



Article

Age and Platform Effects on Syntactic Complexity in Uzbek Social Media Discourse

Abdurayimova Durdona

1. PhD Researcher, Alisher Navoi Tashkent State University of Uzbek Language and Literature
- * Correspondence: abdurayimovadurdona21@gmail.com

Abstract: This study explores the influence of age and communication platform on syntactic complexity in Uzbek social media discourse, focusing on Telegram and Facebook. Data were collected from 150 participants across five age groups, each contributing approximately 300 words, resulting in a 50,000-word corpus. Three syntactic indicators were looked at: the percentage of simple sentences, the proportion of compound sentences, and the average sentence length. Age and platform had an important effect on syntactic structure ($p < 0.001$), according to descriptive statistics, two-way Welch ANOVA, and Games–Howell post-hoc tests. Younger users preferred shorter, simpler constructions, while older users produced longer, more complex sentences. In comparison with Telegram users, Facebook users showed more syntactic elaboration. These findings highlight that syntactic variation reflects both cognitive maturity and platform-specific communication norms in Uzbek computer-mediated communication.

Keywords: syntactic complexity, average sentence length, simple sentence, compound sentence, age variation, platform effect, digital communication, Uzbek language, social media, Facebook, Telegram.

Citation: Durdona, A. Age and Platform Effects on Syntactic Complexity in Uzbek Social Media Discourse. Central Asian Journal of Literature, Philosophy, and Culture 2026, 7(1), 94-101

Received: 10th Aug 2025
Revised: 16th Sep 2025
Accepted: 24th Oct 2025
Published: 08th Nov 2025



Copyright: © 2026 by the authors. Submitted for open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>)

1. Introduction

In recent years, digital communication has profoundly transformed the structure and dynamics of written discourse. Social media platforms like Facebook and Telegram have created hybrid environments in which written messages have grown more and more like spoken communication—brief, impromptu, and frequently emotionally charged. The syntax of digital texts was affected by these modifications in addition to lexical and pragmatic choices (Crystal; Herring) [1]. Sentence complexity is a significant indicator of how users adapt their writing styles for online contexts in this new communicative paradigm. Two fundamental syntactic patterns—complex and simple phrases—reflect users' communicative intent, cognitive processing, and ability to speak. While older users usually use longer, more structured, and formally cohesive sentences, younger users typically produce shorter, syntactically simpler, and more expressive constructions (Tagliamonte & Denis, Biber, Halliday) [2], [3]. The communicative economy principle—the tendency to reduce grammatical complexity while maximizing efficiency and immediacy—is an important driver over syntactic simplification in computer-mediated communication (CMC). However, the extent of this simplification varies across age groups and platforms, influenced by socio-pragmatic and technological factors. Consequently, sentence length and clause complexity can serve as potential stylometric markers for identifying users' demographic characteristics, particularly age (Goswami et al., Pennebaker) [4]. Although numerous studies have investigated age-related syntactic variation in English-based social media discourse, there remains a lack of empirical research focused on Uzbek-language digital communication. Given the increasing

linguistic diversity of online discourse in Uzbekistan, the present study seeks to fill this gap by examining how sentence length and syntactic complexity differ across age groups and communication platforms (Telegram and Facebook) [5]. The primary aim of this study is to analyze the impact of age and platform on syntactic complexity in digital communication. Specifically, it seeks to determine:

- a. Whether age significantly influences sentence length and the distribution of simple and compound sentences;
- b. Whether communication platforms (Telegram vs. Facebook) affect syntactic patterns;
- c. Whether an interaction exists between age and platform factors in shaping sentence complexity [6].

The findings are expected to contribute to sociolinguistic profiling, digital stylistics, and forensic linguistic applications by identifying syntactic markers that are sensitive to demographic variation in social media discourse.

2. Materials and Methods

2.1. Participants and Data Collection

The study analyzed digital messages from **150 Uzbek-speaking participants**, divided equally into **five age groups**: 14–19, 20–29, 30–39, 40–49, and 50+ years (30 participants per group)[7]. Each participant contributed approximately **300 words** of naturally produced written communication drawn from **Telegram and Facebook**.

Thus, each age group contained around **10,000 words**, yielding a total dataset of approximately **50,000 words**. To ensure ecological validity, the messages were collected from real, natural social media interactions. The corpus included only personally written messages, not texts that were copied, forwarded, or quoted.

All data were anonymized and used only for linguistic analysis in accordance with rigorous ethical guidelines. The collection process complied with accepted research ethics and digital ethnography principles (Herring; Androutsopoulos)[8].

2.2. Variables and Measures

Three **syntactic indicators** were analyzed for each participant:

Average Sentence Length (M) – the average number of words per sentence.

Proportion of Simple Sentences (%) – the share of sentences containing a single predicative unit.

Proportion of Compound Sentences (%) – the share of sentences containing coordination or subordination structures[9].

Each sentence was manually segmented and coded according to **standard Uzbek syntactic criteria**. Messages consisting only of emojis, interjections, or incomplete clauses were excluded from the syntactic count to maintain analytical consistency.

These indicators were chosen because they have been widely recognized in previous linguistic and stylometric research as reliable measures of **syntactic complexity** and **communicative economy** (Biber, Tagliamonte & Denis, Pennebaker)[10].

2.3. Analytical Procedure

The analysis proceeded in **three main stages**:

Descriptive Statistics – For all variables, descriptive statistics were computed, including means (M), standard deviations (SD), minimums, maximums, and medians by age group and platform (Telegram vs. Facebook)[11].

Inferential Statistics – A **two-way Welch ANOVA** was conducted to test the main effects of **Age** and **Platform** and their interaction (**Age × Platform**) on each syntactic variable. The **Welch** variant was chosen because it is robust against unequal variances and slightly unbalanced group sizes (Cohen)[12].

Post-hoc Comparisons – Pairwise differences between age groups were examined using the **Games–Howell** test, which does not assume equal variances or sample sizes.

Effect sizes were calculated using **eta-squared (η^2)**, following **Cohen's** guidelines:

$\eta^2 \geq 0.14$ — **large effect**

$0.06 \leq \eta^2 < 0.14$ — **medium effect**

$\eta^2 < 0.06$ — **small effect**

All statistical analyses were performed using **SPSS (version 27.0)** and **Jamovi (version 2.5)**. The significance level was set at $p < 0.05$ [13].

2.4. Research Questions

This study addressed the following research questions:

Does **age** significantly influence sentence length and the proportions of simple and compound sentences in Uzbek social media discourse?

Does the **communication platform** (Telegram vs. Facebook) affect syntactic complexity?

Is there a significant **interaction between age and platform** in shaping sentence structure?

3. Results

3.1 Descriptive Statistics

Table 1. Descriptive Statistics for Sentence Length and Syntactic Complexity across Age Groups and Platforms

Age Group	Platform	N	Average Sentence Length (M \pm SD)	Simple Sentences (M \pm SD) (%)	Compound Sentences (M \pm SD) (%)
14–19	Telegram	30	4.62 \pm 2.58	79.34 \pm 19.00	20.66 \pm 19.00
20–29	Telegram	15	5.18 \pm 1.46	74.09 \pm 12.28	25.92 \pm 12.28
20–29	Facebook	15	8.40 \pm 1.35	47.14 \pm 21.77	47.68 \pm 15.98
30–39	Telegram	15	5.36 \pm 1.69	71.86 \pm 21.32	28.14 \pm 21.32
30–39	Facebook	15	7.16 \pm 1.66	70.68 \pm 13.03	29.32 \pm 13.03
40–49	Telegram	15	6.77 \pm 2.34	64.36 \pm 15.29	35.64 \pm 15.29
40–49	Facebook	15	7.73 \pm 1.59	54.38 \pm 12.46	45.62 \pm 12.46
50+	Telegram	15	7.98 \pm 2.91	55.63 \pm 23.86	44.37 \pm 23.86
50+	Facebook	15	9.89 \pm 2.20	53.39 \pm 20.84	46.61 \pm 20.84

Note. M = mean; SD = standard deviation. Sentence length and syntactic proportions show consistent increases in complexity with age and platform (Facebook > Telegram)[14].

Sentence length and syntactic complexity demonstrated different age- and platform-related patterns, according to descriptive analysis. The age group of 14–19 years old produced the shortest sentences (M = 4.62) and the highest number of simple sentences (79.3%) among Telegram users. On the other hand, Facebook users between 20 to 29 used more complicated constructions and longer sentences (M = 8.40). With longer sentences on Facebook (M = 7.16) and shorter ones on Telegram, the 30- to 39-year-old group showed moderate syntactic complexity. Facebook users once more produced significantly longer and more complex sentences (M = 7.73; compound = 45.6%) among those aged 40–49, while Telegram users preferred simpler structures. Lastly, the 50+ group showed the highest percentage of compound clauses and the longest sentences overall (M = 9.89), suggesting a more complex and formal writing style. In overall, sentence length increased gradually with age, **while the percentage of simple sentences fell and the amount of compound constructions grew**. Across platforms, **Facebook messages** were consistently longer and syntactically more complex than those on Telegram.

3.2. Two-Way Welch ANOVA

Table 2. Two-Way Welch ANOVA Results for the Effects of Age and Platform on Syntactic Variables

Indicator	Factor	F (Welch)	df ₁	df ₂ (approx.)	p-value	Result
Average Sentence Length	Age	28.63	4	≈94	< 0.001	Significant
	Platform	41.12	1	≈108	< 0.001	Significant
	Age × Platform	3.84	4	≈98	0.006	Interaction present
Simple Sentences (%)	Age	8.97	4	≈92	< 0.001	Significant
	Platform	22.64	1	≈108	< 0.001	Significant
	Age × Platform	2.91	4	≈96	0.025	Interaction present
Compound Sentences (%)	Age	9.88	4	≈93	< 0.001	Significant
	Platform	19.47	1	≈108	< 0.001	Significant
	Age × Platform	3.12	4	≈95	0.019	Interaction present

Note. All effects are statistically significant at $p < 0.05$. Interaction effects indicate that the influence of platform varies by age group.

The findings indicate that age and platform impact syntactic variations both independently and simultaneously. Importantly, the impact of platform increases with age: Facebook users produced significantly longer and more complex sentences for those aged 40–49 and 50+, whereas platform differences were negligible for younger users (14–19) [15].

Age and Platform both had significant main effects on all three syntactic variables, in accordance to the two-way Welch ANOVA results ($p < 0.001$).

Furthermore, each variable's interaction effect (Age × Platform) was statistically significant ($p < 0.05$), indicating that platform influence varied through age group.

Note = statistically significant ($p < 0.05$)

3.3. Post-hoc Analysis (Games–Howell)

Table 3. Games–Howell Post-hoc Comparison of Age Groups for Sentence Length and Syntactic Ratios

Group 1	Group 2	ΔM (ASL)	p (ASL)	d (ASL)	ΔM (S)	p (S)	d (S)	ΔM (C)	p (C)	d (C)	Overall Effect Size
G1	G2	-1.63	0.021	0.62	+12.8	0.042	0.70	-12.8	0.041	-0.70	Medium
G1	G3	-1.58	0.034	0.60	+8.1	0.128	0.46	-8.1	0.133	-0.46	Small–Medium

G1	G4	-2.79	< 0.001	1.11	+22.3	< 0.001	1.41	-22.3	< 0.001	-1.41	Very Large
G1	G5	-4.52	< 0.001	1.95	+25.0	< 0.001	1.19	-25.0	< 0.001	-1.19	Very Large
G2	G3	-0.05	0.994	0.02	-4.7	0.553	-0.26	+4.7	0.557	0.26	None
G2	G4	-1.16	0.083	0.50	+9.5	0.076	0.57	-9.5	0.079	-0.57	Moderate (approx.)
G2	G5	-2.89	< 0.001	1.15	+12.2	0.023	0.60	-12.2	0.021	-0.60	Large
G3	G4	-1.11	0.101	0.56	+14.2	0.019	0.90	-14.2	0.018	-0.90	Large (S/C only)
G3	G5	-2.94	< 0.001	1.15	+17.0	0.007	0.82	-17.0	0.006	-0.82	Large
G4	G5	-1.83	0.008	0.68	+2.8	0.841	-0.13	-2.8	0.844	0.13	Medium (ASL only)

Note:

ΔM (ASL) = Mean difference in **average sentence length**,

ΔM (S) = Mean difference in **simple sentences (%)**,

ΔM (C) = Mean difference in **compound sentences (%)**.

Positive ΔM (S) values indicate higher simple sentence proportions, whereas negative ΔM (C) values reflect fewer compound sentences in Group 1 compared to Group 2.

Cohen's d was used for effect size classification (Cohen): small (0.2), medium (0.5), large (0.8+) [16].

The **Games–Howell post-hoc comparisons** revealed that the largest differences occurred between the 14–19 and 50+ groups. Effect sizes were substantial:

Average Sentence length: $d = 1.95$ (large effect)

Simple/compound ratio: $d \approx 1.2$ (large effect)

Significant differences were also observed between younger (14–19) and middle-aged groups (30–39, 40–49). However, adjacent age groups (20–29 ↔ 30–39; 40–49 ↔ 50+) showed smaller or statistically nonsignificant differences. These results confirm that syntactic complexity rises gradually with age, becoming most pronounced in the contrast between adolescent and senior users

3.4. Effect Size (η^2) Analysis

Effect-size statistics (η^2) based on Welch ANOVA indicate that both **Age** and **Platform** had strong effects on **average sentence length**, while medium effects were found for **simple** and **compound** sentence proportions. The interaction effect was moderate across all measures.

Table 4. Effect Size (η^2) Summary for Age, Platform, and Their Interaction on Syntactic Variables

Variable	η^2 (Age)	η^2 (Platform)	η^2 (Age \times Platform)	Effect Level
Average Sentence Length	0.31 (large)	0.28 (large)	0.07 (medium)	Strong independent, moderate interaction
Simple Sentences (%)	0.18 (medium)	0.20 (medium)	0.05 (small)	Medium effect
Compound Sentences (%)	0.19 (medium)	0.22 (medium)	0.06 (medium)	Medium effect

Note. According to Cohen (1988), $\eta^2 \geq 0.14$ = large effect; $0.06 \leq \eta^2 < 0.14$ = medium effect; $\eta^2 < 0.06$ = small effect.

3.5. Summary of Findings

The overall results demonstrate that:

Syntactic complexity varies greatly through age, with older users creating longer and more complex sentences.

Syntactic patterns vary significantly by platform; Telegram favors brief, oral-like syntax, while Facebook encourages elaborate sentence structures.

Platform differences become more apparent in older cohorts, due to the platform-age interaction. Each of these findings give support to the idea that syntactic complexity in digital communication is a manifestation of platform-specific communication norms and also cognitive maturity.

4. Discussion

The findings of this study demonstrate systematic variation in syntactic complexity across age groups and communication platforms within Uzbek digital discourse. These results confirm that both **age** and **platform** are powerful determinants of users' syntactic behavior in computer-mediated communication (CMC).

Younger participants (aged 14–19) preferred shorter and syntactically simpler sentences, primarily composed of single-clause utterances. The immediacy, spontaneity, and informality that define social media interactions are apparent in this tendency (Crystal; Tagliamonte & Denis) [17]. Their writing style is very like their spoken discourse, which is in keeping with Herring's concept of hybrid digital communication, which is a combination of written and spoken components that reflect technological limitations.

On the other hand, older consumers—in particular those among the years of 40 and 49 and 50 and over—produced more intricate and longer sentences, indicating higher levels of syntactic elaboration and planning. This pattern corresponds to the frameworks for Biber and Halliday, that connect formality, cognitive maturity, and deliberate discourse structuring to syntactic complexity. Messages from older users frequently displayed subordination and logical connections, suggesting a more considerate communication style (Pennebaker) [18].

Syntactic structure was additionally significantly affected by the platform factor. Regardless of age group, Facebook users consistently produced longer and more complex sentences compared to Telegram users. This finding can be attributed to the semi-public and more formal communicative setting of Facebook, which encourages extended, coherent expression. Telegram, conversely, facilitates rapid, interactive, and speech-like

exchanges where brevity and emotional immediacy dominate (Androutsopoulos, Barton & Lee).

The **interaction between Age and Platform** (Age × Platform) revealed that syntactic variation is not uniform across age cohorts. For younger participants, platform effects were insignificant, but for older users, that they became more apparent. As stated by Herring & Dainas, this suggests that elderly individuals intentionally adapt their language to fit the affordances and intended demographics of each medium.

The effect-size analysis demonstrates how robust these differences are ($\eta^2 = 0.31$ for age; $\eta^2 = 0.28$ for platform). The most sensitive indicator proved to be sentence length, that was followed by the percentage of simple and compound sentences. This is in line with earlier stylometric studies that discovered that clause ratio and sentence length are reliable traits for author profiling and sociolinguistic differentiation (Grieve, Rangel et al., Tschuggnall & Specht) [19].

From a more broad point of view, these patterns show how Uzbek digital writing retains culturally distinctive elements while incorporating global CMC trends. The coexistence of formal and informal syntactic styles across age groups indicates that online language expression is affected by social positioning, communicative intent, and digital literacy.

5. Conclusion

This study investigated how **age** and **platform** influence syntactic complexity in Uzbek social media communication, focusing on Telegram and Facebook messages. The results confirmed that both factors — and their interaction — significantly affect average sentence length and the distribution of simple versus compound sentences. **Age** strongly predicts syntactic complexity: younger users employ shorter, simpler sentences, while older users produce longer and more compound constructions. **Platform** shapes linguistic style: Facebook fosters more structured and elaborated writing, whereas Telegram encourages brevity and conversational tone. **Interaction effects** reveal that older users exhibit greater syntactic adaptability across platforms than younger users.

These findings provide empirical evidence that syntactic complexity reflects both **cognitive development** and **technological adaptation** in Uzbek CMC. From an applied viewpoint, the identified indicators — clause type proportions and average sentence length (ASL) — can be used as reliable stylometric indicators in sociolinguistic studies, forensic linguistics, and author profiling.

In conclusion, the language used on Uzbek social media evolves due to both advances in technology and generational differences. Digital communication still has patterned variation reflecting age, experience, and communicative purpose, yet it promotes simplicity and expressiveness. Future research should extend this analysis by integrating **lexical richness**, **emoji usage**, and **discourse-pragmatic features**, thereby offering a more comprehensive view of linguistic variation in the Uzbek digital sphere.

REFERENCES

- [1] J. Androutsopoulos, "Computer-mediated communication and linguistic variation," in *The International Encyclopedia of Communication*, K. B. Jensen, Ed. Berlin: De Gruyter Mouton, 2014.
- [2] D. Barton and C. Lee, *Language Online: Investigating Digital Texts and Practices*. Routledge, 2013.
- [3] D. Biber, *Dimensions of Register Variation: A Cross-Linguistic Comparison*. Cambridge University Press, 1995.
- [4] E. Cambria and B. White, "Jumping NLP curves: A review of natural language processing research," *IEEE Computational Intelligence Magazine*, vol. 9, no. 2, pp. 48–57, 2014, doi: 10.1109/MCI.2014.2307227.
- [5] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Lawrence Erlbaum Associates, 1988.
- [6] D. Crystal, *Language and the Internet* (2nd ed.). Cambridge University Press, 2006.
- [7] S. Goswami, S. Sarkar, and M. Rustagi, "Stylometric analysis of bloggers' age and gender," in *Proceedings of the International Conference on Weblogs and Social Media (ICWSM)*, 2009, pp. 214–217.
- [8] J. Grieve, "Quantitative authorship attribution: An evaluation of techniques," *Literary and Linguistic Computing*, vol. 22, no. 3, pp. 251–270, 2007, doi: 10.1093/lc/fqm020.

-
- [9] M. A. K. Halliday, *An Introduction to Functional Grammar* (3rd ed.). London: Arnold, 2004.
- [10] S. C. Herring, "Discourse in Web 2.0: Familiar, reconfigured, and emergent," in *Discourse 2.0: Language and New Media*, D. Tannen and A. M. Trester, Eds. Georgetown University Press, 2013, pp. 1–25.
- [11] S. C. Herring and A. R. Dainas, "Gender and age variation in interactional features of digital discourse," *Language@Internet*, vol. 18, no. 3, pp. 1–28, 2020.
- [12] J. W. Pennebaker, *The Secret Life of Pronouns: What Our Words Say About Us*. Bloomsbury Press, 2011.
- [13] F. Rangel, P. Rosso, I. Chugur, and M. Potthast, "Overview of the Author Profiling Task at PAN 2017: Gender and Language Variety Identification in Twitter," *CEUR Workshop Proceedings*, vol. 1866, pp. 1–23, 2017.
- [14] S. A. Tagliamonte and D. Denis, "Linguistic ruin? LOL! Instant messaging and teen language," *American Speech*, vol. 83, no. 1, pp. 3–34, 2008, doi: 10.1215/00031283-2008-001.
- [15] M. Tschuggnall and G. Specht, "Enhancing authorship attribution by utilizing syntax tree profiles," in *Working Notes Papers of the CLEF 2014 Evaluation Labs*, 2014, pp. 195–205.
- [16] F. Yus, *Cyberpragmatics: Internet-mediated Communication in Context*. Amsterdam: John Benjamins, 2011.
- [17] A. Boutet, H. Kim, and E. Yoneki, "What's in your tweets? Investigating microblogging usage on Twitter," in *International Conference on Social Informatics*, 2012, pp. 20–33, Springer
- [18] S. C. Herring and J. Androutsopoulos, "Computer-mediated discourse 2.0," in *The Handbook of Discourse Analysis* (2nd ed.), D. Tannen, H. Hamilton, and D. Schiffrin, Eds. Wiley-Blackwell, 2015, pp. 127–151.
- [19] S. M. Mohammad and P. D. Turney, "Crowdsourcing a word–emotion association lexicon," *Computational Intelligence*, vol. 29, no. 3, pp. 436–465, 2013.