained frequency, semantic stability, and cross-demographic dispersion. However, dictionary decisions also carry cultural and ethical weight, as inclusion can be perceived as endorsement, especially for ideologically charged terms. A comparative analysis across English, French, German, and Japanese dictionaries reveals divergent philosophies of linguistic authority, while a proposed practical framework provides lexicographers with tools for balancing responsiveness with caution. Ultimately, the study argues that digital slang survives codification only through endurance and broad intelligibility, underscoring the role of dictionaries as both descriptive record-keepers and cultural arbiters in an era of accelerated linguistic change.

Keywords: Digital Slang; Lexicography; Dictionary Codification; Sociolinguistic Legitimacy; Descriptivism Vs. Prescriptivism; Weak-Tie Diffusion; Semantic Stability; Cambridge Dictionary; Cross-Linguistic Comparison; Cultural Authority

1. Introduction

1.1. Language Change and Legitimacy

Language is never static. It is a living system continually reshaped by social interaction, cultural practice, and, increasingly, by digital technologies. Among the most visible markers of this dynamism is slang lexical innovation that not only reflects creativity and play but also indexes identity and contestation of norms. In traditional sociolinguistic accounts, slang was often treated as ephemeral or peripheral, surviving only in subcultural spaces unless it gradually diffused into mainstream registers. However, in the digital age, this trajectory has been profoundly accelerated and globalised.

Classical theory still offers useful lenses. Saussure's ([1916]/1983) distinction between *langue* and *parole* reminds us that systemic change often begins with individual departures from convention that, through repetition, sediment into the structure itself. Labov (1972) emphasised the role of social evaluation in determining which changes gain legitimacy, while Milroy (2002) demonstrated how weak-tie networks allow innovations to spread beyond localised communities. These perspectives remain foundational, but they were developed in pre-digital environments where diffusion was slower, geographically bounded, and mediated primarily through speech and print.

* Corresponding author: Dinesh Deckker

Today, slang does not merely circulate; it is algorithmically amplified, remixed, and made hyper-visible through global media infrastructures. This shift compels us to ask: under what conditions does such volatility crystallise into legitimacy, culminating in the codification of slang by authoritative dictionaries? The stakes extend beyond descriptive record-keeping. The inclusion of digital slang in a reference work such as the Cambridge Dictionary foregrounds broader debates about institutional authority, intergenerational tensions, and the politics of what counts as “real” language.

1.2. Digital Slang in the TikTok/YouTube Era

Nowhere is this acceleration more apparent than on platforms such as TikTok and YouTube, where a single sound clip, hashtag, or video trend can propel an obscure term into global circulation in days. Words like *skibidi*, *delulu*, and *tradwife* illustrate this dynamic: emerging first in tightly bound online communities, they rapidly migrated outward through algorithmic recommendation systems, remixable formats, and influencer amplification. Unlike earlier sociolinguistic models, these are not merely organic processes of spread but technologically scaffolded cascades.

Empirical studies have begun to map this landscape. Keidar, Opedal, Jin, and Sachan (2022) show that slang terms on digital platforms exhibit sharper frequency spikes and more rapid stabilisation of meaning than traditional neologisms, indicating a compressed life cycle of innovation. Ilbury, Grieve, and Hall (2024), analysing billions of social media posts, demonstrate how slang diffuses nationally through weak-tie networks, predicting offline adoption patterns. Taken together, such work underscores the need to rethink lexicographic timetables: if a term can achieve global currency within weeks, should dictionaries revise their historical model of waiting for long-term durability before codification?

This question is not merely academic. Lexicographers face mounting pressure to keep pace with digital language change, yet premature inclusion risks cluttering reference works with transient curiosities. The balance between responsiveness and caution has therefore become a defining challenge of twenty-first-century dictionary practice.

1.3. Descriptivism vs. Prescriptivism in the Digital Era

The tension between descriptivism and prescriptivism remains central to lexicography, but the digital age reshapes its contours. In principle, descriptivism views dictionaries as empirical records: entries follow usage, rather than editorial fiat. Cambridge’s 2025 additions of *skibidi*, *delulu*, and *tradwife* were justified through corpus attestation, cross-register spread, and “staying power” (Cambridge University Press and Assessment, 2025). This reflects the descriptivist claim that legitimacy is observed, not bestowed.

Yet public reception resists this framing. As VanEyck and Curzan (2025) show in their study of New York Times style manuals, institutional codification has long been understood as a prescriptive act shaping norms rather than merely recording them. That perception persists: surveys show users still treat dictionaries as arbiters of correctness (Burada and Sinu, 2023), and dictionary prefaces themselves reinforce authority through metaphors of guardianship (Ștefănescu and Minică, 2023).

Digital slang magnifies this paradox. Viral forms travel from niche to mainstream in weeks, forcing lexicographers to make rapid decisions under public scrutiny. Editors stress descriptivist neutrality, but media coverage reframes inclusion as cultural endorsement, especially when youth-driven or ideologically charged terms are at stake. As Linde-Usiekiewicz (2023) notes, codification can even redirect usage, confirming that reference works exert normative force regardless of intent.

In practice, then, descriptivism and prescriptivism operate not as opposing poles but as intertwined logics. Corpus evidence provides the technical justification, but every entry carries prescriptive weight in shaping what counts as “real” English. The digital era has not resolved the divide; it has accelerated and exposed it, making the politics of legitimacy more visible than ever.

1.4. Research Question

1.4.1. This paper, therefore asks

What linguistic and sociocultural conditions warrant the codification of digital slang into authoritative dictionaries?

Addressing this question requires bridging foundational sociolinguistic theories of language change with contemporary evidence from digital platforms and corpus-based lexicography. The analysis proceeds in three steps. First, it reviews theoretical and empirical research on the diffusion and stabilisation of slang in online environments. Second, it examines how lexicographers operationalise criteria such as frequency, dispersion, and semantic stability when determining

eligibility for dictionary inclusion. Third, it situates Cambridge's 2025 additions within broader debates over descriptivism, prescriptivism, and digital enregisterment, showing how codification today functions simultaneously as descriptive practice and cultural arbitration.

The inquiry matters for at least three reasons. For linguistics, it illuminates how classical models of diffusion and legitimacy must be adapted for algorithmically mediated environments. For lexicography, it clarifies the criteria and risks of updating reference works in an era of accelerated change. For society more broadly, it underscores how dictionary decisions, far from being neutral, participate in cultural negotiations over authority, identity, and generational divides. By analysing the pathways through which slang travels from niche meme to codified entry, this study seeks to clarify not only when slang becomes legitimate but also *why* its institutional recognition matters in the politics of language today.

2. Methodology

2.1. Design and Rationale

This study employs a narrative/theoretical review design rather than a systematic or scoping review. Narrative reviews are especially appropriate for fields where empirical research is fragmented across disciplines, methods, and genres, and where the aim is to draw conceptual connections rather than exhaustively document all available evidence. The focus here is not on providing statistical generalisations, but on synthesising major theoretical frameworks and recent empirical studies that illuminate the question: *under what conditions does internet slang become "legitimate" enough for dictionary codification?*

This approach is informed by criteria for transparent narrative reviewing, including the SANRA guidelines (Baethge, Goldbeck-Wood, and Mertens, 2019). SANRA emphasises clarity of purpose, structured presentation, methodological disclosure, referencing quality, interpretive balance, and contribution to knowledge. While not applied in checklist form, these principles shaped the review's design: clear articulation of the research problem (legitimation of slang), explicit reporting of sources and scope, and critical synthesis of evidence rather than uncritical aggregation.

By adopting this design, the review is positioned as a conceptual bridge: linking classical sociolinguistic theory (Saussure, Labov, Milroy, Hockett) with cutting-edge work on digital slang diffusion, corpus-based tracking, and lexicographic practice. The narrative format allows foregrounding of tensions between descriptivism and prescriptivism, virality and longevity, digital affordances and institutional inertia that cannot be adequately captured in systematic evidence tallies alone.

2.2. Sources and Sampling Strategy

Sources were identified through a combination of purposive sampling and snowball citation tracking. Unlike systematic reviews, which aim for reproducibility through exhaustive search strings and screening protocols, purposive sampling acknowledges the selective and interpretive role of the researcher. The emphasis here was on relevance, recency, and conceptual richness.

The core evidence base included

2.2.1. Peer-reviewed journal articles and conference papers (2022–2025)

Identified primarily through Scopus, Web of Science, MLA International Bibliography, LLBA, and Google Scholar. Search terms combined topical (e.g., *slang*, *meme*, *neologism*) and methodological (e.g., *lexicography*, *dictionary*, *diffusion*, *virality*) keywords.

2.2.2. Lexicographic series and reports

Especially *American Speech's "Among the New Words"* (Hughes et al., 2024a, 2024b) and Cambridge University Press announcements documenting dictionary updates (Cambridge University Press and Assessment, 2025).

2.2.3. Foundational theoretical works

Saussure ([1916]/1983), Labov (1972), Milroy (2002), and Hockett (1960) were retained not as empirical "data points" but as conceptual anchors that continue to shape debates on legitimacy and language change.

2.2.4. *Adjacent empirical studies*

Research on meme virality (Smith, 2022), digital platform affordances (Quick and Maddox, 2024), and prescriptivism politics (Curzan, 2023; VanEyck and Curzan, 2025) were included as they directly inform the lexicographic legitimation process.

Searches were supplemented by backwards and forward citation tracking, enabling identification of influential works not captured in initial database queries. For example, Ilbury, Grieve, and Hall (2024) were identified through references in Hughes et al. (2024) and then became a key piece of evidence.

2.3. Inclusion and Exclusion Principles

Rather than rigid inclusion/exclusion rules typical of systematic reviews, this study used principles of relevance and representativeness:

2.3.1. *Included*

Studies published in English between 2022 and 2025 directly addressing slang diffusion, neologism tracking, dictionary codification, or related sociolinguistic/lexicographic questions.

- Institutional communications (e.g., Cambridge Dictionary updates, AP coverage of word inclusions).
- Foundational theoretical texts used to frame, not evidence claims.
- Excluded:
- Non-academic commentary without methodological grounding (e.g., blogs, personal essays).
- Studies where “slang” was incidental to another focus (e.g., slang in translation pedagogy).
- Duplicative coverage (e.g., multiple near-identical news articles about the same Cambridge update).

The goal was not “coverage of everything” but a curated corpus of influential and representative sources that speak to the review’s guiding question.

2.4. Selection and Review Process

The selection process unfolded in three iterative stages:

- Initial identification: ~120 records identified across databases and targeted searches.
- Screening: Titles and abstracts assessed for direct relevance to slang legitimation and lexicography, reducing the pool to ~45.
- Full-text review: Articles evaluated for methodological transparency, conceptual contribution, and alignment with the review’s aims. This yielded ~22 peer-reviewed studies, supplemented by ~10 institutional/primary lexicographic documents and 5 theoretical texts.

As this was a single-author review, no intercoder reliability was possible. This introduces bias risk, but is mitigated by (a) triangulating empirical findings with theoretical traditions, and (b) reporting inclusion logic transparently.

2.5. Data Handling and Synthesis

Each retained study was summarised in terms of bibliographic details, methodological approach, and key findings. Thematic synthesis was then conducted across three intersecting strands:

- Diffusion dynamics: Corpus/computational studies that quantify slang frequency trajectories, semantic stabilisation, and survival rates (e.g., Keidar et al., 2022; Baumann, 2024).
- Lexicographic practice: Analyses of editorial criteria, corpus surveillance, and the descriptivism–prescriptivism tension in dictionary-making (e.g., Rundell, 2020; VanEyck and Curzan, 2025).
- Institutional codification: Documentation of dictionary updates and their cultural reception (e.g., Cambridge University Press, 2025).

Themes were not pre-imposed, but inductively identified through repeated reading and cross-comparison. For example, the theme of “staying power” emerged consistently across corpus-based and lexicographic studies, while “cultural arbitration” surfaced in both prescriptivist literature and dictionary press releases.

Where possible, findings were triangulated: e.g., viral slang frequency spikes (Keidar et al., 2022) were linked to editorial statements about monitoring “sustained use” (Cambridge, 2025), highlighting the gap between digital virality and lexicographic caution.

Scope and Limitations

This study is explicitly not a systematic review. It is limited by:

- Language and regional bias: Evidence is overwhelmingly Anglophone, with minimal coverage of non-English or cross-linguistic lexicographic processes.
- Platform bias: Much of the empirical work centres on Twitter/X corpora, with TikTok and YouTube still underrepresented in peer-reviewed research.
- Institutional opacity: Dictionary publishers rarely disclose detailed editorial decision processes, so much of the analysis relies on indirect statements and public-facing updates.
- Single-author subjectivity: Inclusion, synthesis, and interpretation were conducted without intercoder checks, meaning that findings reflect the interpretive lens of the researcher.

These limitations shape the scope of claims: this review should be read as an exploratory, interpretive synthesis rather than a definitive empirical account. Its contribution lies in mapping conceptual terrain and highlighting tensions, not in producing exhaustive coverage or predictive models.

2.6. Summary of Methodological Position

In summary, this review situates itself within the narrative review tradition, guided by SANRA’s call for clarity and transparency, but without the procedural apparatus of systematic reviews. It privileges conceptual synthesis over exhaustiveness, curating a corpus of relevant and influential studies to build a theoretically informed account of slang legitimation in the digital age. By being upfront about scope and limitations, it seeks to avoid the charge of “pseudo-systematicity” and instead present itself as what it truly is: a theoretically grounded, purposively curated synthesis that opens space for further empirical research.

3. Literature review

3.1. Theoretical Foundations

The legitimacy of slang has long been a central concern of sociolinguistics, but digital platforms force us to reconsider the frameworks traditionally used to explain how innovations gain acceptance. Classical theory remains indispensable, yet the conditions of the twenty-first century rapid circulation, algorithmic amplification, and memetic culture require us to rethink their application.

Ferdinand de Saussure’s ([1916]/1983) distinction between *langue* (the structured system) and *parole* (individual speech acts) provides a starting point. Slang exemplifies how playful deviations in *parole* can, through repetition, reshape the *langue*. In digital environments, however, the feedback loop between innovation and systemic absorption is compressed. A coinage like *delulu* begins as an ephemeral in-joke but, within weeks, stabilises across fandoms and news media. The temporal lag between creative departure and institutional recognition, once spanning decades, is now measured in months. Saussure’s framework thus remains valid, but the cycle he described operates at an accelerated pace.

William Labov (1972) reoriented the study of change toward social evaluation, emphasizing that linguistic legitimacy emerges when innovations are ratified by a community. Yet in the TikTok or YouTube era, “community” is not a bounded speech group but a platform-mediated audience whose judgments are algorithmically amplified. Likes, shares, and remixes serve as new indices of evaluation, compressing what Labov described as gradual, face-to-face processes into globalized acts of recognition and resistance. The trajectory of *skibidi* from a niche meme to a dictionary headword demonstrates how evaluative practices can now scale almost instantaneously, creating legitimacy not through slow diffusion but through collective, platform-visible uptake.

Milroy’s (2002) social network theory clarifies why such accelerated uptake is possible. Innovations diffuse most rapidly in weak-tie networks, where connections are loose and norms are less tightly enforced. Digital spaces are dominated by precisely these conditions: high turnover, fleeting ties, and porous boundaries. A user need not belong to a fandom or subculture to encounter *tradwife*; algorithmic recommendation ensures its circulation beyond local

solidarity groups. By contrast, strong-tie networks close-knit communities with shared norms often resist linguistic change. The internet’s architecture, therefore, privileges spread over stability, creating a paradox: slang travels further and faster but risks volatility unless stabilised by repeated, cross-register usage. This insight from Milroy remains essential for explaining why many viral terms disappear as quickly as they appear, while a few cross the threshold into institutional recognition.

Where Saussure, Labov, and Milroy help explain system, evaluation, and diffusion, Charles Hockett’s (1960) design features illuminate the creative dimension of slang. His notion of productivity the ability of language users to generate novel forms resonates strongly in meme culture, where remixing and playful innovation are constant. *Delulu* demonstrates productivity by retooling an existing adjective (*delusional*) into a compact, socially loaded marker of fandom identity. Similarly, displacement the ability to refer beyond immediate context underpins terms like *tradwife*, which condense ideological stances into a single lexical item. Admittedly, Hockett’s features were never intended for analysing digital slang, yet their emphasis on universality helps frame why online discourse produces such a dense stream of neologisms. Platform affordances hashtags, audio templates, duet functions extend productivity by making replication frictionless, enabling forms to circulate with minimal cognitive or creative effort.

Taken together, these theories highlight complementary layers of legitimation. Saussure shows how playful departures can become systemically embedded. Labov underscores that this embedding depends on social ratification, a process now mediated by algorithmic visibility. Milroy explains how weak-tie networks allow innovations to leap between groups, producing breadth of circulation. Hockett reminds us that creativity is not incidental but foundational, given new expressive possibilities afforded by digital media. What distinguishes the digital age is not the mechanisms themselves but their compression and amplification: innovation, evaluation, diffusion, and creativity unfold almost simultaneously, collapsing processes that once took decades into viral cycles visible in real time.

This synthesis also clarifies the lexicographic challenge. Dictionaries must interpret whether the rapid movement from *parole* to *langue*, from playful performance to apparent stability, truly signals lasting legitimacy. If viral diffusion mimics community ratification, do weak-tie endorsements suffice for codification? Moreover, if productivity guarantees endless neologisms, which deserve recognition? Classical frameworks illuminate these tensions but also reveal their limits. Saussure, Labov, Milroy, and Hockett provide useful heuristics, yet their pre-digital assumptions about time, space, and community require reworking in light of algorithmically mediated language change. In short, slang legitimacy remains the interplay of system, evaluation, diffusion, and creativity but in the platform era, these processes are accelerated, entangled, and rendered newly visible to lexicographic institutions.

3.2. Lexicographic Authority and Criteria

Dictionaries occupy a paradoxical role in contemporary culture: formally descriptive, yet functionally prescriptive. Publishers routinely emphasise that their mission is to record language “as it is used,” but the public treats dictionary entries as endorsements of what counts as “real” English. This paradox is heightened in the case of internet slang, where institutional codification transforms fleeting memes into symbols of linguistic legitimacy.

Modern lexicographic practice is anchored in corpus surveillance, but the thresholds that trigger inclusion are neither transparent nor uniform. Frequency, dispersion, and semantic stability remain the canonical criteria, yet their operationalisation varies across institutions. Cambridge, for example, explicitly stresses “staying power” and multisource attestation as prerequisites for new entries such as *delulu* or *skibidi* (Cambridge University Press, 2025). Oxford’s pipelines rely more heavily on the Oxford English Corpus and user-submitted citations. Both claim to be corpus-driven, but the differences reveal that editorial philosophy, institutional resources, and even public branding strategies mediate the criteria as much as raw data.

Crucially, these criteria are not value-neutral. Longevity, often presented as a technical safeguard, doubles as a strategy for preserving institutional authority. As Rundell (2020) argues, without durability checks, dictionaries risk becoming repositories of fads, undermining their credibility as reference works. However, what counts as sufficient persistence is subjective: is six months of cross-platform circulation enough, or must a word penetrate journalism, education, and literature? The insistence on “staying power” effectively privileges forms that have migrated into mainstream registers, meaning slang confined to subcultures can remain invisible regardless of frequency.

This raises a distributional question: who benefits when persistence is defined in this way? Viral terms that leap from youth or niche communities into mainstream media *delulu* moving from K-pop fandom into lifestyle journalism, for instance are fast-tracked toward legitimacy. By contrast, slang that circulates robustly in marginalised or racialised communities often stalls until appropriated by broader audiences. African American Vernacular English, for example,

has supplied global English with some of its most enduring slang, yet lexicographic codification often lags until after mainstream uptake. The “staying power” criterion thus reflects not only linguistic durability but also sociocultural hierarchies about whose language deserves permanence.

Corpus design itself compounds this bias. Most lexicographic corpora remain skewed toward journalism, published literature, and edited online sources. Platforms central to slang innovation TikTok captions, Discord chat, Twitch transcripts are rarely integrated. As a result, what counts as “attested usage” is already filtered through institutional decisions about what registers count as evidence. The claim that dictionaries are “data-driven” therefore obscures the fact that data sources are curated and exclusionary. The authority of the dictionary is exercised not only through editorial decisions but also through the invisible infrastructures of corpus construction.

Furthermore, editorial judgments are not made in a cultural vacuum. The Cambridge announcements surrounding its August 2025 update framed the addition of *skibidi*, *delulu*, and *tradwife* as evidence of institutional responsiveness to “language as it evolves.” Such press releases highlight how dictionary codification is increasingly public-facing, tied to branding and cultural relevance. This signals a shift: entries are not merely scholarly decisions but reputational interventions, positioning the publisher as a responsive authority in fast-moving debates over culture, gender, and digital media.

The tension between descriptivism and stewardship remains at the heart of lexicographic authority. On one hand, editors are expected to follow evidence wherever it leads. On the other, they must protect the stability of the dictionary as a cultural institution. That stability depends on being selective excluding countless viral items despite evidence of short-term popularity. Yet selectivity itself is an ideological act, shaping public perceptions of which communities and practices constitute “proper” English.

In this sense, corpus-driven lexicography does not eliminate subjectivity but reconfigures it. The lexicographer’s role is less about passively recording “what people say” than about adjudicating when usage has matured enough to deserve permanence. Those decisions inevitably privilege some voices over others, reinforcing existing cultural hierarchies while presenting them as neutral data science. In the digital age, therefore, dictionary authority should be read not as a mirror of usage but as a filter that legitimises particular kinds of circulation, persistence, and social recognition.

3.3. Slang in the Digital Age

Digital platforms have radically shortened the lifecycle of lexical change, enabling slang to diffuse across demographics and geographies in weeks rather than years. This acceleration is no longer speculative: it is empirically documented across multiple corpora and platforms.

Ilbury, Grieve, and Hall (2024), drawing on 1.8 billion geotagged tweets, demonstrate how youth-marked features of Multicultural London English diffused nationally through online exposure, showing a clear online–offline feedback loop. Nevertheless, Twitter/X alone cannot capture the multimodal intensity of today’s slang. Baumann (2024), using cross-platform survival models, show that while most neologisms collapse within a few weeks, a select minority survive by embedding themselves in multiple registers. This finding aligns directly with lexicographic criteria of “staying power.” Together, these studies illustrate both the volatility and selectivity of digital diffusion.

Platform affordances make this volatility structurally predictable. TikTok’s duet, stitch, and green screen features create lineage chains where phrases, gestures, or sound clips are replicated and reinterpreted thousands of times. Each replication is simultaneously a citation and a stabilisation: a template that keeps semantic cores intact while permitting playful variation (Quick and Maddox, 2024). YouTube’s comment culture and reaction formats function similarly, embedding slang into multimodal contexts where textual, auditory, and visual repetition anchor meaning. Diffusion, in this sense, is not just about exposure to large audiences but about infrastructures designed to engineer replication.

Computational work confirms this dual dynamic of acceleration and stabilisation. Keidar et al. (2022) show that slang exhibits sharper volatility in frequency trajectories than non-slang, with dramatic spikes tied to trend cycles. However, semantic drift is relatively slow: even ephemeral slang like *skibidi* retains a recognisable core across thousands of citations. This paradox volatile usage but stable meaning helps explain why some terms vanish after their viral moment while others persist long enough to attract lexicographic attention.

The pathways of *skibidi*, *delulu*, and *tradwife* exemplify this hybrid process. Each originated in restricted niches meme culture, fandom discourse, or ideological subcultures and then spread rapidly through affordance-driven cascades. Their eventual codification in Cambridge’s 2025 update underscores a crucial point: editors are not simply counting

raw tokens but are tracking whether a form survives replication, crosses registers, and stabilises semantically. In this sense, digital affordances create the preconditions for lexicographic codification by accelerating both exposure and stabilisation.

Additional evidence from American Speech's "Among the New Words" series supports this claim. Hughes et al. (2024) document how social media has become the dominant incubator for lexical innovation, with editors tracking new forms as they migrate into general-purpose registers such as journalism and entertainment. This corroborates Cambridge's editorial emphasis on cross-register attestation as a marker of permanence.

In short, the digital age reshapes slang trajectories by combining algorithmic acceleration (virality through platform design) with affordance-enabled stabilisation (templates, chains, and memetic anchors). Virality alone may fade, but when replication infrastructures secure semantic coherence across contexts, some terms achieve the durability required for codification. Lexicographers are thus compelled to re-evaluate their timelines: waiting a decade to confirm a slang item's persistence is no longer feasible when diffusion can peak globally within days. The digital ecosystem has not only compressed cycles of slang emergence but also forced institutions to confront a new tempo of legitimacy.

3.4. Empirical Evidence (2022–2025)

Recent scholarship confirms that while digital platforms accelerate the diffusion of slang, only a small fraction of innovations achieve the retention needed for codification. This asymmetry emerges consistently across multiple large-scale corpus and lexicographic studies.

Corpus-based analyses illustrate how slang spreads rapidly through weak-tie networks rather than traditional face-to-face communities. For example, Ilbury, Grieve, and Hall (2024), drawing on 1.8 billion geotagged tweets, show that online exposure predicts offline diffusion of Multicultural London English features evidence that platform visibility can drive national-scale uptake. Complementary computational work by Keidar et al. (2022) identifies the volatility of slang frequency profiles, characterized by dramatic spikes and collapses, even as semantic cores remain relatively stable. Together, these studies highlight both the speed and instability of digital-era lexical circulation.

Other research emphasises that rapid spread rarely guarantees survival. Baumann (2024) demonstrate in a diachronic study of lexical innovation that the overwhelming majority of new forms disappear quickly, with only a small fraction crossing the threshold into long-term adoption. Their findings reframe virality as a filter rather than a guarantee: most forms flare briefly before vanishing, while only a select few demonstrate the persistence required for codification.

Lexicographic studies mirror this selectivity. Cambridge's August 2025 update, which added *skibidi*, *delulu*, and *tradwife*, illustrates how editors prioritise not just frequency but evidence of semantic stability and survival across registers (Cambridge University Press and Assessment, 2025). Longstanding monitoring projects, such as American Speech's "Among the New Words" series (Hughes et al., 2024), further demonstrate how editors track emergent terms as they migrate into journalism and everyday discourse, consolidating their legitimacy.

Taken together, this body of research converges on a paradox: social media accelerates diffusion but also accelerates attrition. Dictionaries, positioned as arbiters of durability, intervene selectively, recording only those items that survive the bottleneck of short-term volatility. Synthesised findings, particularly those of Baumann (2024), sharpen the point that legitimacy is not about speed or reach alone but about long-term survival across multiple registers of language use.

4. Discussion

4.1. Pathways to Legitimacy

The findings suggest that slang legitimacy operates less like a linear ladder and more like a selective filter. New forms typically emerge in niche spaces, gain visibility through algorithmic amplification, diffuse into journalism or broader media, stabilise semantically, broaden across registers, and only then enter institutional codification. At each stage, however, most candidates drop out. This filtration model aligns with Baumann (2024) evidence that only a small minority of innovations survive beyond their first burst. The implication is that dictionary codification reflects not just frequency thresholds but the cumulative winnowing of forms through multiple social and temporal bottlenecks.

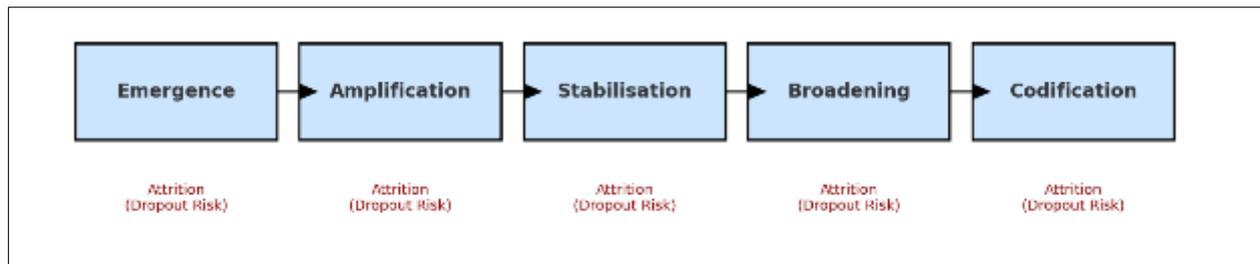


Figure 1 Pathways of Slang Legitimation

4.2. Virality versus Staying Power

Virality remains the entry point to public consciousness, but is insufficient for legitimacy. Keidar et al. (2022) demonstrate that slang items display sharp frequency volatility, spiking more violently than other lexical categories. However, the majority fade before reaching sustained uptake. Baumann (2024) reinforce this point: attrition is the norm, persistence the exception. Dictionary inclusion, therefore, depends on endurance, not visibility. Cambridge’s 2025 decision to add *skibidi* illustrates this selectivity: the word outlasted other meme-born terms by maintaining cross-register presence long after its peak. By contrast, *cheugy* a term heavily viral in 2021 failed to stabilise semantically or broaden across demographics, and was thus excluded. The persistence of *yeet* offers a counterpoint, showing that some items born of virality can consolidate into enduring lexical resources. In short, virality makes a word *noticeable*; staying power makes it *legitimate*.

4.3. Cross-Demographic Uptake

Diffusion across demographics and registers appears crucial in determining whether a term survives the filter. Ilbury, Grieve, and Hall (2024) demonstrate how youth-marked features of Multicultural London English spread nationally via weak-tie networks; however, not all innovations achieve such widespread adoption. Lexicographic codification tends to privilege those items that cross beyond their original subcultural base. *Delulu*, for instance, traveled from K-pop fandom into lifestyle journalism, while *tradwife* migrated from ideological subcultures into mainstream debates on gender roles. These cases illustrate a pattern supported by Hughes et al. (2024): codification is most likely when usage transcends age, geography, and ideology. The decisive factor, then, is not just frequency but whether the item becomes intelligible and relevant across a broad social spectrum.

4.4. Dictionaries as Cultural Arbiters

Dictionaries inevitably act as arbiters of legitimacy, but in the digital era, this role extends beyond language description to cultural positioning. Cambridge’s 2025 update illustrates the point: although editors justified *skibidi*, *delulu*, and *tradwife* through corpus evidence, media coverage reframed the move as a cultural endorsement of youth-driven digital slang. In this way, codification functions as a reputational intervention signalling that a publisher is responsive to contemporary culture while still safeguarding its authority.

Rather than rehashing the descriptivist–prescriptivist debate outlined in 3.2, the key insight here is that dictionaries serve as symbolic actors in broader cultural struggles. Each entry operates simultaneously as a record of usage and as a public statement about what counts as legitimate English, linking editorial decisions to broader debates over identity, authority, and generational change.

4.5. Risks of Rapid Codification

The acceleration of slang diffusion introduces institutional risks. First, ephemerality: premature inclusion could flood dictionaries with transient items, undermining their role as stable reference works. Second, reputational vulnerability: playful entries invite criticism that standards are falling, a concern already visible in public reactions to *delulu* and *skibidi*. Third, institutional mechanics intensify the stakes. Editorial boards must balance scholarly accuracy with commercial imperatives, since brand identity and user trust directly affect market competitiveness (Lew, 2024). Dictionaries are not only scholarly resources but also publishing businesses: the decision to add or exclude a word, therefore, reflects a calculus involving credibility, audience expectations, and long-term brand stability. Recognising these institutional dimensions highlights why codification cannot be reduced to corpus data alone.

4.6. Implications

The broader implications extend across education, sociolinguistics, and lexicography. For pedagogy, new entries provide opportunities to teach register awareness and metalinguistic reflection, illustrating how slang travels from ephemeral contexts to mainstream codification. For sociolinguistics, the findings reinforce theories of change that foreground weak-tie diffusion and selective survival, offering evidence that digital infrastructures compress but do not fundamentally alter the filtering mechanisms of language change. For lexicography, the key implication is the need for transparent criteria. As Lew (2024) argues, user trust depends on visible evidence of how decisions are made. Transparency is therefore not merely a policy recommendation but an operational requirement: without it, dictionary inclusion risks being read as ideological rather than evidence-based. Future work should explore how dashboards, editorial notes, and open corpora can enhance the audibility of the legitimation process, aligning institutional authority with contemporary expectations of accountability.

Table 1 Cross-Linguistic Approaches to Slang Codification

Language / Dictionary	Codification Approach
English (Cambridge Dictionary)	Corpus-driven and relatively responsive; integrates digital slang (e.g., <i>skibidi</i> , <i>delulu</i>) once sustained frequency and dispersion are demonstrated.
French (Académie Française)	Highly conservative; resists inclusion of slang and digital neologisms, prioritising traditional, standardised forms.
German (Duden)	Moderately cautious; slang is considered only after broader acceptance in media, journalism, and education.
Japanese (Kōjien)	Gradual and cautious inclusion of youth slang; typically marked with usage labels (e.g., “informal,” “colloquial”).

To situate Cambridge’s approach within a broader context, it is instructive to compare how other major dictionaries handle slang. As shown in Table 1, English-language dictionaries like Cambridge are relatively agile in responding to digital neologisms, while French authorities such as the Académie Française remain highly conservative, emphasising linguistic purity. German practice, exemplified by Duden, reflects a middle ground, delaying codification until slang penetrates mainstream registers. Japanese dictionaries, such as *Kōjien*, similarly take a cautious approach but often mark entries with explicit usage labels (e.g., “informal”), thereby acknowledging social salience while signalling register boundaries. This cross-linguistic comparison underscores that slang codification is not only a matter of frequency and stability but also reflects divergent cultural philosophies of linguistic authority.

4.7. Ethical Reflection

Beyond questions of ephemerality and reputation, the codification of politically and ideologically charged slang introduces distinct ethical challenges. Terms such as “*tradwife*,” “*woke*,” or “*snowflake*” carry cultural and ideological baggage; their inclusion in authoritative dictionaries may be perceived as institutional endorsement rather than neutral documentation. For instance, *tradwife* indexes not just lifestyle choice but also an ideological position on gender roles; *woke* has been recontextualised from a marker of awareness to a pejorative in certain discourses; and *snowflake* functions as an insult targeting perceived fragility. In each case, the act of codification risks amplifying contested meanings by granting them symbolic legitimacy.

Lexicographers have sought to mitigate this tension by deploying usage notes and labels “informal,” “derogatory,” “disapproving” which signal an evaluative stance without erasing the term’s existence. These paratextual strategies acknowledge the social reality of contested slang while distancing the institution from perceived endorsement. However, such labels themselves can provoke debate: who decides whether a term is derogatory, and in which contexts? The ethical task, therefore, lies not only in documenting usage but also in navigating the cultural politics of legitimacy. Codification becomes a balancing act between scholarly descriptivism, public perception, and institutional responsibility, with editors serving as both cultural mediators and linguistic recorders.

4.8. Practical Framework for Evaluating Slang Legitimacy

To operationalise these insights, we propose a framework that lexicographers and educators can apply when assessing whether slang warrants codification:

- Frequency Threshold– Evidence of sustained use across at least 12 months, avoiding short-lived spikes.
- Dispersion Across Demographics– Documented presence across age groups, regions, and registers (e.g., journalism, social media, education).
- Semantic Stability– Core meaning remains recognisable despite playful or ironic variation.
- Register Breadth– Usage extends beyond niche subcultures and is intelligible in mainstream discourse.
- Cultural Sensitivity– Terms flagged for derogatory, disapproving, or ideological content should include usage notes to contextualise potential harms.
- Transparency of Criteria– Editorial rationale and evidence should be made visible, fostering user trust in the dictionary’s authority.

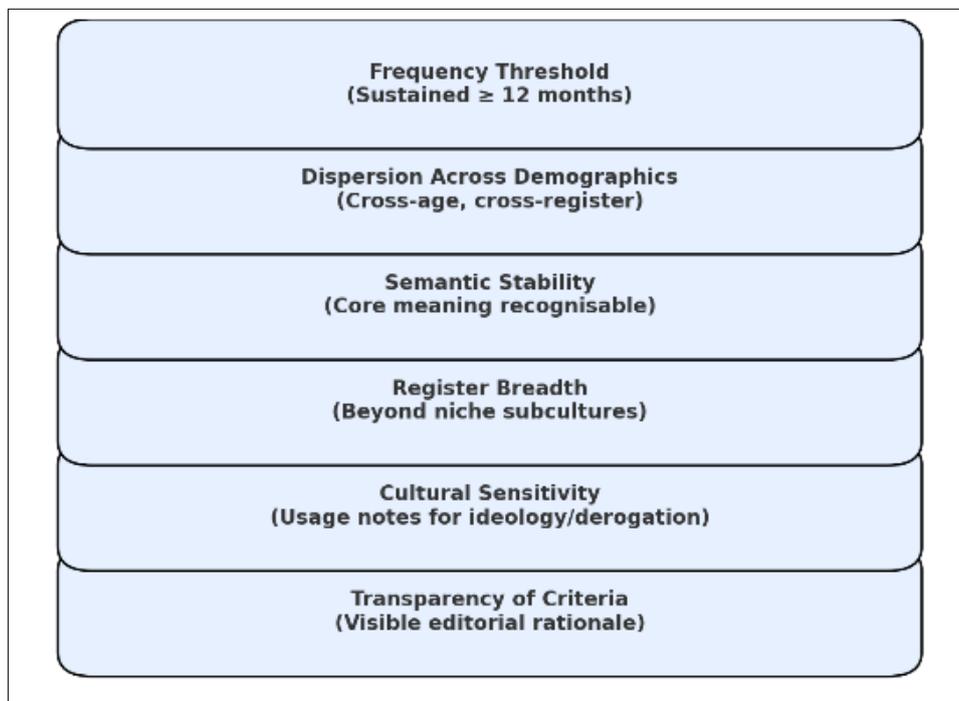


Figure 2 Practical Framework for Evaluating Slang Legitimacy

This checklist highlights that slang codification is not a purely technical matter, but a socio-cultural negotiation that requires clarity, caution, and accountability.

5. Conclusion

This review has demonstrated that digital slang does not achieve dictionary codification solely through popularity. Instead, it survives a demanding sequence of tests involving endurance, breadth of diffusion, semantic stability, and interpretability. While social media platforms accelerate the visibility of new expressions at an unprecedented speed, most fail to gain institutional recognition before collapsing. Only a minority such as *skibidi*, *delulu*, and *tradwife* persist long enough and spread widely enough to satisfy the criteria applied by lexicographers. Codification thus remains both descriptive and normative: dictionaries respond to usage, yet their decisions simultaneously confer cultural authority by defining what counts as “legitimate” English.

Theoretically, this study reframes slang legitimation as a filtering process rather than a linear progression. Each stage emergence, amplification, stabilisation, broadening, codification functions as a bottleneck that eliminates the majority of innovations (Figure 1). Empirical research underscores this rarity: Baumann (2024) demonstrate that only a small fraction of neologisms survives beyond their viral moment, and lexicographic practice institutionalises only those survivors that meet thresholds of frequency, dispersion, and stability.

The digital era compresses these dynamics dramatically. What once unfolded over decades can now occur in weeks, placing new pressures on lexicographers whose decisions are increasingly public and culturally scrutinised. As Table 1 and Figure 3 highlight, cross-linguistic traditions vary in their responsiveness, and codification involves not only

technical judgments but also ethical considerations and editorial frameworks. Inclusion is never neutral: dictionaries are cultural arbiters whose choices shape debates around authority, identity, and generational change.

Future research must move beyond retrospective description toward predictive and comparative approaches. Survival modelling across multimodal corpora (e.g., TikTok audio, YouTube captions, Twitter/X text) could help forecast which innovations are likely to persist. Cross-linguistic comparisons will reveal whether the filtering model observed in English generalises across diverse lexicographic traditions. Experimental lexicography should assess whether dictionary inclusion itself stabilises usage by institutionalising legitimacy. Ultimately, an ethnographic study of editorial boards could shed light on how commercial strategy, brand identity, and reputational risk intersect with corpus evidence in shaping codification decisions.

Methodologically, progress requires combining computational breadth with qualitative depth: frequency trajectories and diffusion models must be supplemented by close analysis of meaning, community evaluation, and editorial interpretation. Central to all of this is transparency. As Lew (2024) emphasises, user trust depends on visible evidence of editorial criteria; without openness, codification risks being perceived as ideological rather than empirical.

In short, the legitimacy of slang in the digital age is not about sudden elevation, but about survival through volatility, dispersion, and institutional judgment. Dictionaries remain the hinge between ephemeral creativity and enduring linguistic record yet in today's accelerated climate, that hinge does not swing smoothly. It creaks under the cultural, commercial, and ethical pressures of deciding what endures.

Compliance with ethical standards

Disclosure of conflict of interest

The author declares no conflict of interest.

Statement of Ethical Approval

This article does not contain any studies with human participants or animals performed by the author.

Funding

No external funding was received for the preparation of this manuscript.

Data Availability Statement

No datasets were generated or analysed during the current study.

References

- [1] Associated Press. (2025, August 19). Cambridge Dictionary adds skibidi and tradwife among 6,000 new words. AP News. <https://www.ap.org/news-highlights/spotlights/2025/cambridge-dictionary-adds-skibidi-and-tradwife-among-6000-new-words/>
- [2] Baumann A (2024) Lexical innovations are rarely passed on during one's lifetime: Epidemiological perspectives on estimating the basic reproductive ratio of words. PLoS ONE 19(12): e0312336. <https://doi.org/10.1371/journal.pone.0312336>
- [3] Burada, M., and Sinu, R. (2023). Prescriptivism in lexicography: A bird's eye view. Bucharest Working Papers in Linguistics, 25(1), 7–19. https://bwpl.unibuc.ro/wp-content/uploads/2024/02/2-BWPL_2023_1_Burada-si-Sinu.pdf
- [4] Cambridge University Press and Assessment. (2025, August 18). Cambridge Dictionary adds skibidi, delulu and tradwife. Cambridge University Press and Assessment. <https://www.cambridge.org/news-and-insights/cambridge-dictionary-adds-skibidi-delulu-and-tradwife>
- [5] Hockett, C. F. (1960). The origin of speech. Scientific American, 203(3), 88–96. <https://doi.org/10.1038/scientificamerican0960-88>
- [6] Hughes, B., Calhoun, K., Fawcett, A., Wright, K. E., Zimmer, B., Brewster, E., McLean, J., and Zhang, L. (2024). Among the new words. American Speech, 99(1), 78–90. <https://doi.org/10.1215/00031283-11186920>

- [7] Ilbury, C., Grieve, J., and Hall, D. (2024). Using social media to infer the diffusion of an urban contact dialect: The case of Multicultural London English lexis. *Journal of Sociolinguistics*, 28(3), 45–70. <https://doi.org/10.1111/josl.12653>
- [8] Keidar, D., Opedal, A., Jin, Z., and Sachan, M. (2022). Slangvolution: A causal analysis of semantic change and frequency dynamics in slang. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics* (pp. 1422–1442). Association for Computational Linguistics. <https://aclanthology.org/2022.acl-long.101/>
- [9] Lew, R. (2024). What lexical factors drive look-ups in the English Wiktionary? *SAGE Open*, 14(1), 1–14. <https://doi.org/10.1177/21582440231219101>
- [10] Linde-Usiekniewicz, J. (2023). In quest of influences of Polish language dictionaries on the usage of native speakers. *International Journal of Lexicography*, 36(4), 447–465. <https://doi.org/10.1093/ijl/ecad015>
- [11] Milroy, J. (2002). Social networks. In J. K. Chambers, P. Trudgill, and N. Schilling-Estes (Eds.), *The handbook of language variation and change* (pp. 549–572). Blackwell.
- [12] Quick, M., and Maddox, J. (2024). Us, them, right, wrong: How TikTok’s Green Screen, Duet, and Stitch help shape political discourse. *First Monday*, 29(3). <https://doi.org/10.5210/fm.v29i3.13360>
- [13] Rundell, M. (2020). Monitoring new words in a digital age: Challenges for lexicographers. *Dictionaries*, 41(2), 1–23.
- [14] Saussure, F. de. ([1916]/1983). *Course in general linguistics* (C. Bally and A. Sechehaye, Eds.; R. Harris, Trans.). Open Court.
- [15] Smith, N. (2022). Memetic moments: The speed of Twitter memes. *Journal of Digital Social Research*, 4(3), 19–35. <https://doi.org/10.33621/jdsr.v4i1.95>
- [16] Ștefănescu, M., and Minică, M. (2023). On “The Fabrick of the Tongue”: Language metaphors used to advocate descriptivism/prescriptivism in English and Romanian dictionaries. *Studia Universitatis Babeș-Bolyai Philologia*, 68(1), 181–194. <https://doi.org/10.24193/subbphilo.2023.1.10>
- [17] VanEyck, K., and Curzan, A. (2025). The politics of prescriptivism: One style manual, one century. *American Speech*, 100(1), 22–41. <https://doi.org/10.1215/00031283-10867229>