BG Pulse Journal USA
ISSN:1947-2536 http://www.bgpulseusa.com/index.php/bg

STYLISTIC FEATURES OF SUBTITLE AND DUBBING TEXTS IN THE ENGLISH LANGUAGE

Bozorova Vasila Ilkhomovna
Teacher, Department of Practical Translation of English Language,
State World Language University.
E-mail: ilkhomovnavasila07@gmail.com

Abstract: This article examines the specialized linguistic, syntactic, and pragmatic features of subtitle and dubbing texts within English audiovisual discourse. Operating within a poly-semiotic and multimodal screen environment, audiovisual translation (AVT) acts as a dynamic system where verbal paths must continuously balance artistic intent against strict technical and cognitive constraints. Through a comparative structural analysis, this study maps the core differences between the visual, cross-modal shift of subtitling and the oral, iso-modal replacement of dubbing. The findings reveal that subtitling relies on intense lexical density updates, syntactic simplification, and the systematic removal of conversational discourse markers (pragmatic pruning) to respect spatial boundaries (35–42 characters per line) and temporal guidelines. Conversely, dubbing focuses on phonetic harmony, mouth shape alignment (lip-sync), and matching the exact timing of breath tracks (isochrony) to maintain natural dialogue cadence and physical realism. Additionally, this article evaluates the ongoing tension between spoken socio-dialectal speech variations and the standardization rules of written screen text. Ultimately, the paper demonstrates that English screen texts are never static, word-for-word copies of spoken dialogue; instead, they are highly specialized communicative frameworks engineered to manage visual cognitive load while fully preserving narrative depth and cinematic immersion.

Keywords: Audiovisual Translation (AVT), Media Discourse, Subtitling Text, Dubbing Scripts, Stylistic Compression, Multimodal Cohesion, Pragmatic Pruning, Isochrony, Lip-Sync Harmony, Conversational Registers.

The global distribution of digital screen media has transformed how linguistic texts are produced, transmitted, and consumed. At the heart of this media landscape sits audiovisual translation (AVT), a specialized branch of applied linguistics that manages the transfer of verbal communication within multimodal environments. [1]Within this field, English functions not only as a primary global source language for cinematic distribution but also as a highly structured linguistic system that must be re-engineered across two distinct translation modalities: subtitling and dubbing.[2] Subtitling involves a cross-modal shift where spoken dialogue is transformed into synchronous on-screen text, while dubbing requires an iso-modal, oral-to-oral substitution that replaces the original vocal track with a localized performance.[3]

[1] Tavadze, L., Diasamidze, I., Katamadze, N., & Davitadze, L. (2024). Modern tendencies in media discourse. International Journal of Innovative Technologies in Social Science, 1(41). https://doi.org/10.31435/rsglobal_ijitss/30032024/8123
[2] Yusupov, U. K. (2013). Contrastive Linguistics of the English and Uzbek Languages. Tashkent: Akademnashr.
[3] Chaume, F. (2012). Audiovisual Translation: Dubbing. Manchester: St. Jerome Publishing.

Audiovisual discourse is fundamentally poly-semiotic, meaning that language never operates in a vacuum.[4] On screen, verbal strings must coexist harmoniously with visual tracks, body language, musical scores, atmospheric sound design, and rapid editorial cuts. [5]This complex interaction gives rise to a compelling linguistic paradox: screen text must successfully simulate the spontaneous, fluid, and unpredictable nature of natural human speech while remaining bound by rigid technical, physical, and cognitive constraints.[6] This article explores the complex stylistic features of English subtitle and dubbing texts, analyzing how the strict demands of time, space, and sound reshape syntax, lexical density, conversational markers, and sociolinguistic realism.

The Architectural Divide: Spatial and Temporal Constraints

The structural properties of an audiovisual text are shaped directly by its operational channel. Subtitling forces an oral, transient acoustic signal into a permanent graphic layout on screen.[7] This transition is governed by strict spatial and temporal boundaries designed to match human cognitive processing speeds. In global subtitling standards, an on-screen text block is restricted to a maximum of two lines, with each line containing between 35 and 42 characters, inclusive of spaces and punctuation.[8] Temporally, subtitle lines must adhere to the six-second rule, which dictates that a full two-line block should remain visible for no longer than six seconds, and no less than one second, to prevent cognitive fatigue or visual "flashing." Furthermore, the exposure velocity must remain within a comfortable reading threshold of 12 to 17 characters per second.[9]

These technical constraints demand a high degree of stylistic compression.[10] To ensure the viewer can read the text while simultaneously absorbing the visual composition of the scene, the text designer must systematically strip away linguistic redundancies. This process requires a careful re-engineering of natural phrasing. Spoken clauses are stripped of elaborate modifiers, complex nested subordinate clauses are broken down into simple sentences, and progressive aspect verbs are systematically flattened into simple past or present tenses.[11] Despite this severe structural pruning, the text must fully preserve the narrative's underlying semantic core and emotional urgency.[12]

In contrast, the stylistic framework of dubbing operates on an entirely different axis, shifting focus from visual compression to physical and acoustic synchronization. The dubbing scriptwriter is not bounded by line-length parameters but must achieve strict labial harmony (lip-sync) and isochrony.[13] Labial harmony demands that the translated vocabulary match the visible mouth movements of the on-screen actor, a constraint that becomes exceptionally acute during tight close-up shots. For instance, if an actor closes their mouth on a bilabial consonant (b, p, m) or opens it wide on a prominent open vowel, the translated text must utilize words that mirror those exact

[4] Delabastita, D. (1989). Translation and Mass Communication: Film and TV Translation as Evidence of Cultural Dynamics. Babel, 35(4), pp. 193-218.
[5] Dobrosklonskaya, T.G. (2008). Medialinguistics: A Systematic Approach to the Study of the Language of the Media. Moscow: Flinta: Nauka.
[6] Bakiyev, F. J. (2026). Subtitr tarjimasi tadqiqotlarida korpusga asoslangan yondashuvlar: “O'tkan kunlar” filmi misolida. Qo'qon Davlat Pedagogika Instituti Ilmiy Xabarlari, 1(1), pp. 237-241.
[7] Gottlieb, H. (1992). Subtitling — A new university discipline. In Teaching Translation and Interpreting (pp. 161–170). Amsterdam: John Benjamins.
[8] Diaz Cintas, J., & Remael, A. (2020). Subtitling: Concepts and Practices. London: Routledge.
[9] Perego, E. (2019). Information loss in subtitling: A corpus-based study on the translation of modern audiovisual texts. Perspectives: Studies in Translation Theory and Practice, 27(3), pp. 412–429.
[10] Nida, E. A., & Taber, C. R. (1974). The Theory and Practice of Translation. Leiden: E.J. Brill.
[11] Mamatov, A. E. (2021). Zamonaviy tilshunoslik muammolari [Problems of Modern Linguistics]. Tashkent: O'qituvchi.
[12] Catford, J. C. (1965). A Linguistic Theory of Translation. London: Oxford University Press.
[13] Chaume, F. (2012). Audiovisual Translation: Dubbing. Manchester: St. Jerome Publishing.

physiological movements.[14] Isochrony requires that the total duration of the spoken target utterance match the precise time frame of the original speaker's breath track, ensuring that the audio begins and terminates in perfect alignment with the visual performance.[15] Consequently, while subtitle texts favor visual conciseness and lexical pruning, dubbing scripts prioritize acoustic fluidity, natural cadence, and rhythmic elasticity.[16]

Lexical Density and the Mechanics of Pragmatic Pruning

Natural, unscripted human speech is characterized by low lexical density—meaning it features a high proportion of grammatical function words relative to content-carrying words—as well as frequent repetitions, self-corrections, and phatic expressions. However, when these loose conversational structures are transformed into subtitle blocks, they undergo an intense process of pragmatic pruning. This stylistic transformation alters the structural balance of the dialogue, significantly increasing its lexical density.

To stay within technical character limits, standard English screen dialogue frequently sacrifices polite modal hedging and indirect syntactic structures. For example, a spoken line such as, "I was wondering if you could perhaps tell me where the station is?" accumulates 65 characters and requires significant time to process aurally. When condensed into a subtitle, this string is stripped of its modal padding ("I was wondering," "perhaps") and reshaped into a direct interrogative: "Where is the train station?", compressing the line to a lean 28 characters.

While this structural compression effectively manages the viewer's cognitive load, it introduces a significant linguistic tradeoff. Stripping away modal verbs and hesitations fundamentally alters the interpersonal nuances of character interaction. By removing polite hedging, the character's voice is made to sound far more direct, assertive, and less vulnerable than the original acoustic performance implies, illustrating how technical constraints directly shape characterization.

The Pragmatic Function of Conversational Discourse Markers

Discourse markers, such as well, you know, look, I mean, honestly, and by the way, serve as the vital connective tissue of organic English conversation. These elements rarely carry denotative, dictionary-defined semantic meaning. Instead, they operate on a pragmatic level to regulate dialogue pacing, signal transitions in focus, establish interpersonal power hierarchies, and soften confrontational assertions.

In audiovisual translation, these markers present a recurring structural dilemma. Within subtitling matrices, where every character space is highly contested, discourse markers are typically the first elements targeted for complete deletion. If an actor exclaims, "Well, honestly, I just don't think it's a good idea, you know?", the compressed subtitle will almost invariably read: "I don't think it's a good idea." Although this deletion saves valuable on-screen space, it completely flattens the speaker’s hesitation, self-doubt, or desire for validation, removing the subtle psychological layers embedded within the scene.

In the domain of dubbing, however, the wholesale omission of discourse markers can cause severe communicative failure. Because these conversational anchors are frequently accompanied by physical gestures—such as a nod of the head, a wave of the hand, or a specific pause in breath—deleting them from the audio track creates a jarring disconnect between what the viewer sees and

[14] Akhmanova, O. S. (2014). Slovar lingvisticheskikh terminov [Dictionary of Linguistic Terms]. Moscow: KomKniga.
[15] Remael, A. (2010). Audiovisual translation. In Handbook of Translation Studies (Vol. 1, pp. 12–17). Amsterdam: John Benjamins.
[16] Salimov, R. D. (2024). Kinodiskursda milliy koloritni ifodalovchi leksik birliklar tarjimasi [Translation of lexical units expressing national color in cinematic discourse]. O’zbekistonda Xorijiy Tillar, 3(52), pp. 114–121.

what they hear. If a character shifts their body language to deliver an explanation while saying "Well...", the acoustic track must provide a corresponding vocalization to fill that physical space. Therefore, dubbing scripts cannot simply discard these markers; instead, they must systematically replace them with pragmatically equivalent target markers that preserve the natural, rhythmic cadence of human interaction.

The Conflict Between Vernacular Registers and Graphic Convention

The representation of socio-dialectal variation, regional accents, and urban slang highlights a core conflict in audiovisual discourse: the tension between auditory reality and graphic convention. Screen writers frequently employ non-standard dialects, such as African American Vernacular English (AAVE), Cockney rhyming slang, or specific regional variations, to instantly anchor a character within a precise socio-economic, geographic, or historical context.

[Oral Domain: High Realism] ⟶ [Cross-Modal Processing] ⟶ [Graphic Domain: Standardization]

Non-Standard Dialects	Phonetic Compression	Strict Orthographic Constraints
(e.g., "gonna", "wanna")	vs. Orthographic Clutter	(Preserves Scannability)

When these highly informal oral structures are converted into graphic subtitles, they encounter the inherent standardization bias of the written word. Printing non-standard phonetic spellings literally on screen (such as "gonna," "wanna," "cha," or "innit") often creates visual clutter. Rather than register a regional accent, the viewer's brain frequently flags these spellings as typographical errors, disrupting the automatic scanning process required for comfortable subtitle reading.

To resolve this conflict, professional AVT script designers rely on the principle of stylistic compensation. Instead of attempting to replicate non-standard syntax or broken orthography on screen, translators embed the informal register within the vocabulary itself. By selecting specific slang terms or idiomatic expressions that carry equivalent social meaning, they preserve the character's background and identity without violating standard graphic conventions or disrupting the reader's visual experience.

Conclusion

The stylistic features of subtitle and dubbing texts in the English language are shaped by an ongoing negotiation between artistic expression and technical boundaries. Subtitling relies on strategic compression, syntactic simplification, and pragmatic pruning, transforming loose oral dialogue into a clean, lexically dense, and highly scannable graphic layout. Dubbing operates on an entirely different plane, prioritizing phonetic harmony, acoustic pacing, and conversational flow to preserve the physical illusion of lip synchrony.

Ultimately, these transformations show that an audiovisual text is never a static, word-for-word copy of a spoken script. It is a highly specialized linguistic system engineered to balance the constraints of time, space, and sound—ensuring that global audiences can navigate complex screen narratives seamlessly without ever breaking the cinematic spell.

References

1. Akhmanova, O. S. (2014). Slovar lingvisticheskikh terminov [Dictionary of Linguistic Terms]. Moscow: KomKniga.
2. Alimova, I. X. (2025). Audiovizual tarjima matnlarida uchraydigan diskurs markerlar tahlili [Analysis of discourse markers encountered in audiovisual translation texts]. Journal of Universal Science Research, 3(4), pp. 254–259.
3. Ashurova, D. U., & Galieva, M. R. (2018). Stylistics of Literary Text. Tashkent: Alisher Navoiy nomidagi O'zbekiston Milliy kutubxonasi nashriyoti.

4. Bakiyev, F. J. (2026). Subtitr tarjimasi tadqiqotlarida korpusga asoslangan yondashuvlar: “O'tkan kunlar” filmi misolida [Corpus-based approaches in subtitle translation studies: On the example of the film "O'tkan kunlar"]. Qo'qon Davlat Pedagogika Instituti Ilmiy Xabarlari, 1(1), pp. 237–241.
5. Catford, J. C. (1965). A Linguistic Theory of Translation. London: Oxford University Press.
6. Chaume, F. (2012). Audiovisual Translation: Dubbing. Manchester: St. Jerome Publishing.
7. Delabastita, D. (1989). Translation and Mass Communication: Film and TV Translation as Evidence of Cultural Dynamics. Babel, 35(4), pp. 193–218.
8. Diaz Cintas, J., & Remael, A. (2020). Subtitling: Concepts and Practices. London: Routledge.
9. Dobrosklonskaya, T. G. (2008). Medialinguistics: A Systematic Approach to the Study of the Language of the Media. Moscow: Flinta: Nauka.
10. Gambier, Y. (2013). The position of audiovisual translation studies. In The Routledge Handbook of Translation Studies (pp. 61–75). London: Routledge.
11. Gottlieb, H. (1992). Subtitling — A new university discipline. In Teaching Translation and Interpreting (pp. 161–170). Amsterdam: John Benjamins.
12. Karamuratova, J. (2023). Ingliz va o’zbek tillaridagi frazeologizmlarning lingvomadaniy xususiyatlari [Linguacultural characteristics of phraseological units in English and Uzbek languages]. Inovatsion Texnologiyalar va Nashriyot Davri, 2(1), pp. 89–94.
13. Mamatov, A. E. (2021). Zamonaviy tilshunoslik muammolari [Problems of Modern Linguistics]. Tashkent: O'qituvchi.
14. Nida, E. A., & Taber, C. R. (1974). The Theory and Practice of Translation. Leiden: E.J. Brill.
15. Perego, E. (2019). Information loss in subtitling: A corpus-based study on the translation of modern audiovisual texts. Perspectives: Studies in Translation Theory and Practice, 27(3), pp. 412–429.
16. Remael, A. (2010). Audiovisual translation. In Handbook of Translation Studies (Vol. 1, pp. 12–17). Amsterdam: John Benjamins.
17. Salimov, R. D. (2024). Kinodiskursda milliy koloritni ifodalovchi leksik birliklar tarjimasi [Translation of lexical units expressing national color in cinematic discourse]. O’zbekistonda Xorijiy Tillar, 3(52), pp. 114–121.
18. Tavadze, L., Diasamidze, I., Katamadze, N., & Davitadze, L. (2024). Modern tendencies in media discourse. International Journal of Innovative Technologies in Social Science, 1(41). https://doi.org/10.31435/rsglobal_ijitss/30032024/8123
19. Toury, G. (2012). Descriptive Translation Studies — and beyond (Revised ed.). Amsterdam: John Benjamins.
20. Yusupov, U. K. (2013). Contrastive Linguistics of the English and Uzbek Languages. Tashkent: Akademnashr.