

**Borys Grinchenko Kyiv Metropolitan University**  
**Faculty of Romance and Germanic Philology**  
**Linguistics and Translation Department**

*Цими підписами засвідчую, що подані  
на захист рукопис та електронний  
документ є ідентичні.*  
19.05.2026 *[Signature]*

**INNOVATIVE AND TRANSFORMATIONAL TRENDS IN THE  
LANGUAGE OF SOCIAL NETWORKS: THE CASE OF TWITTER/X**

MA Paper  
Viktoriia Basarab  
Group MLAm-1-24-2.0d

Research supervisor  
Professor R. K. Makhachashvili

Approved for defense at  
the Linguistics and  
Translation Department  
26, May, 2026, Minutes 5

\_\_\_\_\_ Andrii KOZACHUK

Kyiv 2026

Abstract.....	3
Introduction.....	3
Theoretical background.....	5
Research Methods.....	10
2.1.1. Rationale for the choice of platform and time frame.....	10
2.1.2. Justification of the time frame.....	11
2.1.3. Data collection procedure.....	12
2.1.4. Composition of the dataset.....	12
2.3. Voyant Tools: Frequency and Thematic Analysis of Corpora.....	13
2.4. ChatGPT as an analytical assistant: principles and limitations of use....	15
2.5. Critical discourse analysis: procedure and principles.....	16
2.6. Practical implementation: from data to analysis.....	17
2.7. Ethical aspects of working with social media content.....	19
Results.....	19
3.1. Overall picture: what is specific to the language of Twitter/X.....	20
3.2. Neologisms: mechanisms of emergence and consolidation.....	21
3.2.1. Doomscrolling: Case Studies of Lexical Innovation.....	21
3.2.2. Enshittification: Neologism as Social Criticism.....	24
3.2.3. Rizz and goblincore: neologisms from the subcultural environment..	27
3.3. Semantic shifts: when an old word acquires a new meaning.....	30
3.3.1. Theoretical reminder and operational criteria.....	30
3.3.2. Ratio: From Mathematics to Social Judgment.....	30
3.3.3. Slay: Semantic amelioration through Subcultural Diffusion.....	33
3.3.4. Era: Semantic Shift Through Construction.....	35
3.4. Pragmatic formulas: conventional units of pragmatic identification.....	36
3.4.1. Theoretical characterization of memetic expressions.....	36
3.4.2. Understood the assignment and no notes.....	37
3.4.3. No cap, IYKYK: grammaticalized abbreviations as meme units...	39
3.5. Syntactic and punctuation levels: compression and pragmatic drift.....	41
3.6. Chronological dynamics: the relationship of linguistic innovations with the sociocultural context.....	43
3.7. Summary results: systematization and confirmation of the hypothesis..	45
References.....	50

## Abstract

This master's paper is devoted to the study of innovative and transformative linguistic processes on the social media platform Twitter/X, based on English-language posts from 2019 to 2025. The aim of the work is to identify, systematise and linguistically interpret the key trends characterising the platform's language during the period under study, as well as to identify the extralinguistic factors that determine these trends.

The subject of this study is the language of English-language Twitter/X posts. The analysis focuses on innovative lexical units (neologisms, internet slang, media abbreviations), transformations in pragmatic communication strategies, as well as specific syntactic and punctuation patterns that have emerged or undergone significant changes under the influence of the digital environment. The data consists of authentic tweets collected via the official Twitter/X platform across various thematic categories — ranging from everyday communication to news and humour.

The research question is formulated as follows: what linguistic changes have occurred on English-language Twitter/X between 2019 and 2025, and what linguistic and extralinguistic factors have driven them? The working hypothesis is that Twitter/X is an independent linguistic environment with its own mechanisms of language creation and the spread of innovations, and that the patterns identified are systematic rather than random.

The research methodology combines corpus-linguistic methods (frequency analysis, keyword analysis, collocation analysis) using Voyant Tools, the use of ChatGPT as an analytical assistant for classifying examples, and critical discourse analysis of selected micro-contexts. The analysis revealed that three interrelated transformational trends are clearly discernible in the language of Twitter/X: the active creation of new words and the reinterpretation of existing ones (semantic shifts); and the reduction of syntactic structures and the pragmatic reorientation of punctuation marks.

**Keywords:** *digital discourse, Twitter/X, social media language, corpus linguistics, neologisms, AI analysis.*

## Introduction

Relevance of the topic. Social media today is not merely a channel of communication – it has become an environment where language thrives and evolves at an unprecedented pace. New words emerging on Twitter/X are constantly being added, with hundreds of new words and expressions appearing; some remain platform-specific slang, whilst others rapidly enter general linguistic usage. *Doomscrolling* made it into the Oxford English Dictionary less than a year after appearing on Twitter. *Cancel*, *ratio*, *understood the assignment*, *no cap* – these terms didn't just appear in dictionaries; they changed the way millions of people describe their experiences.

Alongside lexical changes, syntax is also undergoing transformation: sentences are becoming shorter and grammatically less conventional, whilst punctuation marks are taking on pragmatic functions not covered by any textbook.

These processes are occurring at an extraordinary pace – faster than any linguistic change in the pre-literate and even the print eras. At the same time, they are systematic: researchers are observing not a chaotic 'deterioration' of the

language, but the emergence of new patterns, new norms and new genres. This is precisely why linguistic analysis of the language of Twitter/X is not only academically interesting but also practically important: it allows us to understand the direction in which language is moving in the context of digital communication.

The relevance of this topic stems from the fact that social media today is not merely a channel of communication, but a fully-fledged linguistic environment where processes take place on a daily basis that traditional linguistics has spent years studying. The language of Twitter/X is changing rapidly and noticeably: new words that appear here find their way into dictionaries; expressions that begin as jokes within niche communities become part of the everyday speech of millions. However, the study of these processes has been hampered by the lack of convenient and reliable tools for working with a large corpora of informal texts. Advances in corpus linguistics and artificial intelligence methods have opened up new possibilities for such research.

The object of the study is the language of English-language posts on the Twitter/X platform from 2019 to 2025.

The subject of the study is innovative and transformative linguistic trends on social media: lexical innovations (neologisms and semantic shifts), syntactic compression, and pragmatic shifts in punctuation.

The aim of the study is to identify, systematise and linguistically interpret the main innovative and transformational trends in the language of English-language Twitter/X from 2019 to 2025, as well as to determine the linguistic and extra-linguistic factors that underpin them.

To achieve this objective, the following tasks have been identified:

1. to conduct a critical review of the academic literature on the linguistics of digital discourse and social media language;
2. to examine theoretical approaches to the study of linguistic innovation, neologism and semantic shift in the digital environment;
3. to define research methods and justify the choice of material;
4. to identify characteristic lexical and thematic trends using Voyant Tools;
5. to conduct a qualitative analysis of lexical innovations and semantic shifts;
6. to analyse syntactic transformations and the pragmatic drift of punctuation marks;
7. to verify the working hypothesis and formulate final conclusions.

**Research data.** The data consists of authentic English-language tweets collected from the official Twitter/X platform between 2019 and 2025. The corpus includes original tweets spanning various thematic categories: from everyday conversation and humour to news, mental health and pop culture. The corpus covers both language segments and is balanced across years, allowing changes over time to be tracked. The sample includes only original posts from

verified and unverified real individuals, excluding retweets and accounts belonging to bots and organisations.

**Working hypothesis:** social media (Twitter/X) constitute a distinct linguistic environment with its own mechanisms for language creation and the dissemination of innovations; the linguistic transformations observed within it are systemic and are determined both by the platform's technological features and by the sociocultural context of the period under study.

**Research methods:** frequency and thematic analysis using Voyant Tools (Cirrus, Trends, Contexts); analytical support from ChatGPT; critical discourse analysis of micro-contexts based on Norman Fairclough.

**Theoretical significance of the study.** This paper provides a theoretical analysis, systematisation and justification of the linguistic foundations for the study of digital discourse as an independent linguistic phenomenon; it refines the conceptual framework for describing linguistic innovation in the context of platform-based communication; a theoretical justification is proposed for the integration of corpus methods and artificial intelligence technologies into the linguistic study of online text.

**Practical significance of the results.** The materials and conclusions of this study can be used in courses on media linguistics, pragmatics and sociolinguistics; in the development of tools for the automatic analysis of English-language online texts; and as a methodological basis for further research into the language of social media.

**Scientific relevance of the results achieved.** This study: (a) conducts a systematic diachronic analysis of the language used on English-language Twitter/X over a six-year period, utilising AI tools; (b) employs GPT-4 as a tool for the qualitative linguistic annotation of a large corpus of Twitter texts.

## Theoretical background

1.1. The emergence of digital discourse linguistics: key stages and approaches.

The study of the language of online communication as a distinct academic field began relatively recently – in the late 1990s, when the widespread adoption of email and chat rooms prompted researchers to ask what, in fact, was happening to language in this new environment. Initial responses were polarised: some scholars saw internet language as a threat to linguistic standards and evidence of a decline in literacy, whilst others viewed it as a fundamentally new linguistic phenomenon worthy of serious analysis. Over time, the latter view prevailed, and today the linguistics of digital discourse is a recognised and rapidly developing branch of linguistics.

The early stages of the discipline's development are primarily associated with the name of David Crystal. In his book *Language and the Internet* (Crystal, 2001), he introduced the concept of *Netspeak* – a distinct form of linguistic activity that combines features of both written and spoken language, yet is not

identical to either. Crystal emphasised that this was not a ‘corrupted’ version of standard language, but rather a linguistic adaptation to new conditions of communication: limited time for writing, the absence of non-verbal cues, and a specific relationship between author and audience (Crystal, 2001, p. 48). In the second edition of the book (Crystal, 2011), the researcher confirmed and expanded on these conclusions, focusing separately on various genres of online communication – from email to chat rooms and blogs.

At the same time, Susan Herring developed a methodology for analysing computer-mediated communication – Computer-Mediated Discourse Analysis (CMDA). Herring convincingly argues that the technological tool itself – the platform’s architecture, its rules and limitations – influences language just as significantly as the social context (Herring, 2007). The researcher emphasises the importance of taking into account the *situational factors* of each platform: who is writing, for whom, for what purpose, and under what technical conditions (Herring, 1996, p. 154). This observation proved particularly valuable for later research into social networks, where architectural decisions (such as Twitter’s character limit) directly shape linguistic practices.

The next significant stage was the emergence of ‘social media linguistics’ in the early 2010s. With the rapid spread of Twitter, Facebook and Instagram, researchers began to distinguish not simply between ‘internet language’, but between the language of specific platforms. It turned out that the same user writes differently depending on the intended audience and the norms accepted on each platform. Jannis Androustopoulos (2014), in his study of digital language practices and online communication, demonstrates that different social platforms shape specific linguistic norms and styles, determined by their technical capabilities and audience characteristics.

Since then, the linguistics of digital discourse has developed along two main lines. The first – sociolinguistic – analyses the language of social media through the prism of identity, communities of practice and power relations. The second – computational – develops and applies quantitative methods for analysing large corpora of digital texts. This work deliberately combines both approaches: quantitative methods provide breadth of coverage, whilst qualitative discourse analysis offers depth of interpretation.

## **1.2. Social media as a linguistic environment.**

The language of social media has several key characteristics that distinguish it from previous forms of written communication. Firstly, its public-private nature: a person writes as if for acquaintances, but in fact for an unlimited audience. This gives rise to ironic self-distancing and meme-like phrases as structural responses to this communicative paradox. Secondly, algorithmic mediation: the algorithms of recommendation systems reward apt and resonant linguistic phrases with wider reach. Taina Bucher (2017) calls this *algorithmic imaginaries* – the unconscious adaptation of language to the

platform's requirements. Thirdly, textual compression: competition for the audience's attention encourages conciseness and semantic brevity.

### **1.2. Twitter/X as a subject of linguistic research.**

Among all social media platforms, Twitter occupies a special place in linguistic research. Firstly, it is an exclusively text-based platform: even when a tweet contains an image, text remains its central element. Secondly, the 280-character limit (140 characters until 2017) has fostered a unique culture of linguistic compression: here, users learn to express the maximum with minimal means. This has directly influenced the development of specific lexical, syntactic and pragmatic strategies, which constitute the main focus of this study.

Jacob Eisenstein et al. (Eisenstein et al., 2011) were the first to conduct a large-scale corpus analysis of Twitter language: they demonstrated that non-standard spelling (abbreviations, phonetic spelling, lack of punctuation) is not a chaotic phenomenon, but rather a sociolinguistically determined one with clear patterns. Later, J.Grieve, A.Nini and D.Guo in 2017, using a large Twitter corpus, found that the platform is a leading hub for lexical innovation: words first recorded on Twitter eventually enter general linguistic usage significantly more often than neologisms from other sources. M. Zappavigna (2012) analyses the pragmatics of hashtags and demonstrates that they serve not only as search tags but also as community markers – a way of signalling affiliation with a particular group of views or interests.

Another key concept for this work is the idea of *ambient affiliation*, proposed by Zappavigna (2011,2012): Twitter allows people to connect with an undefined, broad audience, which in turn shapes the specific pragmatics of posts – designed to appeal to both close friends and strangers. This paradox of public-private discourse gives rise to particular communicative strategies, which are reflected in all the identified trends – from ironic constructions to meme-based formulas.

### **1.3. The Theory of Linguistic Innovation: from Labov to memes.**

To accurately describe linguistic changes on Twitter/X, there is a need for a theoretical framework to explain how and why linguistic innovations arise and spread. First and foremost, the terminological foundations of this study must be established: in this study, we understand a linguistic innovation to be any new or significantly altered linguistic form – a new word, a new meaning of an old word, a new grammatical construction, or a new pragmatic function of an existing linguistic unit – that arises and becomes established within a particular linguistic community (Labov, 2001, p. 8).

William Labov, in his analysis of the mechanisms of linguistic change in spoken language, formulated several principles that have proved extremely useful for the analysis of digital discourse as well. In particular, he demonstrated that linguistic innovations spread via social networks: from *linguistic leaders* – people with high social authority within the community – to the rest of the participants (Labov, 1994). On Twitter, this mechanism is reproduced in an

accelerated and transparent form: when a well-known account uses a new word or construction, it instantly reaches thousands of followers and can become established in the platform's language much faster than happens in real life.

Everett Rogers' theory of the diffusion of innovations (Rogers, 1962) offers another useful framework. Rogers describes the spread of any innovation – including linguistic ones – following an S-shaped curve: first it is adopted by a small group of *innovators*, then *early adopters*, followed by the *early majority*, and finally the *late majority*.

To put this in the context of social media: a new word is first used within a narrow subcultural community, then spreads beyond its boundaries and either becomes part of the platform-wide vocabulary or disappears. On Twitter, this process can take anywhere from a few days to several months – in contrast to traditional language diffusion, which takes years or even decades.

Limor Shifman (Shifman, 2014) proposes another explanatory model within the framework of meme theory. Shifman reinterprets classic concepts and demonstrates that internet memes are not merely funny images, but units of cultural information that replicate, vary and compete for attention. Language patterns on Twitter behave in the same way: a successful construction replicates – it is copied, varied, reinterpreted and in the process becomes a stable unit of platform discourse. Shifman, analysing meme-based communication, rightly points out that a shared meme vocabulary acts as a social glue: knowing the current linguistic formulas means being inside the community, not knowing them means being outside it (Shifman, 2014).

#### **1.4. Semantic shift and neologism in the digital environment.**

The two central phenomena that form the main focus of practical analysis in this study are neologism (the emergence of new words) and semantic shift (a change in the meaning of existing words). A neologism is traditionally defined as a new word or expression that has appeared in a language and has not yet been fully assimilated into the linguistic consciousness of the majority of speakers (Algeo & Pyles, 2010, p. 252). In the context of Twitter, this definition requires clarification: the boundary between *new* and *assimilated* is extremely fluid here, as a word can go from zero usage to widespread use in a matter of weeks. V.Hofmann (2018) notes that Twitter-specific shifts often have a clear *tipping point* – a specific event after which a word begins to be used predominantly in its new meaning.

For this study, an operational criterion is applied: a neologism is defined as a word that was absent from the standard language reference corpora (BNC, COCA) prior to 2019, or that appears in them with a significantly lower frequency than in the Twitter corpus for the period under study.

Semantic change – a shift in a word's meaning without a change in its form – is an even older and extensively studied phenomenon. Classic types of semantic change include broadening of meaning, narrowing of meaning, melioration (improvement of connotations) and pejoration (deterioration of

connotations) (Ullmann, 1962, p.132). On Twitter, these processes occur particularly clearly and rapidly: a word can radically change its meaning within a few months under the influence of a specific cultural context or a viral event.

### **1.5. Syntactic compression and the pragmatics of punctuation in Twitter discourse.**

In addition to lexical changes, the language of Twitter/X exhibits significant syntactic and punctuation-related transformations, which also require theoretical analysis. The tendency towards syntactic compression – the shortening of sentences and the simplification of their structure is well known from studies of spoken language and written communication (Baron, 2010, p.17). However, on Twitter, this tendency takes on a particularly pronounced character: the character limit forces users to economise on syntactic resources to the utmost.

In the context of digital communication, punctuation marks are gradually acquiring new pragmatic functions not provided for by grammatical norms. In particular, a full stop at the end of a sentence – traditionally a neutral grammatical mark can, in a Twitter context, signal sharpness, insistence or passive aggression. H.P. Grice (1975) explains this through *conventionalised meaning*: a mark acquires a new function through the widespread repetition of a particular pattern.

This phenomenon is referred to as the *metapragmatic drift* of punctuation marks: the mark formally retains its grammatical function, but acquires an additional and sometimes dominant pragmatic meaning. Similar processes have parallels in spoken communication, where prosodic devices (intonation, tempo, pauses) perform pragmatic functions parallel to grammatical ones.

### **1.6. Critical discourse analysis as an interpretative tool.**

The quantitative methods employed allow us to identify trends and patterns. However, to understand the social and cultural context behind them, a qualitative level of analysis is required. To this end, this study employs critical discourse analysis (CDA) in the tradition of Norman Fairclough and Teun A. van Dijk. Fairclough (1992, 2003) argues that any text exists simultaneously on three levels: as a linguistic artefact (a set of lexical, grammatical and stylistic choices), as a discursive practice (a mode of producing, circulating and consuming text) and as a social practice (the reflection and construction of social relations, identities and power). This three-dimensional approach allows us to analyse a tweet not merely as a linguistic unit, but as a social action.

Teun A. van Dijk (2008) emphasises the importance of analysing the ideological dimensions of discourse: whose views and values it reproduces, whom it includes in the we and whom it excludes. Although the focus of this study is primarily linguistic rather than ideological, Teun A. van Dijk's approach is useful for interpreting the pragmatic strategies identified for example, how ironic linguistic formulas construct the identities of us and them within Twitter communities.

The language of social media is a distinct linguistic environment with its own mechanisms of innovation. It has been established that the linguistics of digital discourse has evolved from describing *Netspeak* as a general phenomenon to analysing platform-specific linguistic practices. The conceptual framework of the study has been defined: the concepts of linguistic innovation, neologism, semantic shift, syntactic compression and pragmatic drift. It has been established that the theoretical approaches of Labov, Rogers and Shifman explain the mechanisms of the spread of linguistic innovations, the approaches of Herring and Androutsopoulos – the role of the technological environment in the formation of linguistic practices, and the approaches of Fairclough and van Dijk – the socio-discursive dimension of the phenomena under analysis. The current state of research on this issue indicates that a systematic diachronic study of the language of English-speaking Twitter/X from the perspective of corpus linguistics and AI methods is a relevant and innovative field of research.

## Research Methods

### **2.1. Research material: selection criteria and corpus composition.**

This section describes the practical stages of the research: the principles governing the selection of material, the sequence of analytical steps, and the specific methods used to identify and interpret linguistic trends on Twitter/X.

#### 2.1.1. Rationale for the choice of platform and time frame.

The subject of this study – innovative and transformative trends in the language of contemporary social media – is broad and requires a focus on specific material. Among several possible platforms for analysis, Twitter/X has the greatest analytical value from a linguistic perspective, and this decision requires detailed justification, as the subject of the study covers social media in general. The first and most important reason is its text-centric nature. Twitter/X is the only major social network where text is not a supplementary element but a fundamental component of communication. On Instagram, text serves as a caption for an image. On TikTok, it acts as a comment on a video. On Twitter, text is the post itself and even when accompanied by media, it is the words that carry the main semantic load. This makes Twitter/X ideal material for linguistic analysis: here, the researcher deals with text in its purest form, rather than as part of a multimedia ensemble.

The second reason is its documented influence on the language as a whole. As demonstrated in the theoretical section, J.Grieve and I.Clarke in 2017 empirically demonstrated that Twitter is a leading source of lexical innovations in modern English: words first recorded on Twitter are significantly more likely to enter general usage than neologisms from other sources. This means that the trends identified in Twitter data are representative of broader processes in the language of social media and are not isolated phenomena. In other words, by analysing Twitter, we are analysing a leading vector of linguistic change in the digital space.

Thirdly, Twitter/X provides researchers with broad and reproducible access to data, which is a prerequisite for scientifically sound corpus research. The fourth reason is public accessibility: the vast majority of posts on Twitter/X are publicly available, which allows for the creation of a sample free from systematic biases associated with content privacy.

It is fundamentally important to emphasise that the trends identified in this study are not viewed as purely Twitter phenomena, but as manifestations of broader patterns in the language of social media. Lexical innovations that emerge spread to other platforms and into general usage; syntactic and punctuation conventions that have developed on Twitter are also observed by researchers on other digital communication platforms – Facebook Messenger and Reddit. In this paper, the platform is referred to as Twitter/X, as it changed ownership in October 2022 and officially changed its name from Twitter to X in July 2023. As the research material covers both periods of the platform’s existence – before and after the rebranding, the dual name Twitter/X is used in this paper as the standard method in academic literature for referring to the same platform following a change in its branding.

#### 2.1.2. Justification of the time frame.

The time frame of 2019–2025 was not chosen arbitrarily, but on the basis of several factors that make this particular period particularly significant for the study of linguistic innovations on social media.

The year 2019 serves as a starting point for several reasons. Firstly, it was then that Twitter increased the character limit from 140 to 280 for all users (this change took place at the end of 2017, but the widespread adaptation of linguistic practices to the new limit occurred throughout 2018–2019). Secondly, 2019 serves as a kind of pre-pandemic baseline – a point against which the changes brought about by the pandemic can be measured. Thirdly, it was precisely in 2019–2020 that Twitter’s audience grew sharply, making the platform an even more significant linguistic environment.

The years 2020–2021– the period of the COVID-19 pandemic are arguably the most significant for the language of social media in recent history. The mass shift of communication online, extremely high levels of social anxiety, and collective focus on the same topics and events created a unique environment for linguistic innovation. New words emerged and spread at an unprecedented rate: doomscrolling, quarantine fatigue, super-spreader (in a new, metaphorical sense), social distancing (in new syntactic constructions). An analysis of this period is key to understanding the mechanisms of linguistic innovation under conditions of social stress.

The years 2022–2025 brought two further significant factors. The first was the widespread adoption of generative AI tools (ChatGPT, Midjourney, Copilot), which introduced a new wave of lexical borrowings and reinterpretations into Twitter discourse: hallucinate (in the sense of generating false information), prompt, vibe coding, AI-generated. The second was Twitter’s

rebranding to X (2023), which in itself triggered numerous linguistic reactions on the platform and left a noticeable mark on the material under study. Thus, the six-year period spans several waves of linguistic innovation, allowing us to identify not only individual phenomena but also systemic trends.

### 2.1.3. Data collection procedure.

Data collection was carried out via the official Twitter/X account. The material was selected on the basis of four complementary principles. The first is relevance: each selected microcontext must directly illustrate one of the phenomena under investigation (lexical innovation, semantic shift, syntactic compression or pragmatic punctuation drift). The second is diversity: the selected examples must demonstrate different aspects, subtypes and contexts of use of the phenomenon under study, rather than repeating the same thing. Search queries were constructed to ensure thematic diversity in the sample. The third is authenticity: the material is taken exclusively from real, unedited posts. Only original posts, not retweets were included in the corpus. Fourth – chronological diversity: the examples cover the entire period under study (2019–2025), allowing the dynamics of the phenomena to be traced over time.

In practice, the selection process took place in two stages. In the first, exploratory stage, Voyant Tools was used to identify words and phrases with atypical frequencies or noticeable changes in frequency across different time periods—these became candidates for detailed analysis. In the second, analytical stage, specific tweets were manually selected that most clearly and comprehensively illustrated the phenomenon under investigation. ChatGPT assisted at this stage, in particular for the rapid selection of examples from a specific thematic cluster or to verify whether a particular example truly illustrated the phenomenon under investigation, rather than a similar but distinct one.

### 2.1.4. Composition of the dataset.

The dataset comprises seven balanced annual time series – 2019, 2020, 2021, 2022, 2023, 2024, 2025 – and is evenly distributed in terms of subject matter. The distribution across the years is relatively even, allowing for accurate comparison of the different time periods. Table 2.1 provides an overview of the data.

<b>Year</b>	<b>Approximate number of tweets</b>	<b>Remarks</b>
2019	280	baseline, pre-pandemic period
2020	310	the COVID-19 pandemic, lockdown

2021	290	post-pandemic, social movements
2022	300	AI tools, social tensions
2023	210	the widespread adoption of AI, rebranding on X
2024- 2025	300	latest data, partial snapshot

*Table 2.1. General characteristics of the Twitter/X corpus (2019–2025).*

It should be noted that the figures are approximate: the exact sample size varied depending on the availability of data via the API at the time of collection and the results of filtering out duplicates and irrelevant content. This is a common occurrence in social media research (McEnery & Hardie, 2012, p. 14).

### 2.3. Voyant Tools: Frequency and Thematic Analysis of Corpora.

#### 2.3.1. General Description of the Tool.

Voyant Tools is a free web-based text analysis tool developed by Stefan Sinclair and Jeffrey Rockwell (Sinclair & Rockwell, 2016). It is widely used in Digital Humanities and corpus linguistics thanks to its combination of powerful functionality and accessibility: no programming or installation of specialised software is required to use it. The platform allows users to upload a text corpus – individual files or groups of files and obtain several types of automated analysis in real time.

For this study, Voyant Tools was used in the initial, exploratory phase: to understand which words and expressions are most characteristic of the collected material and where the main analytical focus should be directed.

Specifically, the following programme features were used. Cirrus – a word cloud that clearly shows which lexemes are the most frequent in the uploaded text (after removing function words articles, prepositions, conjunctions using the built-in stop-word list). Comparing word clouds for different time periods provides an initial, indicative overview of how the lexical profile of the material has changed. In particular, a comparison of Cirrus clouds for 2019 and 2020 immediately reveals the emergence of new central words – doomscrolling, pandemic, quarantine, which are virtually absent from the 2019 cloud. Similarly, the 2022–2023 cloud shows an increase in words related to artificial intelligence: AI, ChatGPT, prompt, generate, hallucinate. This shift in the thematic core is the first indication that the language of Twitter responds to external socio-cultural events.

It is important to highlight the limitations of this feature: Cirrus displays only absolute frequency and does not take usage context into account. The word

*getting* may appear frequently for various reasons (it forms part of constructions such as *getting ratioed*, *getting tired* and *it's getting dark*). Therefore, the word cloud is merely a starting point for analysis, not the final result.

The Trends feature is the most valuable tool for this study, as it allows us to track how the frequency of a specific word has changed from one time period to another. The graph shows normalised frequency (the number of occurrences per total number of words), which allows for a fair comparison of subsamples of different sizes. It is Trends that demonstrates the very phenomenon described theoretically as the S-curve of Labov and Rogers: a gradual start, rapid growth and stabilisation. The Contexts (Key Word in Context, KWIC) function is a tool that allows you to view the immediate context of a specific word in a text: a few words to the left and a few to the right. This enables you to quickly assess the structures and meanings in which the word under investigation is used, without having to review each tweet individually. KWIC is particularly valuable for analysing semantic shifts. Let's look at how this works using the example of *ratio*. When reviewing the KWIC lines for this word in the 2019 data, we see the following context: price-to-earnings ratio, debt-to-equity ratio, compare the ratio of, the ratio between. All of this is a mathematical and financial context. In the 2021–2025 material, the same positions are occupied by: getting ratioed, ratio incoming, this tweet is getting, community ratio, hard ratioed. No mathematical context whatsoever. This change in lexical context is direct evidence of a semantic shift, even before the researcher begins to analyse specific tweets. A similar KWIC analysis was carried out for *slay*, *cancel*, *vibe*, *touch grass* and other identified keywords. In each case, a comparison of the contexts from 2019 and 2023–2025 confirmed or refined the hypothesis of a semantic shift. Voyant Tools is thus not only an exploratory but also, to some extent, an analytical tool.

The Summary function provides general statistics on the uploaded text: total word count, number of unique lexemes, and average document length. This data is useful for analysing syntactic trends, in particular, for tracking changes in average sentence length across different time periods.

A comparison of Summary statistics for the 2019 and 2025 subsets confirmed a noticeable reduction in average sentence length: in the 2025 material, the average number of words per sentence is lower than in 2019. The proportion of short sentences (up to 5 words) has increased, whilst the proportion of long sentences (over 20 words) has decreased. These changes are statistically modest but consistent with qualitative observations of syntactic compression and confirm that this trend is real, rather than a subjective impression of the researcher.

In practice, working with Voyant Tools looked like this: the collected text material was uploaded in separate blocks (by year or thematic category), after which word clouds and frequency trends were compared between the blocks. Words that saw a sharp increase in frequency in a given year or deviated significantly from expected standard language usage became candidates for

further detailed analysis. It was Voyant Tools, for example, that helped reveal that *doomscrolling* rose sharply in 2020, whilst *ratio* and *slay* became noticeably more frequent from 2021–2022 onwards.

It is important to be aware of the limitations of this tool. Voyant Tools does not compare frequencies against a standard language corpus, in other words, it shows what is common in our text, but cannot automatically confirm that this is specific to Twitter. Furthermore, the programme does not analyse meaning: it recognises the word *ratio*, but does not distinguish between its mathematical and Twitter-specific uses. Therefore, Voyant Tools' results should be viewed as a starting point for further qualitative analysis, rather than as a self-contained conclusion.

## 2.4. ChatGPT as an analytical assistant: principles and limitations of use.

### 2.4.1. The role of ChatGPT in the study.

ChatGPT (GPT-4, OpenAI, 2023) was used in this study as an analytical assistant and this role requires a clear conceptual definition. ChatGPT is not a research tool in the same sense as Voyant Tools: it does not perform automated calculations nor does it provide objective numerical indicators. Rather, it is an experienced *consultant* that helps the researcher quickly orient themselves within the cultural and pragmatic context of unfamiliar expressions, verify their own interpretation, and formulate hypotheses for further verification.

The theoretical basis for this application is provided by the study by F.Gilardi, M.Elizadeh and M.Kubli (2023), which showed that GPT-4 performs text classification tasks at a level comparable to that of a human analyst. This indicates the model's suitability for supporting analytical roles in linguistic research – albeit subject to mandatory verification of its responses.

ChatGPT (GPT-4 model, OpenAI, 2023) was used in the study in two roles. The first was to assist in classifying the selected examples. For each micro-context from the corpus, ChatGPT was asked to describe the pragmatic function of the tweet: whether it was ironic, humorous or sincere; what strategy the author was using; and to which cultural or linguistic tradition the expression used belonged. The responses were used not as a definitive interpretation, but as one voice in a dialogue alongside the reader's own analysis.

The second task was to establish the cultural and subcultural origins of expressions. Twitter slang often originates in specific subcultural contexts (AAVE, LGBTQ+ culture, fandom, gaming, financial communities), and understanding these origins is crucial for accurate linguistic interpretation. ChatGPT proved to be a useful tool for initially identifying such connections: for example, it confirmed the link between *slay* and AAVE (African American Vernacular English) and drag culture, as well as between *no cap* and basketball slang, which was subsequently verified using linguistic sources (Green, 2016).

It is necessary to describe the limitations of ChatGPT's use in this study honestly and in detail. Firstly, the model is generative: its responses are the result of statistical generation based on training data, rather than analysis in the

strict scientific sense. This means that ChatGPT can make mistakes, particularly regarding new or subcultural expressions that have emerged since the model's training was completed. Secondly, ChatGPT is prone to *hallucinations* – the confident assertion of false information. Therefore, any factual statement made by ChatGPT (for example, regarding the origin of a word or the specific date of its emergence) was always verified against linguistic sources or the corpus material itself.

Thirdly, ChatGPT is not a neutral tool in an ideological sense: it reflects certain cultural biases embedded in the training data. In particular, when analysing subcultural expressions from AAVE or LGBTQ+ culture, the model may produce overly normalised interpretations that smooth out the specificity of the cultural context. To compensate for this limitation, specialist academic sources were consulted when analysing such examples (Blodgett, Green, O'Connor, 2016; Smitherman, 2000).

The general principle governing the use of ChatGPT in this study can be summarised as follows: the model serves as a useful starting point for reflection and hypothesis testing, but it is not an authoritative source. All conclusions of the study are based on the analysis of real-world data and are supported by linguistic literature; ChatGPT serves as a supporting tool, not as a basis for argumentation.

## 2.5. Critical discourse analysis: procedure and principles.

### 2.5.1. Selection of a theoretical framework.

Critical Discourse Analysis (CDA) is the primary analytical method used in this study. The choice of CDA, rather than, for example, a purely corpus-based or purely pragmatic approach, is based on several considerations. Firstly, CDA allows text to be analysed simultaneously at several levels: linguistic, discursive and social, which corresponds to the complex nature of the phenomena under investigation. Secondly, CDA is context-sensitive: it does not consider linguistic units in isolation from their socio-cultural environment, which is of fundamental importance for the analysis of Twitter language, where context is often decisive for understanding meaning. Thirdly, CDA requires the researcher to be reflective regarding their own position and perspective, which is a methodologically honest approach for humanities research.

The methodological basis is Fairclough's three-dimensional model of discourse analysis (Fairclough, 1992, 2003). According to this model, each micro-context under analysis is examined at three levels: textual (specific linguistic features – vocabulary, grammar, punctuation, sentence structure), discursive (genre, rhetorical strategies, intertextual links, platform norms) and social (cultural and sociological context, identity construction, relations between the speaker and the audience).

### 2.5.2. Procedure for analysing microcontexts.

Each selected microcontext is analysed using a standard procedure that ensures consistency and reproducibility. The first step is to identify the

phenomenon under investigation: in the quoted tweet, the element that is the subject of the analysis is highlighted. The second step is the textual level: specific linguistic characteristics relating to the phenomenon under investigation are described. For lexical innovations, this includes the word-formation model, grammatical category and form; for semantic shifts, the nature of the change in meaning; for syntactic features, deviations from standard grammatical norms and the type of construction; and for punctuation phenomena, the specific punctuation mark and its position in the sentence.

The third step – the discursive level: the genre of the tweet is identified (ironic complaint, meme-based response, commentary on a pop-cultural event, etc.), intertextual links are established (which linguistic tradition or subculture the expression refers to), and compliance with platform conventions is assessed (whether this expression is typical of Twitter discourse in general or specific to a particular thematic cluster). The fourth step is the social level: an analysis is conducted to determine what identity the author constructs through the choice of this expression, which community they are appealing to, and what social or cultural context lies behind this usage. This step is the most interpretative and requires support from theoretical sources or additional examples.

The fifth step is comparative: if several examples are given to analyse a single phenomenon, they are compared with one another to identify common features and differences. It is precisely this comparison that is key to identifying the systematic nature of the phenomenon: if several different examples demonstrate the same pattern in different contexts, this indicates that we are dealing with a trend rather than an isolated instance.

### 2.5.3. The logic of combining methods: from breadth to depth.

The three methods used in this study – Voyant Tools, ChatGPT and critical discourse analysis. They do not form a parallel set, but rather a sequential analytical chain, where each subsequent method deepens and refines the results of the previous one. Voyant Tools provides a quantitative starting point and a roadmap of the material under investigation: which words are characteristic, where there are noticeable changes in frequency, and what thematic shifts occur across different time periods. ChatGPT helps users quickly navigate the cultural and pragmatic context of unfamiliar expressions. Qualitative analysis of examples provides depth: it explains exactly what is happening in a specific text and why this is linguistically significant. It is important that this sequence is unidirectional: Voyant Tools and ChatGPT serve as preparatory tools, not as the final authority. All conclusions are based on the analysis of real examples from the corpus.

## 2.6. Practical implementation: from data to analysis.

### 2.6.1. First stage: exploratory analysis in Voyant Tools.

In the first, exploratory stage, all the collected material was uploaded to Voyant Tools in the form of text files one per year (2019–2025). Initially, word clouds (Cirrus) were generated for each annual slice and for the entire dataset as

a whole. Comparing the clouds revealed words that had risen or fallen sharply between the time periods.

Next, the Trends function was applied to specific words that had attracted attention in the previous step. Frequency graphs were generated for: *doomscrolling*, *ratio*, *slay*, *cancel*, *vibe*, *touch grass*, *understood the assignment*, *no cap*, *NPC*, *enshittification*, and others. Each graph was interpreted in the context of chronological events: the sharp rise in March 2020 correlates with the onset of the pandemic; the rise in AI-related vocabulary in 2022–2023 with the widespread adoption of ChatGPT. These correlations are a key part of the argument: they confirm that linguistic innovations on Twitter are not arbitrary, but are directly linked to broader social processes.

In the third stage of the exploratory analysis, the KWIC function Contexts was used to identify semantic shifts: the lexical contexts of candidate words in early and late time periods were compared. Where the contexts changed significantly, this served as confirmation of a semantic shift and provided grounds for including that word in further detailed analysis.

#### 2.6.2. Stage Two: Formulating hypotheses with ChatGPT.

For each phenomenon selected in the previous stage, a preliminary hypothesis was formulated: what exactly is the linguistic change taking place and why? ChatGPT was used at this stage to verify and supplement these hypotheses. For example, the hypothesis that *slay* changed its meaning under the influence of AAVE was formulated independently based on an analysis of the material and then verified using ChatGPT and linguistic sources.

The sequence was always as follows: first, my own hypothesis; then, consultation with ChatGPT. Reversing this order could have introduced bias into the analysis: if ChatGPT offers an interpretation first, the researcher might unconsciously seek confirmation for that interpretation rather than for an independently formulated hypothesis. A match between my own hypothesis and ChatGPT's response reinforced my confidence in its correctness; any discrepancies became grounds for more detailed analysis and consultation with academic sources.

#### 2.6.3. Stage Three: Detailed Discourse Analysis.

In the third, main analytical stage, a detailed discourse analysis of representative micro-contexts was carried out for each selected phenomenon. For each lexical innovation or transformational trend, two or three (or more) tweets were selected that demonstrated different aspects of the phenomenon. The analysis was carried out according to the procedure described above – at the textual, discursive and social levels and concluded with a synthetic conclusion regarding the nature and systematic nature of the phenomenon under study.

An important part of this stage is the link between the analysis of specific examples and the overall argument. Following the analysis of each lexical unit or phenomenon, an interim conclusion is formulated that relates to the working hypothesis.

## 2.7. Ethical aspects of working with social media content.

Research into social media language raises several specific ethical questions that need to be addressed. The first concerns privacy: do researchers have the right to use public tweets without the authors' explicit consent? According to Twitter/X's Terms of Service and generally accepted principles of research ethics in the digital sphere (Association of Internet Researchers, 2019), publicly available posts are a legitimate subject of academic analysis and do not require specific consent from the authors, provided that the research is academic rather than commercial, and that the authors' personal data is not used in ways that could cause them harm.

Several measures have been taken in this study to minimise potential risks. Firstly, the focus is on linguistic practices and constructions rather than on the identities of the tweet authors: the analysis examines what is written and how it is written, rather than who wrote it. Secondly, for examples that are not public statements by well-known figures but 'private' messages from ordinary users, abbreviated or anonymised identifiers are used. Thirdly, any information that could identify the author outside their Twitter account is excluded from the analysis.

The second ethical issue concerns the adequacy of interpretation: does the researcher distort the meaning of tweets by taking them out of context? To minimise this risk, the study consistently adheres to the principle of minimal necessary context combined with the principle of sufficient explanation: each example provided is accompanied by a detailed description of the context necessary for its understanding.

### Results

This section provides a detailed account of the results of the analysis of a sample of English-language Twitter/X posts from 2019 to 2025. In accordance with the objectives set out in the introduction, the analysis was conducted at three levels of linguistic organisation: lexical, syntactic and punctuation, as well as in a diachronic dimension encompassing all levels. Each subsection follows a consistent structure: a general description of the identified phenomenon, quantitative data from Voyant Tools, a detailed qualitative analysis of representative micro-contexts, a theoretical interpretation with references to the academic literature.


The sample comprises purposefully selected microcontexts from five thematic categories. The principle of purposeful sampling in qualitative linguistic research is widely accepted and well-established in methodological literature: it allows one to focus on the most illustrative examples rather than on a statistically averaged picture (Patton, 2014, p. 277). All the examples given are authentic posts, verified against the original source; references to the source are provided immediately after the quotation. The phenomenon under investigation is highlighted in each example; the analysis unfolds in the paragraph following the example.

### 3.1. Overall picture: what is specific to the language of Twitter/X.

The starting point for the analysis is the question: what makes the language of Twitter/X different from standard English? Analysis using Voyant Tools (the Cirrus function and comparison of frequency profiles of the material by thematic blocks) revealed a layer of vocabulary that is characteristic of the collected sample and atypical for the general language. It is important to highlight a methodological limitation: Voyant Tools does not automatically compare the material with a reference corpus; therefore, the conclusion regarding the specificity of certain forms is based on a combination of quantitative observations (high frequency in specific years, absence in general dictionaries with the recorded meanings). This is an acceptable approach for a qualitatively oriented study, which forms the methodological framework of this work.

The Cirrus function revealed that, among the 50 most frequent significant lexemes in the sample (after removing function words), a significant proportion are either neologisms or words used in non-standard meanings. The Trends function confirmed a sharp change in frequency profiles between 2019–2020 and 2022–2025: a number of terms are virtually absent from the 2019 material and become noticeable from 2020–2021 onwards. The Contexts function allowed us to examine the specific contexts in which the words key to the analysis are used, which proved to be an important tool for identifying semantic shifts.

Based on the results of quantitative and qualitative analysis, the lexical innovations identified in the sample were classified into four types, the approximate ratio between which is shown in Fig. 3.1.

Neologisms		~34%
Semantic changes		~31%
Memes		~22%
Abbreviations		~13%

*Fig. 3.1. Approximate distribution of identified lexical innovations by type (based on an analysis of the sample in Voyant Tools and qualitative analysis).*

Neologisms (around 34%) are new lexical forms that were not included in authoritative dictionaries (Merriam-Webster, Oxford English Dictionary) prior to 2019, or were listed there only with meanings fundamentally different from those identified in the sample. This group includes: *doomscrolling*,

*enshittification*, *deinfluencing*, *finsta* (fake + Instagram), *goblincore*, *gaslit* (in a broader figurative sense), *rizz*. Semantically reoriented words (around 31%) – existing lexemes that have acquired new, sometimes radically different meanings within the context of the platform: *ratio*, *slay*, *cancel*, *vibe*, *tea* (meaning gossip, scandalous information), *sus* (suspicious), *era* (in expressions such as *in my healing era*). Meme formulas (around 22%) – expressions that are grammatically complete sentences or phrases, but pragmatically function as conventional units of approval, disapproval or self-identification: *understood the assignment*, *no notes*, *no cap*, *main character energy*, *not the...*, *living rent free*. Grammaticalised abbreviations (around 13%) – abbreviations and acronyms that have ceased to be mere abbreviations and function as full-fledged grammatical units: *NPC* (non-playable character) in a figurative sense, *POV* (point of view) as a genre marker, *IYKYK* (if you know you know) as a formula of esoteric solidarity.

These four types are not mutually exclusive: some units combine features of several types. For example, *era* in the expression *in my villain era* is both a semantic shift (a new meaning for a word that existed previously) and a memetic formula (a conventional construction with a fixed pragmatic function). Such overlap is to be expected and corresponds to the principle of category prototypicality in cognitive linguistics (Rosch, 1971): clear boundaries between categories are an artefact of classification, not a reality of the linguistic system.

### 3.2. Neologisms: mechanisms of emergence and consolidation.

#### 3.2.1. Doomscrolling: Case Studies of Lexical Innovation.

Doomscrolling is perhaps the most studied neologism that emerged on Twitter during the period under study, and at the same time the most illustrative case for understanding the mechanisms of lexical innovation in social networks. The word is a morphological blend of *doom* (death, destruction, or a sense of inevitable fate) and *scrolling* (continuous navigation of digital content, including news feeds or social media interfaces): both parts are phonologically and semantically motivated. The first use of doomscrolling on Twitter is documented by researchers around October–November 2019 (Merriam-Webster Online, 2020), although it has been significantly present in the sample only since March 2020, which coincides with the beginning of mass quarantine in connection with the COVID-19 pandemic.

The corpus findings demonstrate the rapid stabilization of this lexical innovation within digital discourse related to social media consumption and crisis communication. The analyzed subcorpus displayed a vocabulary density of 0.330, a readability index of 9.902, and an average sentence length of 18.1 words, reflecting the comparatively concise and conversational style characteristics of Twitter/X communication. The most frequent lexical items in the corpus were *doomscrolling* (138 occurrences), *Twitter* (24), and *news* (16). The frequent co-occurrence with *news* and *Twitter* further demonstrates that the

term is strongly connected to the experience of continuous exposure to negative information within a platform-mediated environment.

The Trends function in Voyant Tools recorded a characteristic S-shaped dynamics of the spread: practically absent until February 2020, rapid growth in March–April 2020, gradual stabilization in 2021–2022. This dynamics exactly corresponds to the Rogers model (Rogers, 2003): a new lexical item goes from first use (innovators) through mass diffusion (early and late majority) to stabilization. Significantly, doomscrolling has moved from Twitter to common usage. After three years on the Merriam-Webster watch list, the word *doomscrolling* was recognized as an official word in September 2023. Dictionary.com selected it as the top monthly trend in August 2020. Macquarie Dictionary named doomscrolling the 2020 Committee Word of the Year. This supports the thesis by Grieve et al. (2017, p.100) that Twitter is a leading incubator of lexical innovation.

The TermsBerry visualization provides evidence for the sociocultural grounding of neologism. The presence of lexical associations related to the COVID-19 pandemic confirms that doomscrolling emerged and gained prominence in direct connection with the pandemic information environment, where users were repeatedly exposed to alarming news, statistics, and crisis-related updates. This supports the broader argument that major social disruptions function as catalysts for accelerated lexical innovation in online communication.

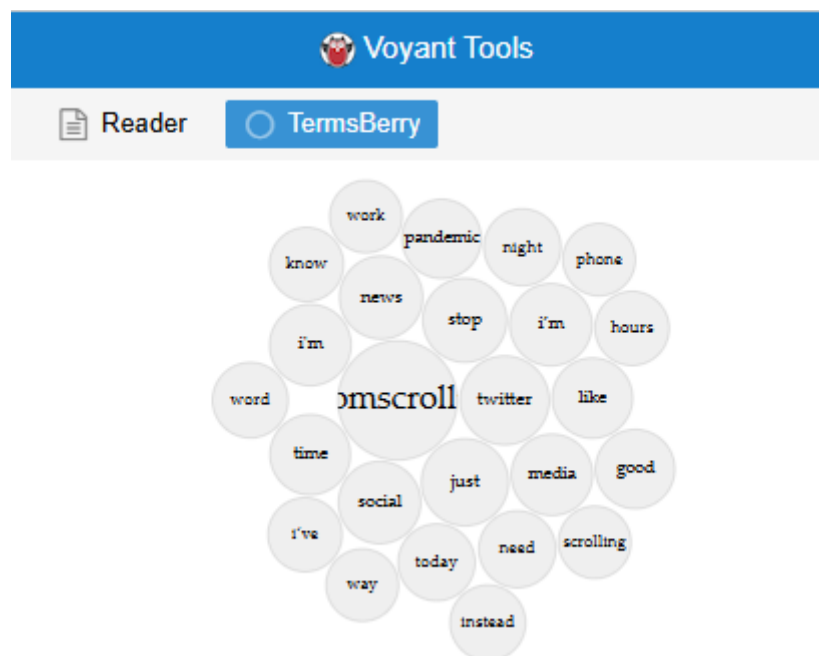


Figure.3.2. Voyant Tools TermsBerry visualization.

Morphologically, *doomscrolling* is a blend – an extremely productive word-formation model in the language of the Internet. Other blend neologisms in the sample include: *finsta* (fake + Instagram, a private Instagram account); *fomo* (fear of missing out, although this word appeared earlier, in the sample it

undergoes a new semantic expansion); *fleabag* (in a figurative sense, from the name of the TV series). Blends are an attractive word-formation model for social networks for several reasons: they are concise (one word instead of a phrase), easily recognizable (both parts are transparent), and semantically capacious (preserve the connotations of both components).

*doomscrolling is such a terrible and fantastic word to describe what i've been doing for the past week* (Twitter/X, @PatosRoxas ,8:21 PM · Mar 21, 2020.)

In this case, it is advisable to start the analysis from the metalinguistic level. Unlike the previous examples, where doomscrolling functions as an element of action description, here it becomes an object of reflection: the linguistic unit is understood as a word, and not only as a designation of experience. Such a metalinguistic perspective is an important feature of the initial stage of consolidation of a neologism, when speakers not only use a new lexeme, but also evaluate its form, sound and correspondence to the phenomenon denoted.

Grammatically, doomscrolling functions as a subject in the structure *doomscrolling is...*, i.e. it implements the syntactic position typical for a noun. At the same time, in the second part of the statement (to describe what I've been doing) it corresponds to a continuous action, expressed through the Present Perfect Continuous (I've been doing). Thus, there is an interesting tension between nominalization (as the name of a phenomenon) and processivity (as a repeated behavior), which reflects the hybrid nature of this neologism.

At the semantic-pragmatic level, the key is the combination of the adjectives *terrible* and *fantastic*, which creates an oxymoronic effect. This combination reflects the speaker's ambivalent attitude towards the phenomenon itself: on the one hand, doomscrolling is assessed as a negative, harmful practice (terrible), on the other hand, as an apt, even *brilliant* linguistic nomination (fantastic word). Such a double assessment indicates a high degree of semantic adequacy of the neologism: it accurately captures the experience, even if the experience itself is undesirable.

From a pragmatic point of view, the utterance performs the function of self-reflexive recognition: the speaker does not simply describe his behavior, but distances himself from it through linguistic assessment. This corresponds to the trend described in studies of digital discourse, where naming the problem (naming the experience) is the first step towards its comprehension and collective discussion. Thus, doomscrolling in this example functions not only as a lexical innovation, but also as a tool for conceptualizing experience.

*Its been really difficult to keep from doomscrolling everyday. The constant barrage of injustice, corruption, stress, and feeling helpless to change it has been difficult to learn how to navigate. Here's some pictures I've taken that help me feel calm. What makes you feel calm?* (Twitter/X, @Kumbamonster\_, 12:15 PM · Aug 2, 2020)

In this example, *doomscrolling* is integrated into a more complex syntactic and discursive structure. At the grammatical level, it functions as a gerund after the construction *keep from + V-ing* (*keep from doomscrolling*), which indicates an attempt to avoid a certain action. This construction implies controllability of behavior, but the adverb *really* in combination with the adjective *difficult* emphasizes that this control is problematic.

At the semantic level, the causal structure of the statement is important. *Doomscrolling* appears not as an isolated action, but as a reaction to *the constant barrage of injustice, corruption, stress*, that is, to the information overload of negative content. In this context, the word acquires an additional semantic dimension: it denotes not only the user's behavior, but also a broader media ecological phenomenon – a constant flow of disturbing information. Pragmatically, this tweet differs from the previous ones in that it moves from an individual confession to interactive communication. The final question *What makes you feel calm?* transforms the statement from a monologue into a dialogue, involving the audience in a joint discussion. The insert *Here's some pictures I've taken that help me feel calm* performs the function of a positive counterpoint: it offers an alternative behavioral strategy in response to the negative impact of *doomscrolling*.

Discursively, this example demonstrates the therapeutic function of social networks described in digital communication studies: the platform is used not only to express anxiety, but also to find emotional support and exchange coping strategies. Thus, *doomscrolling* here acts as a trigger for a broader conversation about the psycho-emotional state of users.

Comparing these two examples allows us to clarify the dynamics of the functioning of neologism *doomscrolling*. In the first case, it is an object of metalinguistic evaluation and a means of conceptualizing experience, while in the second, it is an element of a causal narrative and a trigger for social interaction. Such variability of functions confirms that the word has not only become entrenched at the lexicon level, but has also been integrated into various pragmatic scenarios of digital discourse. It is this functional flexibility that is the key condition for its successful adaptation in the linguistic system of social networks.

### 3.2.2. Enshittification: Neologism as Social Criticism.

*Enshittification* is an example of a neologism that arose not spontaneously, but in response to a specific conceptual question: what to call the systematic deterioration of the quality of large online platforms over time? The word was introduced by Canadian writer and journalist Cary Doctorow, who first used the term *Enshittification* in November 2022 in one of his blog posts, which was republished in *Locus* magazine in January 2023. He describes a three-stage process: first, the platform attracts users with high-quality service; then worsens conditions in favor of partner companies; finally, degrades for both, extracting maximum profit before decline. The word is morphologically

expressive: prefix en- (as in enlarge, enhance) + compound shit + suffix -ification (a process that transforms something into something else, as in *beautification*, *gentrification*). The Trends function recorded the appearance of *enshittification* in the sample in 2023 and a stable presence in 2024.

The corpus analysis of *enshittification* illustrates the rapid consolidation of the term within contemporary technological and platform-critical discourse. The analyzed subcorpus consisted of one compiled document containing 3,616 total word tokens and 1,270 unique word forms, with a vocabulary density of 0.351, a readability index of 11.919, and an average sentence length of 22.6 words. Compared to the corpora of more conversational innovations such as *doomscrolling*, the higher readability index suggests a more analytical and commentary-oriented style, characteristic of discussions surrounding technology, digital platforms, and online culture.

The most frequent lexical items in the corpus were *enshittification* (118 occurrences), *AI* (22), *Twitter* (18), and *tech* (13). The dominance of technology-related vocabulary demonstrates that the neologism is strongly embedded in discussions about digital infrastructures, platform economics, and the perceived decline in the quality of online services. Particularly noteworthy is the frequent association with AI, which reflects the expansion of the term beyond its original platform-centered usage. In the analyzed material, *enshittification* is increasingly applied to broader technological anxieties connected with algorithmic systems, generative AI, and automated content production. This semantic expansion demonstrates that neologism has evolved into a generalized critical label for perceived technological decline.

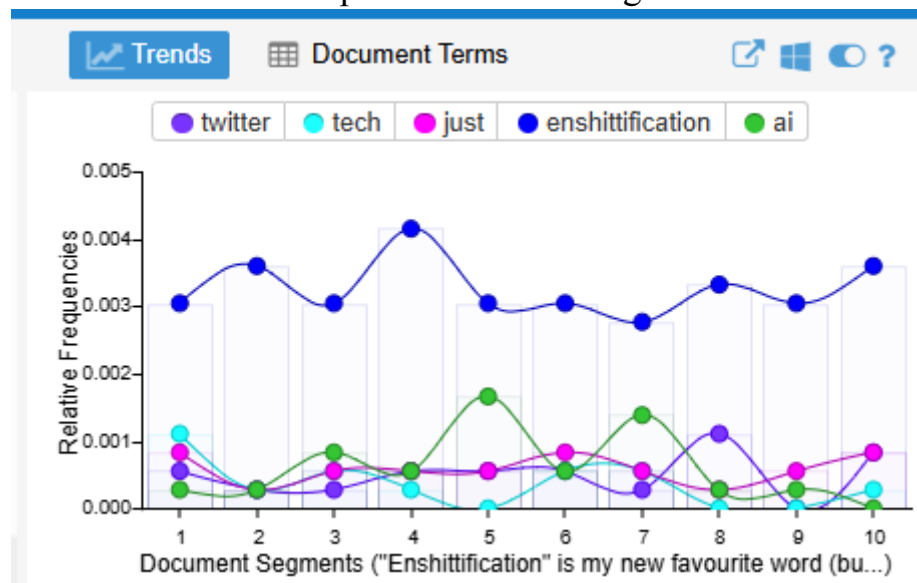


Figure.3.3. Voyant Tools Trends graph.

*The gradual degradation of the internet and "enshittification" is a decent case study of how capitalism determines technology, I think. We've never seen how technology could change outside of capitalism, and it could be quite different than it is.* (Twitter/X, @ComradeQuail, 7:23 PM · Dec 11, 2024)

This example is indicative primarily in terms of the conceptualization of a new word in a broader ideological discourse. First, *enshittification* is integrated into a complex nominative construction (the gradual degradation of the internet and enshittification), where it functions as an abstract noun alongside the more traditional lexeme degradation. Such coordination indicates a high degree of lexicalization: the neologism is already able to function as a full-fledged term in a theoretical generalization.

Secondly, the use of quotation marks (“*enshittification*”) signals a certain distance of the speaker from the word: it is simultaneously used and marked as new or stylistically atypical. This corresponds to an intermediate phase of lexicalization, when the unit is already understandable to the audience, but still retains a touch of novelty or jargon.

On a pragmatic level, the statement performs the function of argumentation: *enshittification* is presented as an empirical case study for a broader thesis about the impact of capitalism on technology. Thus, the word moves from the level of everyday description of experience (as in the case of *doomscrolling*) to the level of analytical tool. This indicates a semantic expansion: it no longer denotes only the process of deterioration of digital platforms, but also becomes a category for interpreting socio-economic phenomena.

Discursively, this tweet demonstrates the integration of the neologism into political and critical discourse. The construction I think performs the function of mitigation (hedging), characteristic of online discussions, where the author simultaneously expresses a position and leaves space for dialogue. Thus, *enshittification* here functions as an element of semi-formalized theorizing, which is an important stage in the process of its normalization.

*"Enshittification" is my new favorite word. A year old but sharing bc it's that good & still relevant, so I implore you to read in case you missed it too. Also don't be deterred by the "TikTok" title, it's only part of the story (cc: @WIRED @doctorow) . (Twitter/X, @jaimelin, 3:34 AM · Feb 16, 2024 )*

This example represents the metalinguistic and affective dimensions of the functioning of a neologism. As in the case of the previous analysis of doomscrolling, the word *enshittification* acts as an object of evaluation (my new favorite word), and not just a tool for describing reality. This indicates a high level of awareness of the speakers of the linguistic innovation itself. Grammatically, *enshittification* functions as a subject in a simple nominal predicate (is my new favorite word), which corresponds to the typical structure of metalinguistic statements. At the same time, quotation marks again perform the function of marking the novelty or stylistic expressiveness of the word.

At the pragmatic level, the presence of appeal constructions (*I implore you to read, in case you missed it*) is important, which transform the statement from an individual statement into a call to action. The abbreviation bc (because) and the symbol & are manifestations of graphic compression, characteristic of

Twitter discourse, where economy of linguistic means is combined with high speed of communication.

Mention (cc: @WIRED @doctorow) performs an intertextual function: it embeds the tweet in a broader media and expert context associated with the popularization of the term. This is an example of the network circulation of neologisms, when a word spreads through references to authoritative sources and influential users.

Pragmatically, this tweet performs a popularization function: the author not only uses the word, but also actively promotes its spread, assessing it as *successful* and *relevant*. Such an assessment is important for the process of diffusion of innovation, since, according to Everett Rogers (Rogers, 2003), a positive perception of novelty significantly accelerates its adoption by the community.

### 3.2.3. Rizz and goblincore: neologisms from the subcultural environment.

Not all neologisms in the sample are associated with social crises or technological changes. *Rizz* and *goblincore* are examples of words that came to Twitter from narrow subcultural environments and spread thanks to the mechanism of meme replication described by Shifman (Shifman, 2013). *Rizz* is a contraction of charisma (or a foreign root related to charm), which refers to natural attractiveness and the ability to easily get along with and attract people. The word comes from New York street slang and became widespread on Twitter after being used by streamer Kai Cenat in 2022.

A corpus analysis demonstrates the high level of integration of this unit into contemporary online youth discourse. It contains 621 word occurrences and 305 unique word forms, with a vocabulary density of 0.491, indicating a relatively diverse lexical composition given the small size of the corpus. The readability index (7.181) reflects the tendency towards syntactically simplified, colloquial speech characteristic of Twitter communication, whilst the average sentence length – 41.4 words, demonstrates users' propensity for emotionally expansive utterances and metacomments. The most frequent unit in the corpus is *rizz* (48 occurrences), whilst the accompanying vocabulary is dominated by *like* (8), *just* (8), *word* (4) and *women* (3). The frequency of the discourse markers *like* and *just* further confirms the informal and affective nature of the speech under investigation.

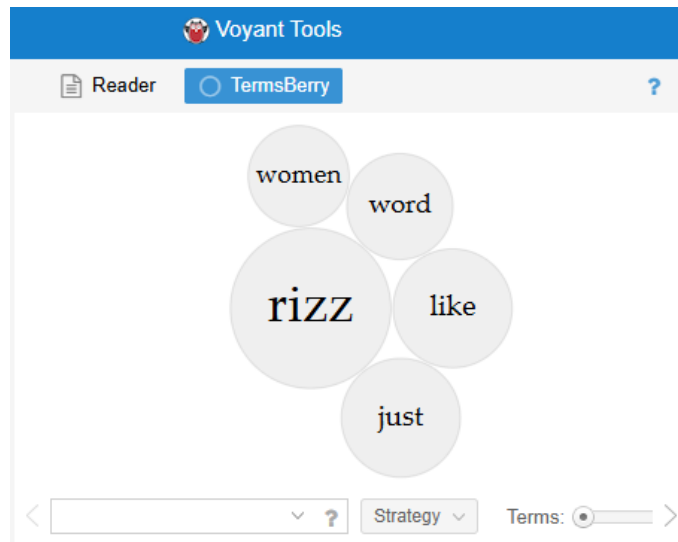


Figure 3.4. Voyant Tools TermsBerry visualization demonstrating the dominant lexical associations surrounding *rizz*.

Trends found a gradual increase in the frequency of *rizz* in the sample since 2022, with a relatively flat profile, unlike the explosive *doomscrolling*. This suggests a gradual process of diffusion through early followers and early majority in Rogers' terms, without a clear explosion moment.

*I just feel like people with real rizz would use the word 'charisma' instead.* (Twitter/X, @hastingsmhs\_, 5:35 PM · Dec 28, 2022 )

The lexeme *rizz* functions as a key marker not only of a certain personal quality, but also of the sociolinguistic identity of the speaker. At the grammatical level, *rizz* is used as a noun in the semantic construction *people with real rizz*, where the adjective *real* performs an intensifying function. Such a combination is indicative: it assumes the existence of gradations within the concept itself (*real* vs. *implicit incomplete* or *pseudo-rizz*), which indicates the semantic development of the lexeme. Thus, the word no longer only names a phenomenon, but also allows for internal differentiation, which is a sign of its active integration into the language system.

At the pragmatic level, the statement is constructed as a subjective evaluation (*I just feel like*), which serves as a hedging device and at the same time signals belonging to an informal, colloquial register. At the same time, the sentence itself contains an implicit normative setting: a real carrier of this quality does not use the slang designation *rizz*, but prefers the standard equivalent *charisma*. Thus, the statement forms an opposition between two lexical codes – youth slang (*rizz*) and neutral standard (*charisma*).

This opposition is sociolinguistically significant. The use of *rizz* marks belonging to a certain online community or generational layer, while *charisma* is associated with more formal, “educated” or traditional speech. The paradox of the statement is that the speaker himself uses *rizz* while simultaneously claiming that “real” bearers of this quality would choose a different word. This creates an effect of ironic distance and a metalinguistic game with registers.

From a cognitive-semantic point of view, *rizz* functions as a generalized label for a complex of traits (charisma, attractiveness, social confidence) that do not have clearly defined boundaries. In this sense, it can be interpreted within the framework of prototype theory (Rosch, 1971): the meaning of a word is not determined by strict necessary and sufficient features, but rather by a prototype – an intuitively recognizable image of a “person with rizz”. That is why it is possible to modify it by real, which appeals to the authentic manifestation of this quality.

Goblincore is an example of an aesthetic neologism – a name for a holistic cultural and stylistic phenomenon. It denotes a subcultural aesthetic that rejects traditional standards of beauty and order and instead glorifies the ugly, natural, chaotic – mushrooms, swamps, stones found in the forest, silted objects. Goblincore originated on Tumblr around 2019–2020 and spread through Twitter and TikTok. According to Ammanda Brennan, a digital cultural researcher and widely recognized expert on internet mem culture, the aesthetic and linguistic phenomenon associated with *goblincore* began emerging in online communities in the spring of 2019 and gained significant momentum during the COVID-19 pandemic, when users increasingly encountered and circulated the term across social media platforms. This development illustrates how subcultural neologisms can rapidly diffuse through digitally mediated communication. Structurally, it is a complex word: goblin + core (a suffix denoting a subcultural aesthetic, following the model of cottagecore, darkcore, normcore). Core as a productive suffix of subcultural names is a separate interesting phenomenon: it appeared around the 2010s and actively spread in the TikTok-YouTube-Twitter space.

*Spotify really said that my music taste was “mystical ominous goblincore” and I’ll never forget that.* (Twitter/X, @amytcun, 7:10 PM · Dec 16, 2022.)

The key object of the analysis is the complex nominative construction *mystical ominous goblincore*, where the neologism *goblincore* functions as the core of the attributive chain.

At the grammatical level, a sequence of prepositional meanings (*mystical, ominous*) typical for the English language is observed, modifying the main noun *goblincore*. Such a structure is indicative, as it indicates the full integration of the neologism into the syntactic system: *goblincore* behaves like a regular noun, capable of accepting meanings and acting as part of a complex noun group. This is an important indicator of lexicalization – the word not only exists, but is also actively combined with other elements.

At the pragmatic level, the construction *Spotify really said that...* performs the function of personifying the platform: the algorithm is represented as a speech agent that expresses a judgment about the user. The adverb *really* here acts as an intensifier, enhancing the effect of surprise or even indignation: the speaker emphasizes that such a classification seems to him at the same time unexpected and revealing. The use of quotation marks (“*mystical ominous goblincore*”), which mark quotability and at the same time distance, is also key.

The speaker reproduces the label provided by the algorithm, but does not fully internalize it, which creates an ironic effect. The statement *I'll never forget that* is hyperbolic, typical of Twitter discourse, and emphasizes the emotional reaction to the categorization.

From a semantic point of view, *goblincore* is an example of a subcultural neologism denoting an aesthetic associated with imperfect nature. In this context, it is combined with the adjectives mystical and ominous, which expand its meaning, adding shades of mystery and gloom. This indicates the openness of the concept: *goblincore* does not have rigidly fixed boundaries and is easily modified depending on the context, which corresponds to the characteristics of prototype theory (Rosch, 1971).

In a discursive sense, this tweet illustrates the interaction between algorithmic categorization and the user's self-presentation. The speaker simultaneously accepts and ridicules the identity imposed on him, which is typical of the digital environment, where algorithms form symbolic labels. The ironic reproduction of this category creates an effect of solidarity with the audience: the recognizable experience of strange algorithmic descriptions becomes the basis for humorous communication.

3.3. Semantic shifts: when an old word acquires a new meaning.

3.3.1. Theoretical reminder and operational criteria.

Before proceeding to the analysis of specific examples, it is necessary to clarify the operational criteria for identifying a semantic shift in the material. A word is considered to have undergone a semantic shift if: (a) in Voyant Tools (Contexts function) its typical collocates in the dataset differ significantly from those expected under its conventional dictionary meaning; (b) ChatGPT, when analyzing the context, cannot use the dictionary meaning of the word to adequately describe the situation in the tweet; (c) in at least three examples of the sample the word is used in a new meaning without any explanation - which indicates the conventionality of the new meaning for the target audience. These three criteria, taken together, provide the minimum necessary evidentiary basis for concluding about a semantic shift.

3.3.2. Ratio: From Mathematics to Social Judgment.

*Ratio* is perhaps the most telling example of platform reinterpretation – a type of semantic shift in which a word takes on a specific meaning that is understandable and relevant only in the context of a particular platform. In common English, *ratio* is a mathematical term: the ratio of two quantities to each other. In Twitter/X, it has taken on a new meaning: a situation where the number of replies to a post significantly exceeds the number of likes. On the platform, this is a conventional signal: if a tweet is replied to more than liked, it means that it is a failure, scandalous, or one that causes mass disagreement. This mechanism is closely tied to the Twitter architecture: it is on this platform that the *ratio* between replies and likes is visible and meaningful to all participants.



At the pragmatic level, the adverb constantly plays an important role, reinforcing the sense of continuity and intrusiveness of the process. *Ratioed* in this context is no longer limited to its primary meaning in Twitter discourse (a situation where a reply receives more interaction than the original tweet), but is expanded to a metaphorical meaning – to be outperformed, to be devalued. The addition *by your own low self esteem* creates an internal agent effect: instead of an external audience, the source of *ratio* becomes one's own low self-esteem.

This leads to an important semantic shift: the social-network mechanism of evaluation is transferred to the intrapsychological space. Thus, *ratioed* functions as a conceptual bridge between digital and emotional experience, which is characteristic of the language of social networks, where technical terms are often reinterpreted as markers of personal states.

Discursively, the statement implements a strategy of ironic self-reflection. Comparing external metrics (views) with an internal state (low self esteem) creates the effect of hyperbolized self-esteem, which is both humorous and critical. This type of utterance can be attributed to the genre of performative self-deprecation (Marwick & Boyd, 2011), where the speaker publicly constructs his own vulnerability as a form of communicative interaction.

*It is the mass positive feedback Elon gets via his audience, which is one of the largest on Twitter period, that gives him confidence. Obviously he is tacking hard to the right but he's functionally "ratio-proof" in any direction. Sounds stupid but this is the game on Twitter.* (Twitter/X, @PopulismUpdates, 10:25 PM · Dec 11, 2022)

The key to the analysis is the adjective *ratio-proof* as a derivative form of the already lexicalized *ratio*. The analysis should begin at the morphological level. *Ratio-proof* is formed according to the productive word-formation model N + -proof (cf. waterproof, bulletproof, foolproof), where the suffixoid -proof denotes immunity or resistance to a certain influence. In this case, we are talking about immunity to *ratio* as a form of public disapproval. The very fact of the appearance of such a derivative form is indicative: the new meaning of *ratio* has already stabilized to such an extent that it has become the basis for further word-formation. As L.Bauer (2001) notes in *Morphological productivity*, the formation of derivative units through productive word-formation patterns can be regarded as evidence of a lexical item's increasing entrenchment within the linguistic system.

At the grammatical level, *ratio-proof* functions as a predicative designation in the nominal part of the predicate (*he's functionally ratio-proof*). The adverb functionally acts as a modifier, clarifying that it is not about absolute, but about practical, empirically observed unavailability for negative reaction. This gives the expression an analytical character and brings it closer to the quasi-scientific style typical of political commentary on social networks.

Semantically, *ratio-proof* represents a further stage of abstraction of meaning: if to be *ratioed* means to experience public disapproval through

interaction metrics, then *ratio-proof* means to be structurally protected from such a mechanism. In the text, this is explained by mass positive feedback and a large audience, that is, by the asymmetry of communicative power. Thus, the word goes beyond the description of a separate situation and begins to denote a stable social position within the platform.

At the pragmatic level, the statement performs a rationalization function: the author explains the behavior of a public figure through the logic of platform mechanisms (this is the game on Twitter). Such a metaphor (game) conceptualizes social interaction as a system of rules and strategies, within which *ratio-proof* acts as an advantageous position. The insertion sounds stupid but is a marker of discursive mitigation (hedging), which foresees possible criticism and at the same time preserves the authority of the statement.

In discursive terms, this example demonstrates the institutionalization of the neologism: *ratio* and its derivatives are already used not only in a humorous or everyday context, but also in the political analysis of digital platforms. Thus, *ratio-proof* illustrates the final phase of lexicalization, when innovation is not only fixed in the language, but also becomes a productive basis for the formation of new conceptual categories within the framework of a broader social discourse.

### 3.3.3. Slay: Semantic amelioration through Subcultural Diffusion.

*Slay* constitutes a classic example of semantic amelioration – the process of a positive re-evaluation of a word’s connotative meaning, which Stephen Ullmann classified as one of the principal types of semantic change (Ullmann, 1962, p. 230) as one of the basic types of semantic change. The word comes from a Proto-Indo-European root meaning *to beat, kill* and in Standard English means *to kill* or *to slay with laughter*. In the context of AAVE and drag performance culture, where performers *slay* their performance (i.e., captivate the audience into oblivion), it has acquired an exclusively positive meaning of to be stunning, to look or act flawless. This cultural root is described in detail in *Demographic Dialectal Variation In Social Media* (Blodgett, Green, O’Connor, 2016, p. 1127) and is supported by the analysis in ChatGPT.

Corpus analysis conducted using Voyant Tools on Twitter/X posts further confirms the high degree to which the word’s new semantics has become established. The microcorpus under investigation consists of 729 word occurrences and 379 unique word forms, indicating relatively high lexical diversity (vocabulary density – 0.520). The readability index is 8.578, and the average sentence length is 28 words, reflecting a combination of colloquial expressiveness and syntactic elaboration characteristic of affective Twitter discourse. The most frequent unit in the corpus is the form *slayed* (52 occurrences), whilst the base form *slay* appears only 4 times. This disparity is significant, as it indicates the active functioning of the verb’s reinterpreted, evaluative meaning in everyday digital communication.

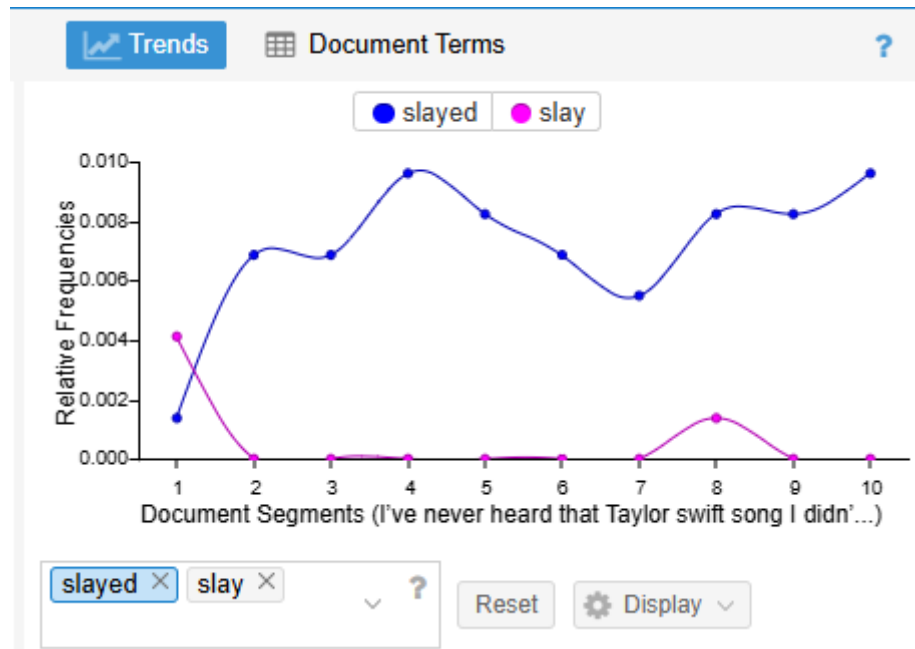


Figure 3.6. Voyant Tools Trends visualization showing the frequency distribution of *slayed* and related lexical items in the analyzed Twitter/X corpus.

The *Trends* feature in Voyant Tools has recorded a noticeable increase in the frequency of *slay* in new meanings since 2021, with a steady peak in 2022–2023. This is consistent with the general chronology of the word's spread: Merriam-Webster added it to the dictionary in a new meaning in September 2022.

*I GOT MY SEMESTER FINALS RESULTS AND FAIR TO SAY I ABSOLUTELY SLAYED THIS TIME (thank god I did cuz otherwise I'd be jumping off a cliff)* (Twitter/X, @evecelia6, 9:28 PM · Dec 28, 2023 ·)

The analysis should begin at the grammatical level. *Slay* is used in the Past Simple form (*slayed*), which is indicative: the word, which originally meant *to kill*, in modern digital discourse functions as a verb of evaluation with the meaning *to achieve significant success, to accomplish something brilliantly*. The regular form *slayed* (despite the existence of the normative form *slew* in the base sense) indicates a complete reanalysis of the verb: the new sense functions independently of the traditional grammatical paradigm. This is an important indicator of lexicalization, as it demonstrates that speakers process *slay* as a separate, autonomous unit.

At the pragmatic level, the construction *absolutely slayed* performs the function of an intensified positive assessment. The adverb *absolutely* enhances the categorical nature of the statement, creating the effect of total success without any caveats or critical components. This formula is typical of the affective style of Twitter discourse, where emotional expression is often hyperbolized.

The paragrapheme level is also significant: the use of capital letters (*I GOT MY SEMESTER FINALS RESULTS*) performs the function of prosodic amplification, similar to increasing the volume or intonation stress in oral

speech. This creates the effect of an exclamation and emphasizes the emotional richness of the message. The abbreviation cuz (because) additionally marks the informality and colloquial nature of the statement.

The parenthetical insertion (*thank god I did cuz otherwise I'd be jumping off a cliff*) serves a dual pragmatic function. On the one hand, it reduces the pathos of the main statement through self-irony; on the other, it enhances the significance of the achievement by implying that the alternative outcome would be emotionally catastrophic. This technique corresponds to Twitter's typical strategy of balancing self-praise and self-deprecation. Discursively, this example illustrates the use of *slay* as a marker of youth identity and speaks for the people. The lexeme comes from African-American Vernacular English (AAVE) and pop-cultural discourse, where it has acquired the meaning of enthusiastic appreciation. Its use without explanation indicates that the word has already become part of the shared lexicon of online communities and does not require glossing.

#### 3.3.4. Era: Semantic Shift Through Construction.

*Era* is an example of a less obvious but theoretically significant semantic shift: the lexical item itself has not changed (with *era* in the sense of *a period of time* remaining fully conventional), but in the expression *in my X era* it has acquired a specific constructive semantics. The construction in *my* + [noun or adjective] + *era* denotes *a time-limited phase of personal experience characterized by a particular quality, identity, or behavioral orientation*. *In my healing era* refers to the ongoing process of psychological recovery; *in my villain era* denotes a phase in which the speaker prioritizes self-interest over social approval; *in my hot girl era* refers to a period characterized by heightened self-confidence and self-presentation.

A corpus analysis of Twitter/X posts conducted using Voyant Tools confirms the high prevalence of this structure in online discourse among young people. The vocabulary density score stands at 0.459, indicating moderate lexical diversity alongside a high degree of repetition of key patterns. The readability index (4.741) is relatively low, reflecting the tendency of such posts towards a style that is as conversational and accessible as possible. At the same time, the average sentence length – 56.2 words, demonstrates Twitter/X users' tendency towards syntactically elaborate, emotionally charged utterances in which personal experience is combined with reflection and self-description. The most frequent unit in the corpus is *era* (33 occurrences), whilst the accompanying vocabulary – *girl* (8), *love* (4), *want* (3) and *soft* (3), indicates a close connection between the construction and discourses of self-care, emotional authenticity and digital self-presentation.

*About to be in my healing era (traveling back up north this week to visit my friends)* (Twitter/X, @ghostbunnytat2s, 4:12 AM · Dec 27, 2023)

The key to the analysis is the construction *in my healing era*, which represents a new type of nomination of personal experience. The analysis should

begin at the grammatical level. The construction *in my X era* is a prepositional group, where *era* acts as a main noun, modified by a possessive pronoun (*my*) and a prepositional definition (*healing*). Formally, it does not differ from standard noun groups, but its frequency and patterning indicate the formation of a new productive model. The expression *about to be* adds a future perspective, transforming the state into an intention or an imminent change in identity.

At the semantic level, *era* undergoes a noticeable semantic broadening. Traditionally, the word denotes a long historical period, but in this context it is used to describe an individual, often short-lived phase of life. Thus, *era* functions as a label for a complex state that includes behavioral practices, emotional attitudes, and lifestyle. Combined with *healing*, it forms the concept of self-healing, which is recognizable in modern digital discourse.

Pragmatically, the construction performs a declarative function: the speaker not only describes a state, but also proclaims a new behavioral strategy. The specification in brackets (*traveling back up north...*) serves to concretize this *era*, translating the abstract concept into the plane of real actions. Such a structure (general declaration + specific example) is typical of Twitter discourse, where personal experience is presented as part of a broader, recognizable narrative.

From a discursive perspective, *in my X era* functions as a tool for self-presentation and identification. The utterance appeals to a collectively understood template that allows the audience to quickly interpret the speaker's state without further explanation. This corresponds to the concept of ambient affiliation (Zappavigna, 2012), where the connection between users is established through shared discursive formulas and cultural codes.

### 3.4. Pragmatic formulas: conventional units of pragmatic identification.

#### 3.4.1. Theoretical characterization of memetic expressions.

Memetic expressions occupy a distinct place within the system of lexical innovations on Twitter/X. Unlike neologisms and semantic shifts, which primarily affect the denotative and semantic dimensions of a word, memetic expressions are predominantly pragmatic innovations. Their meaning is determined not by the lexical content of their components, but by the convention of usage within a specific discursive space. Within Shifman's framework (Shifman, 2014), such expressions may be understood as linguistic templates: relatively stable communicative patterns characterized by formal recognizability and predictable pragmatic function, while remaining open to contextual variation in lexical realization. The broader scholarly understanding of the meme as a cultural phenomenon originated in *The Selfish Gene*, where Richard Dawkins introduced the term to describe a unit of cultural transmission that spreads through imitation, in a way analogous to the replication of genes in biological evolution. In this conceptualization, memes function as replicable cultural units capable of caritative, selection, and dissemination through social interaction. This principle of replication through imitation is particularly

relevant for understanding the circulation of linguistic patterns in digital communication.

Shifman (2014) reinterpreted this concept in relation to the digital environment and demonstrated that internet memes should be viewed not simply as humorous artifacts, but as structured communicative forms with their own patterns of replication and mutation. Memetic expressions on Twitter/X may be regarded as a specific linguistic manifestation of this broader phenomenon. They spread through mass imitation, where each new instance simultaneously reproduces an established pattern and introduces contextual modification. In this process, expressions compete for discursive visibility, and those that prove sufficiently recognizable, adaptable, and pragmatically efficient tend to stabilize within collective communicative practice, while less effective forms gradually disappear. In this respect, memetic expressions constitute both linguistic and cultural phenomena, as their dissemination follows principles comparable to those underlying the spread of other cultural innovations.

From a grammatical point of view, memetic expressions often preserve the formal structure of complete syntactic constructions, including full clauses or phrasal combinations with clear subject-predicate organization. However, pragmatically they function as highly compressed communicative units, whose role is closer to discourse markers or evaluative signals than to fully propositional statements. This is an example of pragmatic grammaticalisation in the terms of Hopper and Traugott (Hopper & Traugott, 2003), according to which formally complex constructions may, through repeated contextual use, develop reduced and specialized discourse-pragmatic functions. An important feature of memetic expressions lies in their dual communicative function. On the one hand, they convey specific evaluative meanings, such as approval, irony, solidarity, or stance-taking. On the other hand, they simultaneously serve as markers of group affiliation, signaling the speaker's participation in a communicative community that recognizes and appropriately interprets these conventionalized forms. It is precisely this combination of semantic expressiveness and social indexing that accounts for their durability and productivity within digital discourse.

#### 3.4.2. Understood the assignment and no notes.

Pragmatic formulas display considerable contextual adaptability and rely heavily on intertextual references and shared digital culture. One such unit is the phrase *understood the assignment*, which is used in contemporary online discourse to express approval, recognition or admiration for someone's fulfilment of a particular social, aesthetic or communicative role. The phrase is significant from the perspective of the pragmatics of digital discourse, as it combines evaluative, intertextual and group solidarity mechanisms. The corpus consists of a single document containing 516 word occurrences and 258 unique word forms. The vocabulary density is 0.500, indicating moderate lexical diversity given the high frequency of key meme constructions. The readability

index (12.234) and average sentence length (25.8 words) point to a syntactically more complex type of Twitter/X communication compared to the platform's typical short replies. The most frequent items in the corpus are *understood* (30 occurrences) and *assignment* (30 occurrences), confirming the formulaic nature of the construction *understood the assignment* and its status as an established unit of digital discourse. The frequency of lexemes such as *like*, *birthday* and *album* also demonstrates the wide thematic variety of contexts in which this formula is used – ranging from pop culture and fan discourse to everyday online communication.

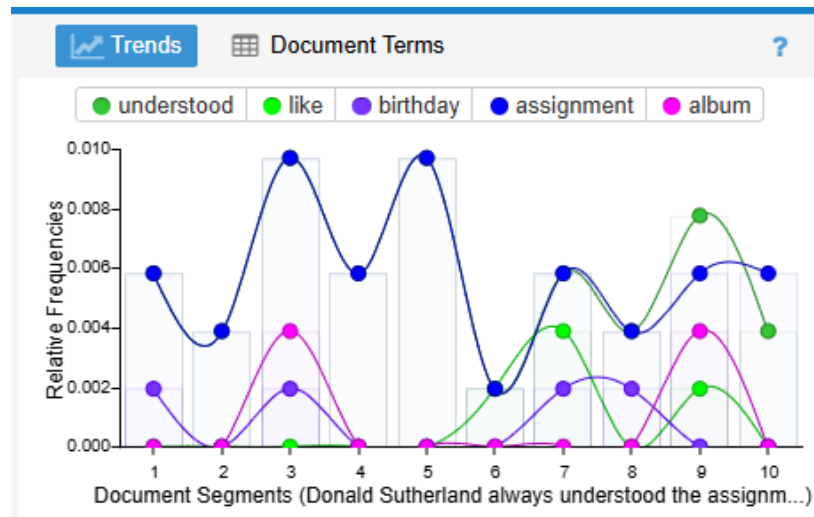


Figure 3.7. Voyant Tools Trends visualization showing frequency distribution.

*Finland, I love everything about your performance. No notes. You clearly understood the assignment! #SBSEurovision* (Twitter/X, @goodonestu, 2:08 PM · May 14, 2023)

The key to the analysis are the memes *no notes* and *understood the assignment*, which function as established means of positive evaluation.

The analysis should begin with the pragmatic-syntactic level. The statement consists of short, segmented sentences (I love everything... / No notes. / You clearly understood the assignment!), which is typical of Twitter discourse. Such segmentation enhances expressiveness and creates the effect of oral speech, where each fragment performs a separate rhetorical function. The address Finland personifies the object of evaluation (it is about the country's performance), which is a common technique in fan discourse.

The expression *no notes* is a metaphor borrowed from the editorial or professional environment, where notes means remarks or corrections. In this context, its use means the complete absence of criticism - that is, unconditional approval. Grammatically, it is an elliptical construction (full form: I have no notes), which emphasizes the brevity and formality of the statement. It is such a reduction that is characteristic of meme units that strive for maximum economy of linguistic means.

The construction understood the assignment is also metaphorical and comes from educational or work discourse, where it means the correct understanding of the task set. In the digital context, it undergoes a semantic shift: it is not about the literal task, but about the ability to adequately interpret the situation or the expectations of the audience and exceed them. The adverb clearly performs an intensifying function, reinforcing the categorical nature of the assessment.

The hashtag *#SBSEurovision* performs both a search and a discursive function: it embeds the tweet in the broader context of discussion of a specific event and signals belonging to a fan community. This corresponds to the concept of ambient affiliation (Zappavigna, 2012), where the connection between users is established through shared thematic markers.

It is important that both formulas (no notes, understood the assignment) are used without any explanation or detail. The subject of the assessment (a specific performance) is not named directly, but is easily reconstructed through the context. Such situational implicitness is a key characteristic of meme speech: meaning is formed not only by linguistic means, but also by the shared knowledge of the audience.

#### 3.4.3. No cap, IYKYK: grammaticalized abbreviations as meme units.

A separate subgroup of meme formulas is made up of grammaticalized abbreviations – abbreviations and reductions that have turned into full-fledged discursive markers. *No cap* is an abbreviation of *no cap on that* (literally without a lid, i.e. without deception): in modern usage it means *it's true, seriously, no joke*. Functionally, *no cap* is a discursive marker of confirmation of the authenticity of a statement – an analogue of *fr* (for real), *deadass* or *I'm serious*. It is important that *no cap* is used as an independent unit (in the function of confirmation at the end of a sentence) or as an interjection/phrase.

*No cap I really only talk to my parents and like 3 other ppl regularly and I've never been at peace like this b4* (Twitter/X, @KhorWal16, 12:18 AM · Dec 29, 2023)

*No cap* is used in a preposition – at the beginning of a statement, where it functions as a global modifier of the entire sentence. Its meaning (without exaggeration, honestly, truth) signals that the speaker emphasizes the sincerity and non-constructed nature of the message. Such a position is typical: moving the marker to the beginning creates a framework for interpretation for the entire subsequent statement.

At the grammatical level, *no cap* is not integrated into the syntactic structure of the sentence, but functions as an interstitial element (discourse marker), which does not affect the propositional content, but significantly modifies its pragmatic perception. Similar units are characteristic of colloquial speech and digital communication, where they serve as signals of the speaker's position.

The general syntactic organization of the utterance is also indicative: the lack of clear punctuation, the use of colloquial abbreviations (ppl, b4), and constructions such as *like 3 other ppl* as a marker of approximation. All this indicates an orientation towards oral speech and confirms the tendency towards syntactic informalization described by Baron (2008).

At the pragmatic level, the combination of no cap with the intensifier really (I really only talk to...) is important. Both elements perform the function of enhancing credibility, which creates the effect of double marking of sincerity. Such excessive explication is indicative of Twitter discourse, where irony, hyperbole, and performativity are so common that the authenticity of the statement often requires additional signals. Semantically, the statement represents a minimalist social interaction (only talk to my parents and like 3 other ppl), which, however, is evaluated positively (I've never been at peace like this). This creates a counterintuitive effect: what could be interpreted as social isolation is presented as a source of inner peace. This is where no cap plays a key role – it guarantees the sincerity of this non-trivial assessment.

Discursively, no cap is a marker of belonging to youth and online slang (in particular, AAVE-origin), and its use signals the speaker's competence in the relevant communicative environment. At the same time, the absence of any explanation indicates that the expression is already sufficiently established and understandable for the target audience.

IYKYK (if you know you know) is the functional opposite of no cap: if the latter opens the meaning of the expression to everyone, then IYKYK deliberately limits it to the initiates. The microcorpus under investigation consists of a single document containing 423 word occurrences and 256 unique word forms. The vocabulary density is 0.605, indicating a relatively high lexical diversity within the corpus, despite its small size. The readability index (8.547) and average sentence length (18.4 words) reflect the combination of conversational spontaneity and syntactic compression characteristic of Twitter/X. The most frequent items in the corpus are iykyk (26 occurrences), just (3), got (3), you're (2) and year (2). The dominance of the abbreviation iykyk confirms its central role as a pragmatic marker of shared experience and group identification in digital discourse.

It is a tool of esoteric solidarization: the speaker makes it clear that there is a certain common knowledge that connects him with a part of the audience and excludes the rest.

*this weekend has been so filler. It has felt like ...episodes slowly leading up to something big. Idk how to describe it but iykyk* (Twitter/X, @bowsnbubbles, 12:07 AM · Dec 29, 2025)

The analysis should begin at the semantic-grammatical level. The word filler, which traditionally functions as a noun (placeholder, secondary content), is in this case used predicatively after a verb-link (has been so filler). Such a conversion (noun to adjective function) indicates the flexibility of the word and

its integration into spoken online discourse. Semantically, filler comes from media discourse (in particular, television series, where filler episodes are episodes that do not advance the main plot), and here it is transferred to denote life experience. Thus, a metaphorical rethinking takes place: personal time is conceptualized as a narrative structure.

The further development of the metaphor (episodes slowly leading up to something big) reinforces this understanding: the weekend is interpreted as part of a series that leads to a future climax. The ellipsis (...episodes) reproduces the speaker's pause and hesitation, imitating the process of finding an adequate description, which enhances the effect of spontaneity and authenticity. The abbreviation *iykyk* (if you know, you know) functions as a key pragmatic marker. It does not simply shorten the expression, but performs the rhetorical function of appealing to shared experience. The speaker acknowledges the difficulties of verbalization (Idk how to describe it), but at the same time delegates part of the interpretation to the audience: understanding depends on the presence of similar experience. Thus, *iykyk* creates an implicit boundary between those who understand and those who do not.

From a pragmatic point of view, this is a strategy of solidarity through partial understatement. The expression simultaneously performs two functions: inclusive (it creates a sense of community among those who are *in the know*) and exclusive one (it leaves out those who do not have the relevant experience or cultural knowledge). This mechanism corresponds to the phenomenon of ambient affiliation described in digital discourse studies (Zappavigna, 2012), where the social bonds between users are established through commonly recognized, but not always explicit, meanings.

### 3.5. Syntactic and punctuation levels: compression and pragmatic drift.

#### 3.5.1. General characteristics of syntactic changes.

Analysis of the syntactic level of the sample revealed two interrelated phenomena: a general trend towards shortening sentences and simplifying grammatical structure, and the emergence of new conventional constructions – meme templates with specific syntactic characteristics. The Summary function in Voyant Tools confirmed that the average sentence length in the 2023–2025 material is lower than in the 2019–2020 material. The proportion of sentences without a formally expressed subject or predicate is increasing. These observations are consistent with the findings of Naomi Baron (Baron, 2008, p. 74), who recorded similar trends in text messages in the early 2000s and predicted their expansion as mobile and social network communication grew.

Syntactic compression in Twitter discourse is not simply a consequence of character limitations: it is a cultural norm of the platform. Long, “well-written” sentences are not associated with successful content on Twitter; brevity and punch are markers of effective communication. This norm has been shaped by algorithmic reward (Bucher, 2018): concise, resonant posts receive more likes and retweets, i.e. are algorithmically incentivized.

### 3.5.2. Pragmatic drift of the final stop.

As Geoffrey Nunberg argues in *The Linguistics of Punctuation* (1990), punctuation marks may extend beyond their purely grammatical function and acquire additional discourse-pragmatic meanings in specific communicative contexts. In digital discourse, the full stop increasingly functions not only as a marker of syntactic completion but also as a potential indicator of stance, emotional distance, seriousness, or irritation. This development is closely related to the communicative conventions of Twitter/X discourse, where short messages frequently appear without terminal punctuation. Since a post already constitutes a complete communicative unit, the omission of the full stop becomes pragmatically unmarked, whereas its explicit use acquires additional interactional significance. In this sense, the full stop undergoes a process of pragmatic enrichment through repeated contextualized use.

Jeff Shayble (2015) argues that punctuation in digital communication is increasingly taking on culturally and interactionally determined meanings that go beyond its traditional grammatical functions. Through repeated use in such online contexts, punctuation marks develop relatively stable pragmatic associations within digital discourse communities. This semantic-pragmatic expansion can be compared to what Traugott and Dasher (2002) describe as *pragmaticalization*, namely the process through which linguistic forms acquire additional pragmatic functions alongside their existing grammatical meanings.

Variant	Example	Pragmatic meaning
No full stop, lowercase letters	i love this song	warm, open, informal
With a full stop, capital letters	I love this song.	cold, distant, or ironic
ok without full stop	ok	confirmation, neutral consent
ok with full stop	ok.	irritability, reluctance to speak
sure without full stop	sure	neutral or benevolent consent
sure with full stop	sure.	skepticism, passive aggression
Three sentences with full stops.	That's fine. Do whatever. It's your life.	emphasized coldness, resentment

Table 3.8. The pragmatic effect of the full stop: minimal pairs and contexts.

*THREE DUPES BACK TO BACK!! How is it that I've been pulling only dupes the past 3 days???! I need to calm down. Ok bye.* (Twitter/X, @Livheartt, 25:18 PM · Dec 27, 2025)

At the paragrammatic level, the statement is characterized by high expressiveness: the use of capital letters (THREE DUPES BACK TO BACK!!), multiple exclamation and question marks (???!), serves to intensify the emotional state – surprise, frustration, even mild affect. Such graphic excess is typical of Twitter discourse, where punctuation partially compensates for the lack of prosodic means of spoken interaction.

Against this background, the final *Ok bye.* contrasts sharply with the previous segments. Grammatically, it is a simple, elliptical construction, but it is the full stop at the end that performs a crucial pragmatic function. Without the full stop (*Ok bye*), the statement would be perceived as a neutral ending or even an amicable exit from the conversation. By contrast, the presence of the full stop adds a touch of completion, abruptness, and a certain emotional distance: the speaker does not simply leave, but demonstratively closes the communicative act.

Thus, the full stop here functions as a marker of micro-pragmatic meaning, signaling a shift in tone from emotional overload to sharp self-soothing or even self-ironic disengagement. In combination with the preceding *I need to calm down*, this creates the effect of performative self-control: the speaker publicly records his own emotional state and at the same time stages its completion. ChatGPT consistently classified the examples *ok.* as *negatively colored or ironic*, and this classification coincided with our own analysis every time. The number of examples of *ok.* in the sample is sufficient to conclude that this usage is conventional: it is not an individual stylistic whim, but an established norm of the platform.

3.6. Chronological dynamics: the relationship of linguistic innovations with the sociocultural context.

The diachronic dimension of the analysis is important not only to confirm the stability of the identified trends, but also to understand their nature: why do certain linguistic forms arise exactly when they do? The *Trends* function in Voyant Tools allowed us to identify several types of chronological patterns in the sample.

The first type is explosive pattern (rapid rise): a unit is practically absent until a certain point, then grows sharply and subsequently stabilizes. *Doomscrolling* is a classic example. The *Trends* function recorded a similar profile for quarantine fatigue, social distancing in a figurative sense, infodemic – all of them are associated with the pandemic discourse of 2020. This confirms the theoretical observation: certain external events function as linguistic

catalysts, accelerating the emergence and spread of the corresponding lexical units.

The second type is the gradual rise pattern: the unit is present in the dataset from the beginning of the study period, but slowly increases in frequency. *Slay* in the new meaning corresponds to this profile: it is present already in 2019 in the subcultural environment, but noticeably increases from 2021–2022 as it goes beyond the LGBTQ+ communities into the platform mainstream. *Era* in the construction *in my X era* also shows a gradual increase starting from 2021. The third type is stable presence: the unit is present throughout the study period with a relatively flat profile. Syntactic compression and pragmatic drift of the full stop exemplify this pattern: they do not exhibit a clearly identifiable point of emergence within the dataset, as their development evidently predates the period under analysis. The stable presence in 2019–2025 indicates that we are talking about long-term structural changes, not temporary surges.

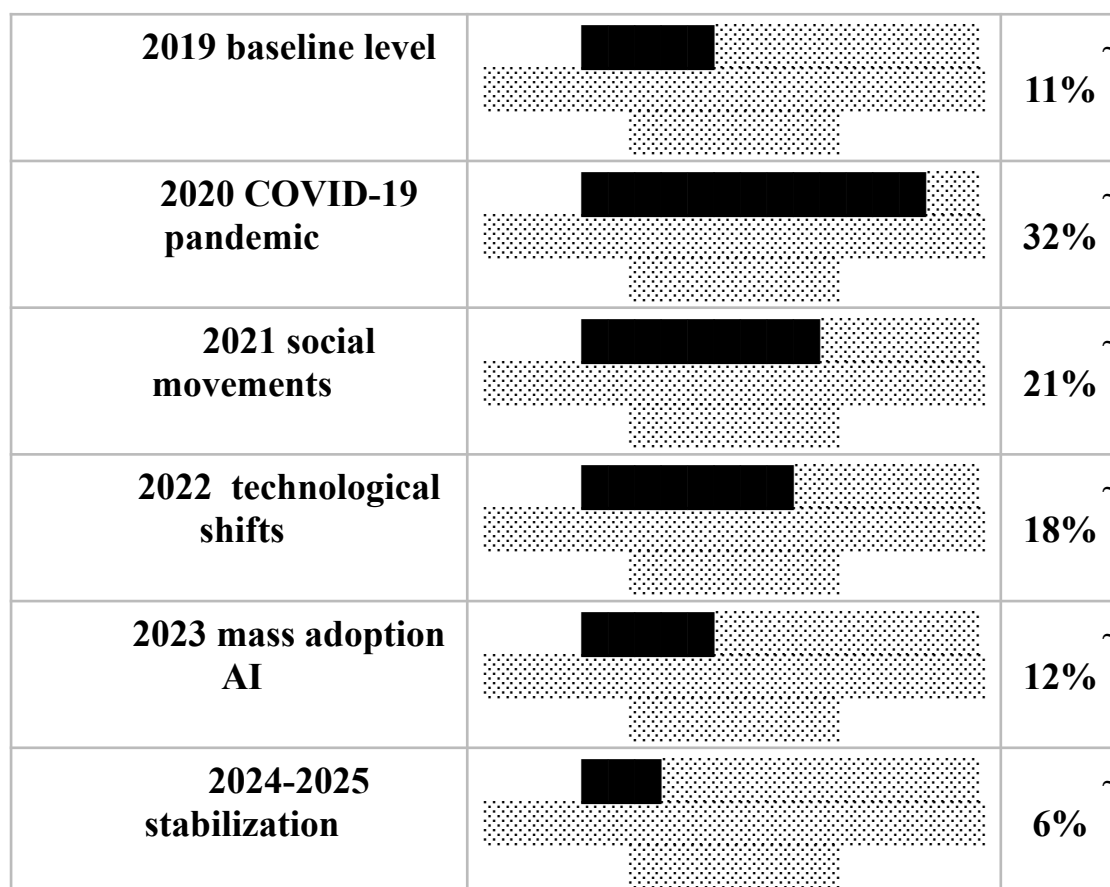


Figure 3.9. Approximate distribution of new lexical items by year of first fixation in the sample.

Figure 3.9 presents the approximate distribution of new lexical items by year of their first significant attestation in the dataset. The dominance of 2020 (about 32%) is statistically evident and theoretically explainable: the COVID-19 pandemic has become an unprecedented linguistic catalyst, which in a short time

has generated or activated a large number of new forms. At the same time, 2019 is not zero baseline: it contains a base layer of platform-specific innovations that existed even before the pandemic (e.g. *no cap*, *vibe* in its extended meaning, *goblincore* in its early stage).

The decrease in the frequency of new units in 2023–2024 (to 12% and 6%, respectively) does not indicate the exhaustion of the platform’s linguistic innovativeness: it is an artifact of the sample collection method (the 2024 material is a partial cross-section and does not cover the full annual cycle) and a general tendency to stabilize the linguistic profile after the explosive 2020–2021. At the same time, the new context – the mass distribution of AI tools and the related public discourse – forms its own lexical layer: *hallucinate* (in the sense of producing fabricated information in relation to AI systems), *prompt engineering*, *AI slop* are all of them are present in the 2023–2024 subset of the data.

### 3.7. Summary results: systematization and confirmation of the hypothesis.

The results of the analysis described in subsections 3.1–3.6 make it possible to systematize the identified trends in the form of a summary table and formulate a final conclusion regarding the working hypothesis of the study.

Level	Trend	Key items in the dataset	Theoretical framework	Verification method
Lexical	Neologisms	<i>doomscrolling</i> , <i>enshittification</i> , <i>rizz</i> , <i>goblincore</i>	Rogers (Diffusion of Innovations); Bauer (word-formation and lexical innovation)	Voyant Trends; example analysis
Lexical	Semantic shifts	ratio, slay, cancel, era, vibe	Ullmann (semantic change typology)	Voyant Contexts + contextual interpretation
Lexical	Memetic expressions	understood the assignment, no notes, no cap, IYKYK	Shifman (meme theory)	example analysis
Pragmatic	Punctuation drift	ok., sure., ellipsis chains	Nunberg (1990) punctuation as discourse/prosodic system	Minimal pair analysis; ChatGPT-assist

				ed pragmatic comparison
Diachronic	Sociocultural conditioning	2020 pandemic cluster; 2023 AI cluster	Labov (sociolinguistic change); Rogers (S-curve diffusion model)	Voyant Trends; temporal clustering analysis

Table 3.10. Summary of research results.

Table 3.10 provides a systematic overview of the identified trends, specifying the corresponding theoretical frameworks and methods of verification. All five rows of the table correspond to phenomena recorded in at least three independent examples of the sample, which is the minimum necessary criterion for concluding that the phenomenon is conventional (rather than individual). Based on the presented results, it can be stated that the working hypothesis of the study is confirmed: Twitter/X and more broadly, social media platforms constitute an independent linguistic environment with its own mechanisms of language creation and dissemination of innovations, and the revealed patterns are systemic. They are reproduced in different thematic categories of the sample (from humor to social events), in different time periods (2019–2025) and at different levels of linguistic organization (vocabulary, syntax, punctuation). These trends are driven by both the technological features of the platform (character limits, algorithmic visibility, public-private nature of communication) and the sociocultural context (pandemic, social movements, technological change).

The section presented the detailed results of the analysis of a sample of English-language Twitter/X posts for 2019–2025. At the lexical level, three types of innovations were identified: *de novo* neologisms (with a detailed case study of *doomscrolling*, *enshittification*, *rizz*, *goblincore*), semantic shifts (*ratio*, *slay*, *era* with an analysis of mechanisms and examples), meme formulas (*understood the assignment*, *no cap*, *IYKYK* with an analysis of pragmatic functions). At the pragmatic level, the drift of the full stop from a neutral grammatical sign to a marker of sharpness and passive aggression is documented. Diachronic analysis revealed a direct connection between sociocultural events (pandemic, AI spread) and the emergence of corresponding lexical clusters. All identified phenomena are theoretically interpreted within the framework of the concepts of the scientific literature mentioned in Section 1, which ensures the theoretical coherence of the work.

#### Conclusions and discussion

The Master's paper is devoted to the study of innovative and transformational trends in the language of modern social networks based on the material of English-language Twitter/X posts for 2019–2025. The conclusions presented below summarize the results of all chapters of the work, justify the

confirmation of the working hypothesis, and outline practical implications as well as directions for further research. The study identified and described two systemic transformational trends that characterize the language of modern social networks.

The first trend is lexical. The sample material recorded active neologization and semantic shifts occurring at an unprecedented speed. Neologisms emerged as a response to new social situations for which there was no precise designation in the language. The *Trends* function in Voyant Tools recorded a characteristic S-shaped dynamics of distribution: practical absence up to a certain point, rapid growth and subsequent stabilization. This dynamics exactly corresponds to the Everett Rogers' diffusion model (Rogers, 2003) and confirms that the mechanism of distribution of linguistic innovations in social networks is fundamentally the same as in traditional language, but significantly accelerated. In addition to neologisms, semantic shifts of existing words were recorded. *Ratio* turned from a mathematical term into a platform-specific designation of a social mechanism for evaluating posts; *slay* made an ameliorative shift from *to kill* to *to perform exceptionally well*; *era* in the construction *in my villain era* acquired a new meaning of personal psychological mode or phase. All these shifts were verified through the *Contexts* function in Voyant Tools, which showed a fundamentally different lexical environment of the words in the sample compared to standard dictionary meanings, and were also confirmed through ChatGPT analysis and checking in Merriam-Webster and Oxford English Dictionary. A separate subgroup is made up of meme formulas – expressions that are grammatically complete sentences (understood the assignment, no notes, no cap), but pragmatically function as conventional markers of approval, authenticity, or group identity. In the terms of Limor Shifman (2014), they are linguistic templates with a fixed structure and variable content.

The second trend is syntactic-pragmatic. Analysis of the syntactic level recorded a general reduction in the average sentence length and an increase in the proportion of constructions without a formally expressed subject or predicate. According to the concept of Zappavigna (2012), such constructions are *ambient affiliation* that appeal to the common experience of the audience and invite solidarity. The pragmatic drift of the full stop is one of the most important results of the study. The full stop – a traditionally neutral grammatical sign has acquired in the language of social networks the additional meaning of sharpness, closure or passive aggression. Analysis of minimal pairs confirmed that identical verbal content is perceived fundamentally differently depending on the presence or absence of a full stop. ChatGPT consistently classified examples with a dot as negatively colored or ironic. According to Nunberg's analysis (Nunberg, 1990), punctuation can be understood as a written system that partially encodes prosodic and discourse-structuring information found in spoken language. In this sense, in the absence of phonological features such as intonation, pragmatic and

discourse-related meanings can be partially realised through paragraphematic features, including punctuation marks.

The diachronic analysis adds an important dimension: most of the identified lexical innovations have a clear point of emergence, which correlates with specific extralinguistic events. The COVID-19 pandemic (2020) became a linguistic catalyst, which gave rise to or activated a whole cluster of new forms in a figurative sense. The social movements of 2020–2021 activated cancel culture. The spread of generative artificial intelligence in 2022–2023 introduced *enshittification* and related forms. Syntactic and punctuation changes, on the other hand, are more gradual and do not have a clear point of origin, which indicates their long-term, structural nature.

The working hypothesis of the study stated: social networks are an independent linguistic environment with their own mechanisms of language creation and dissemination of innovations, and the revealed patterns are systemic and determined by both the technological features of the platforms and the socio-cultural context. The results of the analysis fully confirmed this hypothesis. Reason one: the technological environment of the platform. The 280-character limit is the most obvious structural factor that directly determines linguistic practices: it stimulates grammatical compression, semantic brevity and gives preference to apt formulas over detailed explanations. At the same time, algorithmic visibility is no less important: posts that generate more reactions receive a wider reach. Language forms that are resonant, apt or humorous are rewarded by dissemination, which is a mechanism for selecting successful linguistic innovations. As T.Bucher and A.Helmond (2018) noted, this process is algorithmic accommodation: users unconsciously optimize their language to the requirements of the platform.

Second, the public-private nature of communication. Users write in a context that is simultaneously personal and publicly accessible, which fosters strategies of ironic self-distancing as a form of protection. This explains the prevalence of ironic self-deprecation (as in the examples with *doomscrolling*), passive aggression through a period and humorous meme templates are the dominant pragmatic strategies in the sample. The concept of the imagined audience by A.Marwick and D.Boyd (Marwick & Boyd, 2011) accurately describes this phenomenon: speakers construct a certain reader in their imagination and adjust their language to him, although the real audience may differ significantly. Third, the memetic mode of linguistic diffusion. Successful expressions are replicated, adapted, copied, varied, reinterpreted, eventually stabilizing as conventional units. As demonstrated by Shifman (2014), memes combine efficiency and variability. It is the combination of these qualities that ensures their virality and rapid consolidation as a linguistic norm.

Fourth, the sociocultural context of 2019–2025. The COVID-19 pandemic, mass social movements, and the rapid development of artificial intelligence – each of these events created new social realities that required new

linguistic means to describe. Twitter language reacted quickly and systematically: new words and new meanings emerged in response to new situations.

Beyond its empirical findings, the study makes theoretical contributions that go beyond the scope of the described material. The work theoretically substantiates and specifies the concept of platform reinterpretation as a separate subtype of semantic shift specific to the conditions of digital communication. Unlike classical semantic changes according to Stephen Ullmann (expansion, narrowing, melioration, pejoration), platform reinterpretation is a shift that is relevant and understandable exclusively in the context of a specific platform.

The results on pragmatic punctuation drift offer a new perspective on a more general phenomenon of replacement of prosodic features of spoken language with paragraphematic devices in written digital discourse, suggesting the emergence of a parallel punctuation grammar.

At the same time, several limitations must be acknowledged. The first limitation concerns the nature of the sample. The study relies on a purposive, not a continuous, sample. This is an acceptable approach for qualitative linguistic research, but does not ensure statistical representativeness in a strictly quantitative sense: the conclusions drawn are grounded qualitative observations supported by Voyant Tools frequency data, rather than statistically verified statements.

The second limitation is instrumental: Voyant Tools is a means of exploratory analysis and does not automatically compare the material with a reference corpus of standard language. Conclusions about the specificity of certain forms for social networks are based on a combination of quantitative observations and qualitative analysis, checked against dictionaries and linguistic literature. The third limitation concerns the use of ChatGPT. Although the model was used exclusively as an auxiliary analytical tool with mandatory verification of each answer, it is not a neutral observer and may have its own propensities when working with subcultural or modern expressions. No conclusion of the work is based solely on ChatGPT assessments.

The identified trends are valuable material for courses in pragmatics, sociolinguistics, media linguistics and lexicology of the English language. *Semantic shift ratio, S-curve of doomscrolling, minimal ok/ok pairs.* are visual illustrations of the theories of Ullmann, Rogers and Nunberg. Understanding the language of social networks is an integral part of communicative competence in modern English. The results of the study suggest a systematic approach to teaching the digital layer of language: not as a chaotic set of fashionable expressions, but as a system with its own logic and rules. The identified features of the Twitter/X language are important for the development of NLP systems, in particular, tone analysis and irony recognition. The results of the study can serve as the basis for formulating requirements for annotated training data for such systems.

This work considers Twitter/X trends as representative of social networks in general, but this is a hypothesis that remains to be tested. To what extent are the identified phenomena Twitter and to what extent are they general digital? Is there a pragmatic drift of the dot on Instagram or WhatsApp? Are the constructions as frequent in Reddit or TikTok? A comparative study would answer these questions and significantly clarify the overall picture. This work analyzes text production, but no less important is the question of perception. Do native speakers really interpret ok. as negative and ok as neutral? Are there age or social differences in the perception of meme formulas? A psycholinguistic experiment investigating the assessment of the tonality of minimal pairs by participants of different demographic groups would provide answers to these questions. Research with a large balanced corpus, reference corpora of standard language, and statistical measures of association would allow us to quantitatively verify the qualitative observations of this work and to identify new trends that are not noticeable with a small sample.

Since 2022, Twitter has undergone significant structural changes (rebranding to X, algorithm changes, monetization of verification), which are already affecting the composition and behavior of the audience. Are language practices changing in response to these changes? Is there a fragmentation of language norms as certain groups leave for Threads, Bluesky? These questions require urgent research while the processes are still ongoing.

The language of Twitter/X is not a corrupted standard English or a chaotic set of deviations from the norm. It is an independent, internally coherent system with its own norms, genres, development mechanisms, and its own means of conveying meaning. It is a mirror of its time. The language of social networks is extremely sensitive to social context – more sensitive than any previous form of written communication. It reacts to external events rapidly and systematically, producing new means to describe new realities. That is why social networks are one of the most important and interesting objects of linguistic research today.

#### References

1. Algeo, J., & Pyles, T. (2010). *The origins and development of the English language*. Wadsworth Cengage Learning.
2. Androutsopoulos, J. (2013). *Computer-mediated Communication and Linguistic Landscapes*. *Research methods in sociolinguistics: A practical guide*, 74-90.
3. Baker, P. (2023). *Using corpora in discourse analysis*. Bloomsbury Publishing.
4. Baron, N. S. (2010). *Always on: Language in an online and mobile world*. Oxford University Press.
5. Bauer, L. (2001). *Morphological productivity* (Vol. 95). Cambridge University Press.

6. Blommaert, J., & Varis, P. (2015). Conviviality and collectives on social media: Virality, memes, and new social structures. *Multilingual Margins: A journal of multilingualism from the periphery*, 2(1), 31-31.
7. Bucher, T. (2018). *If... then: Algorithmic power and politics*. Oxford University Press.
8. Burdick, L., Mihalcea, R., Boyd, R. L., & Pennebaker, J. W. (2021). Analyzing connections between user attributes, images, and text. *Cognitive Computation*, 13(2), 241-260.
9. Crystal, D. (2001). *A dictionary of language*. University of Chicago Press.
10. Crystal, D. (2004). *Language and the Internet*.
11. Crystal, D. (2011). *Internet linguistics: A student guide*. Routledge.
12. Dawkins, R. (1981). In defence of selfish genes. *Philosophy*, 56(218), 556-573.
13. Doctorow, C. (2023). Tiktok's enshittification. *Pluralistic: Daily links from cory doctorow*.
14. Dunning, T. E. (1993). Accurate methods for the statistics of surprise and coincidence. *Computational linguistics*, 19(1), 61-74.
15. Eisenstein, J. (2013, June). What to do about bad language on the internet. In *Proceedings of the 2013 conference of the North American Chapter of the association for computational linguistics: Human language technologies* (pp. 359-369).
16. Eisenstein, J., O'Connor, B., Smith, N. A., & Xing, E. P. (2014). Diffusion of lexical change in social media. *PloS one*, 9(11), e113114.
17. Fairclough, N. (1992). Discourse and text: Linguistic and intertextual analysis within discourse analysis. *Discourse & society*, 3(2), 193-217.
18. Fairclough, N. (1992). Intertextuality in critical discourse analysis. *Linguistics and education*, 4, 269-93.
19. Gilardi, F., Alizadeh, M., & Kubli, M. (2023). ChatGPT outperforms crowd workers for text-annotation tasks. *Proceedings of the National Academy of Sciences*, 120(30), e2305016120.
20. Green, J. (2016). *Slang: A very short introduction* (Vol. 465). Oxford University Press.
21. Grice, H. P. (1975). Logic and conversation. In *Speech acts* (pp. 41-58). Brill.
22. Grieve, J., Nini, A., & Guo, D. (2017). Analyzing lexical emergence in Modern American English online1. *English Language & Linguistics*, 21(1), 99-127.
23. Hamilton, W. L., Leskovec, J., & Jurafsky, D. (2016, August). Diachronic word embeddings reveal statistical laws of semantic change. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 1489-1501).

24. Herring, S. C. (1996). Computer-mediated communication.
25. Herring, S. (2007). A faceted classification scheme for computer-mediated discourse. *Language@ internet*, 4.
26. Hofmann, M. (2018). Habermas's Public Sphere versus Trump's Twittersphere: Citizenship in a World of Social Media. *The Journal of Communication and Media Studies*, 3(4), 1.
27. Hopper, P. J., & Traugott, E. C. (2003). *Grammaticalization*. Cambridge university press.
28. Kress, G., & Van Leeuwen, T. (2020). *Reading images: The grammar of visual design*. Routledge.
29. Labov, W. (1994). *Principles of linguistic change (Vol. 1)*. Oxford: Blackwell.
30. Labov, W. (2004). Ordinary events. *Sociolinguistic variation: Critical reflections*, 31-43.
31. Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics*, 159-174.
32. Marwick, A. E., & Boyd, D. (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New media & society*, 13(1), 114-133.
33. Maynard, D. G., & Greenwood, M. A. (2014, March). Who cares about sarcastic tweets? investigating the impact of sarcasm on sentiment analysis. In *Lrec 2014 proceedings*. ELRA.
34. McEnery, T., & Hardie, A. (2011). *Corpus linguistics: Method, theory and practice*. Cambridge University Press.
35. Merriam-Webster. (2020). Doomscrolling. In Merriam-Webster online dictionary. Retrieved December 1, 2024, from <https://www.merriam-webster.com/dictionary/doomscrolling>
36. Merriam-Webster. (2021). Ratio. In Merriam-Webster online dictionary. Retrieved December 1, 2024, from <https://www.merriam-webster.com/dictionary/ratio>
37. Merriam-Webster. (2022). Slay. In Merriam-Webster online dictionary. Retrieved December 1, 2024, from <https://www.merriam-webster.com/dictionary/slay>
38. Merriam-Webster. (2024). Enshittification. In Merriam-Webster online dictionary. Retrieved December 1, 2024, from <https://www.merriam-webster.com/dictionary/enshittification>
39. Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.
40. Naslund, J. A., Bondre, A., Torous, J., & Aschbrenner, K. A. (2020). Social media and mental health: benefits, risks, and opportunities for research and practice. *Journal of technology in behavioral science*, 5(3), 245-257.

41. Nunberg, G. (1990). The linguistics of punctuation (No. 18). Center for the Study of Language (CSLI).
42. Patton, M. Q. (2002). *Qualitative research and evaluation methods*. Thousand Oaks, Cal.: Sage Publications, 4(10.4135), 9781452226058.
43. Rogers, E. (2003). *Diffusion of Innovations* 5th.
44. Rosch, E. (2024). Principles of categorization. In *Cognition and categorization* (pp. 27-48). Routledge.
45. Shifman, L. (2013). *Memes in digital culture*. MIT press.
46. Silverman, D. (2021). *Doing qualitative research*.
47. Stéfán, S., & Rockwell, G. (2016). *Voyant tools*. Boston: Muska & Lipman/Premier-Trade.
48. Squires, L. (2010). Enregistering internet language. *Language in society*, 39(4), 457-492.
49. Tagliamonte, S. A., & Denis, D. (2008). Linguistic ruin? LOL! Instant messaging and teen language. *American speech*, 83(1), 3-34.
50. Thurlow, C., & Mroczek, K. (Eds.). (2011). *Digital discourse: Language in the new media*. Oxford University Press.
51. Traugott, E. C., & Dasher, R. B. (2001). *Regularity in semantic change* (Vol. 97). Cambridge University Press.
52. Ullmann, S. (1962). *Semantics: An introduction to the science of meaning* (p. 30). Oxford: Basil Blackwell.
53. Van Dijk, T. A. (2015). Critical discourse analysis. *The handbook of discourse analysis*, 466-485.
54. Walther, J. B. (2011). Theories of computer-mediated communication and interpersonal relations. *The handbook of interpersonal communication*, 4, 443-479.
55. Yus, F. (2011). *Cyberpragmatics: Internet-mediated communication in context*. John Benjamins Publishing Company.
56. Zappavigna, M. (2012). *Discourse of Twitter and social media*.
57. Zappavigna, M. (2018). *Searchable talk: Hashtags and social media metadiscourse*. Bloomsbury Publishing.