

UDC 811.111'42-028.63:168.35

DOI <https://doi.org/10.24919/2308-4863/103-39>**Daryna MUDRYK,**orcid.org/0000-0002-4962-1620

PhD in Education,

Senior Lecturer at the Department of Foreign Languages
National University of Water and Environmental Engineering
(Rivne, Ukraine) darynka.tararuk@gmail.com

LINGUISTIC CONSTRUCTION OF UNCERTAINTY IN CONTEMPORARY ENGLISH DIGITAL DISCOURSE: A MULTIDIMENSIONAL ANALYTICAL FRAMEWORK

The article addresses the problem of fragmentation in research on the linguistic construction of uncertainty in contemporary English digital discourse, which has traditionally been examined within separate domains, including epistemic modality, hedging, evidentiality, and stance. The aim of the study is to develop an integrative multidimensional framework for analysing uncertainty that combines a taxonomy of formal resources, a seven-category functional scheme, and diagnostic heuristics for distinguishing genuinely epistemic from routinised uses of linguistic markers. These include the substitution, co-occurrence, and sequential-position heuristics. Methodologically, the study is corpus-informed and conceptual-methodological in nature and is based on a synthesis of previously published corpus studies of English digital discourse. The synthesis draws on eight such studies conducted on the Reddit and Twitter/X platforms. The proposed Form → Context → Function → Interaction model conceptualises uncertainty marking as a multifunctional and context-dependent phenomenon whose function cannot be reliably determined from a formal marker alone. Seven functional categories are substantiated: epistemic, evidential, approximative, mitigating, interpersonal-hedging, stance-managing, and responsibility-managing. The synthesis of previous research indicates that variation in uncertainty marking is associated not only with platform but also with topic, community norms, and text length. At the same time, the available evidence does not support an unambiguous interpretation of platform effects, while the multimodal component of the model remains theoretically grounded and requires further empirical validation. The proposed framework provides a basis for more accurate and reproducible coding of uncertainty functions and can be applied in future corpus-based research on digital discourse.

Key words: uncertainty, epistemic modality, epistemic stance, hedging, digital discourse, corpus linguistics, pragmatics.

Дарина МУДРИК,orcid.org/0000-0002-4962-1620

доктор філософії,

старший викладач кафедри іноземних мов

Національного університету водного господарства та природокористування
(Рівне, Україна) darynka.tararuk@gmail.com

МОВНЕ КОНСТРУЮВАННЯ НЕВИЗНАЧЕНОСТІ В СУЧАСНОМУ АНГЛОМОВНОМУ ЦИФРОВОМУ ДИСКУРСІ: БАГАТОВИМІРНА АНАЛІТИЧНА МОДЕЛЬ

Стаття присвячена проблемі фрагментарності досліджень мовного конструювання невизначеності в сучасному англomовному цифровому дискурсі, що традиційно розглядається в межах окремих напрямів – епістемічної модальності, хеджування, евіденційності та стансу. Метою дослідження є розроблення інтегративної багатовимірної моделі аналізу невизначеності, яка поєднує таксономію формальних ресурсів, семикатегоріальну функціональну схему та діагностичні евристичні для розрізнення власне епістемічних і рутинізованих уживань мовних маркерів. Зокрема застосовано евристичні субституції, сумісної появи та послідовної позиції маркера в репліці. Методологічно дослідження має корпусно-інформований концептуальний і методологічний характер та ґрунтується на синтезі раніше опублікованих корпусних досліджень англomовного цифрового дискурсу. Синтезовано вісім таких досліджень, виконаних на матеріалі платформ Reddit і Twitter/X. Запропонована модель Форма → Контекст → Функція → Взаємодія трактує маркування невизначеності як поліфункціональне та контекстно залежне явище, функція якого не може бути надійно встановлена лише за формальним маркером. Обґрунтовано сім функціональних категорій: епістемічну, евіденційну, апроксимативну, пом'якшувальну, міжособистісно-хеджувальну, станс-регульовальну та відповідальнісно-регульовальну. Синтез попередніх досліджень показує, що варіювання мовного маркування невизначеності пов'язане не лише з платформою, а й із тематикою, нормами спільноти та довжиною

тексту. Водночас наявні дані не дають підстав для однозначного трактування платформного ефекту, а мультимодальний компонент моделі залишається теоретично обґрунтованим і потребує подальшої емпіричної перевірки. Запропонований підхід забезпечує основу для більш точного й відтворюваного кодування функцій невизначеності та може бути використаний у майбутніх корпусних дослідженнях цифрового дискурсу.

Ключові слова: невизначеність, епістемічна модальність, епістемічна позиція (станс), хеджування, цифровий дискурс, корпусна лінгвістика, прагматика.

Problem Statement. Write I think it's on the third floor into three different digital contexts and the phrase does three different jobs: hedging genuine uncertainty when addressed to someone lost in a building; softening a confidently held claim in an argument thread, pre-empting objection rather than reporting doubt; or, in a third setting, functioning as a bare turn-launching formula with almost no epistemic content. A linguistic analysis that stops at the lexical item – counting occurrences of I think and calling the count a measure of uncertainty – collapses these three uses into one, losing precisely the distinction that makes the phenomenon worth studying. This motivates the article's central claim: linguistic marking of uncertainty is not a transparent readout of a speaker's private epistemic state. Confident speakers hedge for interpersonal, rhetorical, and responsibility-management reasons; genuinely uncertain speakers sometimes assert without qualification, particularly where hedging carries its own social cost. Treating uncertainty marking as a discursive act rather than a cognitive report reframes the analytic task: the question is not which forms exist, but what work a given form is doing in a given sequential and interpersonal context.

Digital discourse is a productive site for pursuing this question because its characteristic conditions – public and often unknown audiences, asynchronous and persistent text, rapid production, frequent argumentative or evaluative framing – plausibly reweight the interpersonal stakes of hedging relative to private spoken conversation, the setting in which much of the classical literature on stance and hedging was developed (Biber, Finegan, 1988; Biber, Finegan, 1989; Hyland, 1998). Existing corpus research has begun to answer this question piecemeal, across studies of Reddit and Twitter/X that differ in topic, method, and platform (Biri, 2022; Clarke, 2022; Dayter, Messerli, 2022; Liimatta, 2019; Liimatta, 2022a; Liimatta, 2022b; Messerli et al., 2025; Visonà, Kurt, 2025), but has not yet produced a shared analytical apparatus letting such studies be compared and cumulated rather than read in isolation. The gap is therefore one of integration rather than neglect, with three components: a conceptual gap, since modality, hedging, evidentiality, and stance are typically studied as separate phenomena though a single form routinely does work these traditions would classify differently; a method-

ological gap, since surface forms cannot be equated with function without an explicit disambiguation procedure – treating every instance of I think as an epistemic hedge, as much corpus-based hedging research still implicitly does, conflates genuinely distinct uses; and a platform-specificity gap, since digital discourse is often treated, even in careful corpus work, as a single undifferentiated register, when the cited literature on Reddit alone shows substantial internal variation by community, topic, and text length (Liimatta, 2019; Liimatta, 2022a; Liimatta, 2022b; Messerli et al., 2025). This article's contribution is to close these three gaps with a shared framework rather than to close any one of them individually.

Analysis of Previous Research. Uncertainty marking is conventionally studied within separate traditions – epistemic modality (Palmer, 2001; Du Bois, 2007), hedging (Lakoff, 1973; Hyland, 1998; Hyland, 2005; Fraser, 2010), evidentiality (Aikhenvald, 2004; Clarke, 2022), and stance (Du Bois, 2007; Kiesling, 2022) – that rarely engage with one another, alongside a complementary digital-discourse methodology (Bolander, Locher, 2014; Herring, 2001; Herring, Androutsopoulos, 2015; Landert et al., 2023). This article's starting distinction is between uncertainty as a private cognitive state and as a public, linguistically constructed act: a speaker who writes pretty sure it's the third one, don't quote me may hold near-certainty, the hedge managing the social consequences of a possible error rather than reporting doubt. Treating stance as socially consequential and interactionally negotiated rather than transparent (Kiesling, 2022) warrants this distinction: frequency counts of “hedges” cannot, alone, support claims about how uncertain speakers on a platform “really are”, since such counts measure discursive practice shaped by audience design and interpersonal risk.

English distributes uncertainty marking across every grammatical level – modal auxiliaries (Palmer, 2001; Du Bois, 2007), epistemic verbs, adverbs, adjectives, approximators (Channell, 1994), and evidential expressions (Clarke, 2022; Aikhenvald, 2004) – summarised in full in Table 1. Epistemic verbs of cognition often detach from their epistemic origin in routinised first-person use; approximators mark imprecision about a quantity rather than doubt about the proposition as a whole (Channell, 1994),

while evidential expressions mark information source rather than a direct likelihood judgment (Clarke, 2022; Aikhenvald, 2004), a distinction surface-form counts routinely discard.

Three further traditions converge on uncertainty marking. Hedging is achievable through heterogeneous formal means rather than a fixed word class (Fraser, 2010; Hyland, 1996), underwriting the decision to treat uncertainty marking as a functional category rather than a closed lexical class. Politeness theory explains the same forms as face-protective rather than epistemic (Brown, Levinson, 1987); since the two cannot reliably be distinguished by lexical form alone, holding them apart is a task the coding scheme supports. Stance research adds that a stance act simultaneously positions the speaker and aligns with an interlocutor's prior stance (Du Bois, 2007; Kiesling, 2022) – together warranting this article's key claim that a single marker often serves several functions at once, which is why scheme allows multiple simultaneous codings.

Aim of the Article. The aim of the article is to propose an integrative, multidimensional framework – a taxonomy of formal resources, an operational functional coding scheme, and a set of diagnostic heuristics – for analysing uncertainty in contemporary English digital discourse, and to synthesise what previously published corpus research on Reddit and Twitter/X does and does not establish about how platform, topic, and text length condition its distribution.

The aim defines the scope of the analysis and the synthesis of previously published corpus research on Reddit and Twitter/X.

Main Body of the Article. This article is a corpus-informed conceptual and methodological study: its content is drawn entirely from previously published, independently verified corpus research, synthesised against a purpose-built framework. Findings attributed to specific prior publications are kept distinct throughout from the article's own theoretical argument and from constructed illustrative examples labelled [constructed example]. The comparative claims below draw on eight previously published corpus studies of Reddit and Twitter/X detailed in Table 5, with every number reproduced from the source's own methodology, not estimated.

The seven categories and their formal indicators, diagnostic principles, and boundary risks are set out in full in Table 2; two of its boundary-risk pairings recur often enough to warrant a brief gloss here. Epistemic uncertainty and mitigation can be realised by identical surface forms; per Table 2, they are distinguished only by co-text – subsequent claims made with more confidence, the absence of correction, the overall assertiveness of the surrounding discourse – which indicates high underlying confidence despite a surface hedge (Brown, Levinson, 1987). Stance management and epistemic uncertainty are likewise conflated wherever a marker such as I think appears turn-initially, launching a contribution to an exchange, rather

Table 1

Taxonomy of Linguistic Resources for Constructing Uncertainty (by Form)

Level	Category	Examples	Primary source(s)
Lexical	Epistemic adverbs	perhaps, probably, possibly, apparently, arguably	(Biber, Finegan, 1989)
Lexical	Epistemic verbs	seem, appear, suppose, assume, believe, think, guess	(Hyland, 1998)
Lexical	Epistemic adjectives	possible, probable, likely, unlikely, uncertain, unclear	(Palmer, 2001)
Lexical	Approximators	about, around, roughly, kind of, sort of	(Channell, 1994)
Lexical	Evidential expressions	apparently, according to, as far as I know, I heard	(Clarke, 2022; Aikhenvald, 2004)
Grammatical	Epistemic modal auxiliaries	may, might, could, must	(Palmer, 2001; Du Bois, 2007)
Grammatical	Extraposed / conditional constructions	it's unclear whether..., if I'm not mistaken	(Fraser, 2010)
Grammatical	Interrogative framing	posing a claim as a question rather than an assertion	(Fraser, 2010)
Pragmatic / discourse	Routinised epistemic phrases	I think, I guess, I suppose, I'm not sure, I'd say	(Fraser, 2010; Kiesling, 2022)
Pragmatic / discourse	Discourse-organisational hedges	I mean, not sure, if I'm not mistaken	(Kiesling, 2022)
Multimodal	Typographic / orthographic cues	ellipsis, capitalisation, repetition, emoji	(Herring, Androutsopoulos, 2015)

than attaching to a specific, checkable proposition (Du Bois, 2007; Kiesling, 2022) – precisely the ambiguity resolves. Elsewhere, dual coding is appropriate whenever a token genuinely serves two functions at once, most often evidential uncertainty together with responsibility management (apparently, allegedly), since forcing a single label onto such tokens would misrepresent the text.

The paradigmatic ambiguous case, and its occurrence is not treated here as sufficient grounds, on its own, for coding a token as epistemic uncertainty. Three diagnostic heuristics – structured prompts for a coder's judgment rather than infallible tests – resolve such cases in sequential context. The substitution heuristic asks whether the token could be replaced by a bare assertion or by I know without changing the speaker's apparent commitment. For example, in [constructed example] I think you're forgetting the

update shipped last Tuesday, substitution with I know changes little, so the token is more plausibly mitigation or interpersonal hedging than epistemic uncertainty. The co-occurrence heuristic asks whether the token appears alongside other, independently epistemic material – as in I think it might have shipped last Tuesday, but I'm honestly not sure – that would be redundant unless the marker itself carried genuine epistemic weight. The sequential-position heuristic asks whether the token appears turn-initially, launching a contribution (more consistent with discourse organisation or stance management), or attaches to a specific, checkable proposition under live dispute (more consistent with genuine epistemic hedging). These heuristics require future empirical validation through systematic annotation and inter-rater reliability testing; no such validation was conducted in the present study.

Table 2

Coding Scheme: Seven Functional Categories

Category	Definition	Linguistic indicators	Diagnostic principle	Main boundary risk
Epistemic uncertainty	Marks assessed likelihood a proposition is true	Modal auxiliaries, epistemic adverbs on a checkable proposition	Substitution with a certainty marker changes apparent commitment	Mitigation (same forms, high underlying confidence)
Evidential uncertainty	Marks information source / reliability	apparently, according to, I heard	Presence of source-marking lexis	Responsibility management (dual coding often appropriate)
Approximation	Marks imprecision about quantity/ boundary, not truth	about, around, kind of, sort of	Attaches to a scalar/ quantified element	Minimal; often co-occurs with epistemic uncertainty
Mitigation	Softens face-threatening force for interpersonal reasons	Same forms as epistemic uncertainty	Co-text shows high underlying confidence	Epistemic uncertainty
Interpersonal hedging	Manages speaker–addressee relationship broadly	correct me if I'm wrong, what do you think	Direct address to interlocutor	Mitigation (orientation to proposition vs interlocutor)
Stance management	Signals investment level, not epistemic content per se	I think, I guess (routinised)	Sequential-position heuristic (turn-initial)	Epistemic uncertainty
Responsibility management	Distributes/disclaims accountability for accuracy	apparently, allegedly, don't quote me	Explicit disclaimer or evidential frame	Evidential uncertainty

Table 3

Constructed Illustrations of the Seven Functional Categories (cf. Table 2)

Category (Table 2)	Constructed illustration
Epistemic uncertainty	No idea if that's still true, honestly.
Evidential uncertainty	Apparently they changed the policy, but I haven't read it myself.
Approximation	It's kind of a grey area, might depend on the country.
Mitigation	I could be wrong, but I don't think that's how it works.
Interpersonal hedging	I think you might be missing a step there – correct me if I'm wrong.
Stance management	I guess it's a caching issue, but that's just my read on it.
Responsibility management	Apparently it's already fixed upstream, though I haven't verified it.

This subsection synthesises, rather than reports, what the eight cited corpus studies establish when read collectively. The clearest evidence on epistemic-marker density and platform comes from a 1.2-million-word corpus of climate-change discourse (Biri, 2022): Reddit posts carried more epistemic stance markers than Twitter, with differences associated with both platform and speakers' ideological position – sceptics used less hedging of likelihood and more evaluative language attacking opponents' reliability than proponents did. Read through the stance-theoretic account, this can be interpreted as strategic positioning within an adversarial exchange rather than as a difference in how certain the two groups genuinely were about the science. This single, topic-bounded finding offers suggestive rather than strong evidence and should not be generalised across all Reddit or Twitter content; read against evidence that Reddit is not internally homogeneous (Liimatta, 2019; Messerli et al., 2025) and that epistemic-stance marking is sensitive to topic contentiousness over time (Visonà, Kurt, 2025).

On register variation, a 100-million-message corpus from r/ChangeMyView found no systematic difference in formality markers between persuasive and non-persuasive posts (Dayter, Messerli, 2022). This null result cautions against treating formal linguistic features as straightforward indicators of persuasive success. Feature frequency has been shown to change systematically with comment length (Liimatta, 2022a; Liimatta, 2022b), independently of the subreddit-level register variation documented separately (Liimatta, 2019), jointly implying that text length and community purpose are independent sources of variation a study must model separately (the warrant for stratified sampling). Finally, no study synthesised above targets the multimodal reinforcement of uncertainty – punc-

tuation, capitalisation, repetition, emoji – and this article analysed no multimodal data of its own; the multimodal component of the model rests on a theoretical account of digital multimodality alone (Herring, Androutsopoulos, 2015), flagged as the component most in need of dedicated empirical testing.

The reviewed literature provides stronger empirical grounding for the Form and Function dimensions, while Context emerges as a necessary methodological component. By contrast, the Interaction dimension, particularly its multimodal component, remains insufficiently tested and requires further empirical validation.

The model proposed here is the article's central contribution. It is not a psychological model of how speakers assemble uncertain utterances, a causal model of what makes speakers uncertain, or a frequency model of how often a marker occurs – it is an analytical and annotation framework relating what a marker looks like to what it is doing, given where and to whom it is said. Its four components follow the sequence Form → Context → Function → Interaction: Form is the four-level taxonomy in Table 1; Context comprises the situational parameters to condition how a form is used and interpreted; Function is the seven-category scheme in Table 2, assigned only once context is taken into account; so that the combination, not any individual marker, is the proper unit of analysis.

The interaction among components appears in a single constructed sentence: [constructed example] I guess it might be around twenty minutes or so... Here I guess operates, once context is taken into account, as a pragmatic-level stance-management device (Kiesling, 2022) rather than straightforward epistemic reporting, since, by the substitution heuristic, little would change if it were dropped; might oper-

Figure 1

The Form → Context → Function → Interaction model of uncertainty construction

Component	What it captures	Elements
FORM	What a marker looks like	Lexical (adverbs, verbs, adjectives, approximators, evidentials); Grammatical (modal auxiliaries, extraposed/conditional constructions, interrogative framing); Pragmatic/discourse (routinised epistemic phrases, discourse-organisational hedges); Multimodal (punctuation, capitalisation, repetition, emoji)
CONTEXT	The situational parameters a marker's use must be read against	Platform and community norms; topic and topic contentiousness; text length; sequential position within a thread; interlocutor relationship
FUNCTION	What a marker is doing, assigned only once context is taken into account	Epistemic uncertainty; evidential uncertainty; approximation; mitigation; interpersonal hedging; stance management; responsibility management
INTERACTION	The proper unit of analysis	Co-occurrence of markers from multiple form-levels within a single utterance; the combination, not any individual marker, carries the functional weight

Table 4

Evidence Base for Each Level of the Model

Level	What it contributes	Theoretical motivation	Cited empirical motivation
Lexical	Open-class vocabulary marking likelihood, approximation, evidential source	(Lakoff, 1973; Coates, 1983; Chafe, 1986)	(Dayter, Messerli, 2022) (lexical markers in r/ChangeMyView)
Grammatical	Closed-class and syntactic marking of modality, conditionality, interrogative force	(Palmer, 2001; Du Bois, 2007; Fraser, 2010)	(Dayter, Messerli, 2022) (suasive verbs, WH-questions, emphatics)
Pragmatic	Interpretive layer assigning function(s); disambiguates routinised forms	(Du Bois, 2007; Kiesling, 2022; Fraser, 2010)	(Brown, Levinson, 1987) (stance × ideology interaction)
Multimodal	Typographic/orthographic reinforcement or modification of uncertainty	(Herring, Androutsopoulos, 2015)	None cited – theoretical only; principal target for future testing

Table 5

Empirical Scope of Relevant Prior Corpus Studies

Study	Platform(s)	Corpus / scale (as reported by authors)	Principal finding	Relevance to present framework
Biri (2022)	Twitter, Reddit	Climate-change discourse; 1.2 million words	Higher epistemic-marker density on Reddit than Twitter within this corpus; interaction with speaker ideology	A qualified platform association
Clarke (2022)	Twitter	English tweets (scale as reported in source)	Tweets vary along recognisable functional dimensions rather than forming one register	Supports treating Twitter/X as internally differentiated
Dayter & Messerli (2022)	Reddit (r/ChangeMyView)	Persuasive/non-persuasive subcorpora; 100 million messages	No systematic formality difference between persuasive and non-persuasive posts	Null result constraining claims about uncertainty and persuasive success
Liimatta (2019)	Reddit	Multi-subreddit sample (scale as reported in source)	Substantial register variation within Reddit, driven by subreddit purpose	Motivates community-level stratified sampling
Liimatta (2022a)	Reddit	Large Reddit comment sample (scale as reported in source)	Feature frequency varies systematically with comment length	Motivates text-length stratification as an independent factor
Liimatta (2022b)	Reddit	Reddit register/function comparison (as reported in source)	Registers differ in the communicative functions text length serves	Reinforces text length as non-reducible to platform
Messerli et al. (2025)	Reddit	Four linked subreddit studies; multi-study synthesis	Reddit is a heterogeneous set of communicative spaces, not a single register	Reinforces internal platform heterogeneity
Visonà & Kurt (2025)	Twitter/X	Tweets on a contested public-institution topic, 2013–2023	Epistemic-stance marking sensitive to topic contentiousness over time	Supports treating topic contentiousness as an additional contextual variable.

ates grammatically as an epistemic modal auxiliary (Palmer, 2001); and around/or so operate lexically as approximators marking imprecision about a quantity rather than doubt about the trip's occurrence (Channell, 1994). A trailing typographic ellipsis, where present, would operate multimodally, iconically enacting the trailing-off of the claim – a plausible but, untested contribution of that level. No single element captures what the utterance as a whole accomplishes; the model's claim is that the combination is the proper unit, extending a heterogeneous-means account of hedging (Fraser, 2010) to the layered conditions of digital

writing. Two design requirements follow for annotation: co-occurring markers should be recorded as a set rather than tallied independently, since combination carries information isolated counts discard; and functional coding should always follow, never precede, contextual annotation, since assigning function from form alone would defeat the model's purpose.

Conclusions. The study has developed an integrative framework for analysing the linguistic construction of uncertainty in contemporary English digital discourse. Its main contribution is a multidimensional analytical apparatus that combines a taxonomy of

formal resources, a seven-category functional coding scheme, diagnostic heuristics for distinguishing genuinely epistemic uses from routinised and interactionally motivated ones, and the Form → Context → Function → Interaction model. The framework conceptualises uncertainty not as a direct linguistic reflection of a speaker's private epistemic state, but as a context-dependent discursive phenomenon whose interpretation depends on the interaction of formal, situational, functional, and interpersonal factors.

The synthesis of previously published corpus research indicates that surface-form frequency alone is insufficient for establishing the degree of a speaker's epistemic uncertainty. Forms such as *I think* may perform epistemic, mitigating, interpersonal, stance-managing, or responsibility-managing functions depending on their co-text and sequential position. The proposed diagnostic heuristics therefore provide a basis for separating formal identification from functional interpretation and for allowing multiple functional coding where a single marker performs more than one discourse function.

The reviewed evidence also suggests that variation in uncertainty marking across digital platforms should not be attributed to platform identity alone. Topic contentiousness, community norms, and text length emerge as important contextual factors that may interact with platform-related variation. The proposed framework consequently treats Context as a set of interacting parameters rather than as a sin-

gle platform variable. At the same time, the available evidence does not justify broad causal claims about platform effects, and the proposed functional categories have not yet been empirically validated through a unified corpus and systematic annotation.

The multimodal component represents a further theoretically motivated extension of the framework. Although typographic and orthographic resources such as punctuation, capitalisation, repetition, and emoji may contribute to the construction or interpretation of uncertainty in digital discourse, their functional contribution remains insufficiently tested. This dimension therefore constitutes an important direction for future empirical research rather than an established finding of the present study.

The practical value of the framework lies in its potential to support more precise and reproducible corpus-based research. Future empirical studies should apply the proposed coding scheme to newly compiled and appropriately stratified digital corpora, test the diagnostic heuristics through double-coding and inter-rater reliability procedures, examine functional co-occurrence across platforms and communities, and assess the contribution of multimodal resources to the interpretation of uncertainty. In this respect, the present study provides a conceptual and methodological foundation for moving from surface-form counts toward a more context-sensitive analysis of how uncertainty is constructed and negotiated in digital English.

BIBLIOGRAPHY

1. Aikhenvald, A. Y. *Evidentiality*. Oxford : Oxford University Press, 2004. 452 p.
2. Biber, D., Finegan, E. Adverbial stance types in English. *Discourse Processes*. 1988. Vol. 11, No. 1. P. 1–34. DOI: 10.1080/01638538809544689.
3. Biber, D., Finegan, E. Styles of stance in English: Lexical and grammatical marking of evidentiality and affect. *Text*. 1989. Vol. 9, No. 1. P. 93–124.
4. Biri, Y. Epistemic stance in the climate change debate: A comparison of proponents and sceptics on Twitter and Reddit. *Register Studies*. 2022. Vol. 4, No. 2. P. 232–262. DOI: 10.1075/rs.22005.bir.
5. Bolander, B., Locher, M. A. Doing sociolinguistic research on computer-mediated data: A review of four methodological issues. *Discourse, Context & Media*. 2014. Vol. 3. P. 14–26. DOI: 10.1016/j.dcm.2013.10.004.
6. Brown, P., Levinson, S. C. *Politeness: Some universals in language use*. Cambridge : Cambridge University Press, 1987. 345 p.
7. Chafe, W. Evidentiality in English conversation and academic writing. *Evidentiality: The linguistic coding of epistemology* / ed. by W. Chafe, J. Nichols. Norwood : Ablex, 1986. P. 261–272.
8. Channell, J. *Vague language*. Oxford : Oxford University Press, 1994. 220 p.
9. Clarke, I. A multi-dimensional analysis of English tweets. *Language and Literature*. 2022. Vol. 31, No. 2. P. 124–149. DOI: 10.1177/09639470221090369.
10. Coates, J. *The semantics of the modal auxiliaries*. London : Croom Helm, 1983. 259 p.
11. Dayter, D., Messerli, T. C. Persuasive language and features of formality on the r/ChangeMyView subreddit. *Internet Pragmatics*. 2022. Vol. 5, No. 1. P. 165–195. DOI: 10.1075/ip.00072.day.
12. Du Bois, J. W. The stance triangle. *Stancetaking in discourse* / ed. by R. Englebretson. Amsterdam : John Benjamins, 2007. P. 139–182. DOI: 10.1075/pbns.164.07du.
13. Fraser, B. Pragmatic competence: The case of hedging. *New approaches to hedging* / ed. by G. Kaltenböck, W. Mihatsch, S. Schneider. Bingley : Emerald Group Publishing, 2010. P. 15–34.
14. Herring, S. C. Computer-mediated discourse. *The handbook of discourse analysis* / ed. by D. Tannen, D. Schiffrin, H. Hamilton. Oxford : Blackwell, 2001. P. 612–634.

15. Herring, S. C., Androutsopoulos, J. Computer-mediated discourse 2.0. The handbook of discourse analysis. 2nd ed. / ed. by D. Tannen, H. E. Hamilton, D. Schiffrin. Oxford : Wiley-Blackwell, 2015. P. 127–151. DOI: 10.1002/9781118584194.ch6.
16. Hyland, K. Writing without conviction? Hedging in science research articles. *Applied Linguistics*. 1996. Vol. 17, No. 4. P. 433–454. DOI: 10.1093/applin/17.4.433.
17. Hyland, K. Boosting, hedging, and the negotiation of academic knowledge. *Text*. 1998. Vol. 18, No. 3. P. 349–382.
18. Hyland, K. *Metadiscourse: Exploring interaction in writing*. London : Continuum, 2005. 230 p.
19. Kiesling, S. F. Stance and stancetaking. *Annual Review of Linguistics*. 2022. Vol. 8. P. 409–426. DOI: 10.1146/annurev-linguistics-031120-121256.
20. Lakoff, G. Hedges: A study in meaning criteria and the logic of fuzzy concepts. *Journal of Philosophical Logic*. 1973. Vol. 2, No. 4. P. 458–508.
21. Landert, D., Dayter, D., Messerli, T. C., Locher, M. A. *Corpus pragmatics*. Cambridge : Cambridge University Press, 2023. 300 p. DOI: 10.1017/9781009091107.
22. Liimatta, A. Exploring register variation on Reddit. *Register Studies*. 2019. Vol. 1, No. 2. P. 269–295. DOI: 10.1075/rs.18005.lii.
23. Liimatta, A. Register variation across text lengths: Evidence from social media. *International Journal of Corpus Linguistics*. 2022. Vol. 28, No. 2. P. 202–231. DOI: 10.1075/ijcl.20177.lii.
24. Liimatta, A. Do registers have different functions for text length? A case study of Reddit. *Register Studies*. 2022. Vol. 4, No. 2. P. 263–287. DOI: 10.1075/rs.22007.lii.
25. Messerli, T. C., Dayter, D., Leuckert, S., Liimatta, A., Mahler, H., Bohmann, A., Kozma, G., Tosin, R. Digital debating cultures: Communicative practices on Reddit. *Digital Scholarship in the Humanities*. 2025. Vol. 40, No. 1. P. 227–240. DOI: 10.1093/llc/fqaf005.
26. Palmer, F. R. *Mood and modality*. 2nd ed. Cambridge : Cambridge University Press, 2001. 236 p.
27. Visonà, M. W., Kurt, Ş. A corpus-assisted discourse analysis of epistemic stances in tweets about U.S. police from 2013 to 2023. *Applied Corpus Linguistics*. 2025. Vol. 5, No. 3. Article 100138. DOI: 10.1016/j.acorp.2025.100138.

REFERENCES

1. Aikhenvald A. Y. (2004) *Evidentiality*. Oxford: Oxford University Press. 452 p.
2. Biber D., Finegan E. (1988) Adverbial stance types in English. *Discourse Processes*. 11. 1–34. DOI: 10.1080/01638538809544689.
3. Biber D., Finegan E. (1989) Styles of stance in English: Lexical and grammatical marking of evidentiality and affect. *Text*. 9. 93–124.
4. Biri Y. (2022) Epistemic stance in the climate change debate: A comparison of proponents and sceptics on Twitter and Reddit. *Register Studies*. 4. 232–262. DOI: 10.1075/rs.22005.bir.
5. Bolander B., Locher M. A. (2014) Doing sociolinguistic research on computer-mediated data: A review of four methodological issues. *Discourse, Context & Media*. 3. 14–26. DOI: 10.1016/j.dcm.2013.10.004.
6. Brown P., Levinson S. C. (1987) *Politeness: Some universals in language use*. Cambridge: Cambridge University Press. 345 p.
7. Chafe W. (1986) Evidentiality in English conversation and academic writing. In: Chafe W., Nichols J. (eds.) *Evidentiality: The linguistic coding of epistemology*. Norwood: Ablex. P. 261–272.
8. Channell J. (1994) *Vague language*. Oxford: Oxford University Press. 220 p.
9. Clarke I. (2022) A multi-dimensional analysis of English tweets. *Language and Literature*. 31. 124–149. DOI: 10.1177/09639470221090369.
10. Coates J. (1983) *The semantics of the modal auxiliaries*. London: Croom Helm. 259 p.
11. Dayter D., Messerli T. C. (2022) Persuasive language and features of formality on the r/ChangeMyView subreddit. *Internet Pragmatics*. 5. 165–195. DOI: 10.1075/ip.00072.day.
12. Du Bois J. W. (2007) The stance triangle. In: Englebretson R. (ed.) *Stancetaking in discourse*. Amsterdam: John Benjamins. P. 139–182. DOI: 10.1075/pbns.164.07du.
13. Fraser B. (2010) Pragmatic competence: The case of hedging. In: Kaltenböck G., Mihatsch W., Schneider S. (eds.) *New approaches to hedging*. Bingley: Emerald Group Publishing. P. 15–34.
14. Herring S. C. (2001) Computer-mediated discourse. In: Tannen D., Schiffrin D., Hamilton H. (eds.) *The handbook of discourse analysis*. Oxford: Blackwell. P. 612–634.
15. Herring S. C., Androutsopoulos J. (2015) Computer-mediated discourse 2.0. In: Tannen D., Hamilton H. E., Schiffrin D. (eds.) *The handbook of discourse analysis*. 2nd ed. Oxford: Wiley-Blackwell. P. 127–151. DOI: 10.1002/9781118584194.ch6.
16. Hyland K. (1996) Writing without conviction? Hedging in science research articles. *Applied Linguistics*. 17. 433–454. DOI: 10.1093/applin/17.4.433.
17. Hyland K. (1998) Boosting, hedging, and the negotiation of academic knowledge. *Text*. 18. 349–382.
18. Hyland K. (2005) *Metadiscourse: Exploring interaction in writing*. London: Continuum. 230 p.
19. Kiesling S. F. (2022) Stance and stancetaking. *Annual Review of Linguistics*. 8. 409–426. DOI: 10.1146/annurev-linguistics-031120-121256. [in English].
20. Lakoff G. (1973) Hedges: A study in meaning criteria and the logic of fuzzy concepts. *Journal of Philosophical Logic*. 2. 458–508.
21. Landert D., Dayter D., Messerli T. C., Locher M. A. (2023) *Corpus pragmatics*. Cambridge: Cambridge University Press. 300 p. DOI: 10.1017/9781009091107.

22. Liimatta A. (2019) Exploring register variation on Reddit. *Register Studies*. 1. 269–295. DOI: 10.1075/rs.18005.lii.
23. Liimatta A. (2022a) Register variation across text lengths: Evidence from social media. *International Journal of Corpus Linguistics*. 28. 202–231. DOI: 10.1075/ijcl.20177.lii.
24. Liimatta A. (2022b) Do registers have different functions for text length? A case study of Reddit. *Register Studies*. 4. 263–287. DOI: 10.1075/rs.22007.lii.
25. Messerli T. C., Dayter D., Leuckert S., Liimatta A., Mahler H., Bohmann A., Kozma G., Tosin R. (2025) Digital debating cultures: Communicative practices on Reddit. *Digital Scholarship in the Humanities*. 40. 227–240. DOI: 10.1093/dlc/fqaf005.
26. Palmer F. R. (2001) *Mood and modality*. 2nd ed. Cambridge: Cambridge University Press. 236 p.
27. Visonà M. W., Kurt Ş. (2025) A corpus-assisted discourse analysis of epistemic stances in tweets about U.S. police from 2013 to 2023. *Applied Corpus Linguistics*. 5. Article 100138. DOI: 10.1016/j.acorp.2025.100138.

Дата першого надходження статті до видання: 28.08.2026

Дата прийняття статті до друку після рецензування: 25.09.2026

Дата публікації (оприлюднення) статті: 16.10.2026

Стаття поширюється на умовах
ліцензії відкритого доступу (CC BY 4.0)

