

AN ANALYSIS OF WORD FORMATION PROCESSES IN THE OPPENHEIMER MOVIE SCRIPT

Areeya Sompamit, Sukanya Kaowiwattanakul

English Department, School of Liberal Arts, University of Phayao, Thailand

Corresponding Author E-mail: areeya120644sompamit@gmail.com

Received April 1, 2026

Revised April 29, 2026

Accepted May 1, 2026

Abstract

Despite extensive morphological research on cinematic discourse, limited attention has been given to scientifically-oriented movie scripts, which differ significantly from general-language films. This study investigates word-formation processes in the scientifically oriented discourse of the film *Oppenheimer*, a domain that has received limited attention in prior film-based morphological research, which has largely focused on informal or general-language scripts. Drawing on the frameworks of Bauer (1983), Delahunty & Garvey (2010), and Yule (2010), the study adopts a mixed qualitative–quantitative approach to identify and analyse eleven word-formation types. Moving beyond simple description, the analysis reveals a "morphological fingerprint" of specialized communication; findings prove a clear dominance of derivation (67.01%), followed by compounding (17.77%) and abbreviation (5.92%). This distribution reflects not merely lexical choice, but the prevalence of technical and formal lexical construction necessary for establishing disciplinary authority in scientific cinematic discourse. The study contributes to morphology and discourse studies by extending word-formation analysis into a specialised genre and highlighting how linguistic patterns align with disciplinary communication. Pedagogically, the findings support the use of film scripts as authentic resources for developing morphological awareness and academic vocabulary in English language learning.

Keywords: Morphology, Movie Script, Word Formation Processes

Introduction

The study of morphology, particularly descriptive linguistics, provides a fundamental framework for understanding the internal structure of words and how they are systematically formed within a language (Gleason, 1955; Katamba, 1993). Recent linguistic research has shifted from general descriptions of language use to examining how language operates within specialised discourse. Word formation, as a key mechanism of lexical expansion, reflects how meaning is constructed across different communicative contexts (Yule, 2010; Fromkin et al., 2018). However, a critical gap remains in film-based morphological research; previous studies on word formation in films have predominantly

focused on informal language, slang, or entertainment-oriented scripts, leaving scientifically oriented discourse largely overlooked (Nugraha, 2025).

Furthermore, while films are recognized as multimodal texts combining spoken and written language, provide valuable linguistic data (Xu & Wijitsopon, 2023), existing literature often fails to employ systematic analysis through advances in corpus linguistics and multimodal discourse analysis. In addition, most existing studies emphasise vocabulary learning outcomes rather than examining how morphological processes vary across discourse types, leaving a significant gap in understanding word formation in specialised contexts.

To address these limitations, the film *Oppenheimer* presents a unique and necessary case for study. Beyond being a popular medium, it integrates scientific terminology, formal register, and historical discourse alongside everyday interaction, offering a suitable and complex corpus for analysing how word-formation processes function in discipline-specific communication (Chaudhari, 2023).

Therefore, this study aims to identify the types and frequencies of word-formation processes in the *Oppenheimer* script and to examine their role in shaping lexical patterns in specialised discourse (McIntosh, 2024). The findings contribute to morphology and discourse studies and provide pedagogical implications for vocabulary and morphological development through authentic media.

Research Objectives

1. To identify the types of word formation processes used in the movie *Oppenheimer*
2. To determine the frequency of word formation types in the movie *Oppenheimer*

Methodology

This study employed a mixed-methods design integrating qualitative and quantitative approaches to analyse word-formation processes in the script of *Oppenheimer*. The qualitative part focused on identifying and classifying word-formation types based on explicit operational definitions, while the quantitative component measured their frequency and distribution across the dataset.

The data consisted of the full movie script, obtained from the Internet Movie Script Database. To ensure data accuracy, the script was systematically cross-checked against the film. The dataset was treated as a small, specialised corpus, using corpus linguistics principles to enable a systematic and replicable analysis of lexical patterns within a scientifically oriented cinematic discourse.

Sampling was conducted using a purposive criterion-based approach. All lexical items in the script that exhibited identifiable word-formation processes were included in the analysis. The unit of analysis was the word token, and each occurrence was counted once. Function words, proper nouns without morphological modification, and items not meeting the operational definitions of word formation were excluded. This ensured consistency and reduced ambiguity in data selection.

To ensure analytical rigor, the study established strict operational definitions for each category based on Bauer (1983), Delahunty & Garvey (2010), Yule (2010), and Audring (2014). Derivation was defined as the addition of affixes resulting in a change of meaning or grammatical category (e.g., refugee, disloyal). Compounding referred to the combination of two or more free morphemes forming a single lexical unit (e.g., wormhole). Crucially, Multiple Processes were defined as lexical items involving two or more distinct morphological mechanisms (e.g., postgrad), thereby distinguishing them from pure compounding or derivation.

To enhance reliability and provide sufficient proof of validity, a second coder independently analysed a subset of the data (20%). The inclusion criteria for sampling and analysis were further validated by recognizing the fundamental role of morphological awareness in the recognition and processing of complex lexical items (Solaja & Crepaldi, 2024). Inter-coder agreement was calculated to assess consistency in classification, achieving a satisfactory level of agreement. Discrepancies were discussed and resolved to refine the coding scheme and ensure analytical precision.

For quantitative analysis, frequency counts and percentage distributions were calculated for each word-formation type using corpus-based measurement. This integration of qualitative classification and quantitative measurement ensured both depth and validity in the analysis.

Results

The researcher presents the research findings in two main points, as follows:

1. Types of word formation processes

The analysis of the Oppenheimer script reveals the use of ten out of eleven word-formation processes proposed in the theoretical frameworks. The identified processes include derivation, compounding, abbreviation, acronym, multiple processes, clipping, borrowing, conversion, coinage, and blending, while backformation was not observed. The data reveals a significant concentration of specific morphological strategies rather than an equal distribution across categories. Derivation appears as the dominant process, followed by compounding and abbreviation. This hierarchy indicates that the script uses a "Morphological Fingerprint" characterized by productive and conventional mechanisms of lexical expansion.

Instead of relying on informal lexical creativity, derivation is strategically used to construct abstract and technical meanings, aligning with the formal and scientific nature of discourse. Compounding functions as a key strategy for Strategic Lexical Condensation, compressing complex concepts into single lexical units particularly within disciplinary discourse related to institutional roles, scientific entities, and technical objects. Abbreviation and acronyms, concentrated in institutional and academic references, further contribute to communicative efficiency within the high-stakes dialogue. In contrast, marginal processes such as clipping,

borrowing, and blending appear only in limited context-specific instances, proving that creative word-formation plays a minor role compared to standardised processes.

2. Frequency of word formation processes

The quantitative analysis identified a total of 1,925 instances of word-formation processes. As shown in Table 1, derivation accounts for the majority of occurrences (67.01%), followed by compounding (17.77%) and abbreviation (5.92%). All remaining processes individually contribute less than 3% of the total.

This distribution confirms a structural trend favoring affixation as the primary mechanism of lexical construction. The high proportion of derivation indicates that meaning is frequently extended through systematic morphological modification to meet the epistemological demands of the genre. Compounding serves as a secondary efficiency-driven strategy, enabling the condensation of complex semantic content into concise forms, while abbreviation supports technical efficiency in recurring institutional or technical references.

Conversely, the minimal representation of coinage and blending are minimally represented, suggesting that the script prioritizes clarity and precision over lexical innovation. The absence of backformation further reinforces the tendency towards expansion rather than reduction in lexical development. Overall, the results reveal a hierarchical pattern of word-formation processes, where highly productive and rule-governed mechanisms dominate, reflecting the formal, scientific, and historically grounded discourse of the film.

Table 1 The frequency and percentage of each type of word formation process.

Type of Word Formation Process	Frequency	Percentage
Derivation	1,290	67.01%
Compounding	342	17.77%
Abbreviation	114	5.92%
Acronym	47	2.44%
Multiple Processes	46	2.39%
Clipping	33	1.71%
Borrowing	25	1.30%
Conversion	23	1.19%
Coinage	3	0.16%
Blending	2	0.10%
Backformation	-	-
Total	1,925	100%

Representative Examples

Representative examples from the dataset illustrate the strategic dispersion of dominant and less frequent word-formation processes function across different discourse

contexts in the Oppenheimer script. The data proves a clear trend where derivation, the most prevalent category, occurs in forms such as refugee, disloyal, and statement. In these instances, affixation enables shifts in grammatical category and the construction of abstract or evaluative meanings essential for political and scientific discourse.

Compounding, observed in items such as chairman, wormhole, and air force, functions as a key strategy for Strategic Lexical Condensation, condensing complex institutional and technical concepts into single lexical units to maintain communicative density. Conversion, though less frequent, appears in forms like talk and channel, reflecting flexible category shifts without morphological modification, while abbreviations and acronyms such as Dr. and TNT prioritize efficiency in referring to professional titles and scientific terminology. Conversely, less frequent processes, including clipping (lab, Yank), borrowing (quantum, nucleus), and blending (Caltech), occur in marginal, context-specific instances, often associated with informal or highly specialised lexical usage. In addition, the presence of multiple processes, as seen in postgrad and uranium ore, illustrate the potential for layered morphological constructions that combine more than one mechanism. Overall, these examples confirm the hierarchical distribution found in the quantitative findings by proving that highly productive processes, particularly derivation and compounding, dominate the dataset, while less productive or more creative processes remain constrained by the formal requirements of the genre.

Discussion

The findings of this study confirm that derivation is the dominant word-formation process, followed by compounding and abbreviation; however, this study moves beyond merely corroborating previous research by critically analyzing how morphological patterns are shaped by discourse type and communicative function. As shown in the data, the high frequency of derivation reflects the need to construct abstract, technical, and evaluative meanings within a scientifically oriented narrative. This supports the view that derivation functions as a productive and conventional mechanism of lexical expansion (Bauer, 1983; Plag, 2003). This finding further resonates with Jelimun (2024), who demonstrates that mastery of such morphological processes is a fundamental requirement for constructing meaning in formal and academic writing. Crucially, the present study interrogates this dominance, proving its specific role in maintaining precision and conceptual clarity in formal and disciplinary discourse, a function often undervalued in general linguistic descriptions.

Similarly, the prominence of compounding must be interpreted not only as a structural process but as a deliberate communicative strategy. In the context of the Oppenheimer script, compounds serve to condense complex scientific and institutional concepts into manageable lexical units, enabling efficient transmission of dense information. This finding aligns with Delahunty & Garvey (2010), who emphasise the relationship between

word formation and communicative purpose. However, the current results challenge the "general tendency" narrative by suggesting that compounding operates as a form of lexical compression in high-information contexts. Abbreviation and acronyms reinforce this pattern, indicating a broader tendency towards efficiency and standardisation in specialised discourse.

A critical discrepancy emerges when identifying the hierarchical distribution of word-formation processes, where highly productive and rule-governed mechanisms contrasts sharply with findings from studies of informal or slang-based films, such as Mahmood & Ahmed (2023), where coinage is more prominent. This divergence provides a robust basis for arguing that word-formation processes are rigidly influenced by register and discourse function, where linguistic choices also serve to establish the specific professional identity and community belonging within the narrative context (Jenkins, 2007). In informal genres, lexical creativity and identity expression are prioritised, whereas in scientifically oriented discourse, clarity, precision, and conventionality are more salient.

Furthermore, while previous studies (e.g., Safitri (2022), Nindyakirana (2022), Prachumchai & Humna (2022)) consistently report the dominance of derivation and compounding, this study problematizes their lack of context-specific depth by linking these processes to the epistemological demands of scientific and historical narratives. The reliance on derivation enables the expression of abstract concepts such as evaluation, authority, and theoretical reasoning, while compounding facilitates the representation of technical entities. This shows that morphological choices are not merely linguistic phenomena but are strategic tools embedded within broader discourse practices.

Overall, the findings suggest that word-formation processes in film discourse are shaped by genre and communicative purpose rather than occurring randomly. By integrating insights from morphology and discourse analysis, this study offers a more nuanced understanding of how lexical structures function within specialised cinematic texts, moving beyond descriptive accounts towards a theoretically informed explanation of linguistic patterns.

Body of Knowledge

This study offers significant theoretical and practical contributions by moving beyond descriptive morphological analysis toward a context-driven exploration of specialized cinematic discourse:

Genre-Specific "Morphological Fingerprint": The study sets up that scientific-historical narratives possess a distinct "morphological fingerprint". Unlike informal films that prioritize lexical creativity through coinage or slang, this genre relies on derivation (67.01%) and compounding (17.77%) as strategic drivers to construct "Technical Credibility" and "Institutional Authority". These findings prove that morphological frequency is rigidly dictated by the epistemological demands of the genre.

"Strategic Lexical Condensation" Model: The research identifies a structural hierarchy utilized as a specialized communicative strategy. Through the systematic use of compounding and abbreviations (5.92%), the discourse achieves "Strategic Lexical Condensation," enabling the efficient transmission of dense scientific constructs and historical data within the inherent temporal constraints of cinematic dialogue.

Pedagogical Framework for High-Stakes English: The study contributes a novel pedagogical template for EFL/ESL instruction by repositioning specialized scripts as authentic tools for advanced "Morphological Awareness". It offers a practical framework for learners to decode academic and bureaucratic "terms of art," providing a systematic method for mastering word structures within high-stake professional and disciplinary contexts.

Conclusion and Recommendations

The findings demonstrate that ten word-formation processes are present in the script of Oppenheimer, with derivation as the dominant process (67.01%), followed by compounding (17.77%) and abbreviation (5.92%). This distribution reflects a strong reliance on productive and conventional morphological mechanisms to construct abstract, scientific, and institutional meanings. The results suggest that word-formation processes are closely aligned with the communicative demands of specialised discourse, where clarity, precision, and conceptual density are prioritised over lexical creativity.

From a pedagogical perspective, the findings can be translated into structured and workable classroom applications that leverage authentic cinematic data to enhance linguistic competence. Teachers can implement Derivation-focused "Category Shifting" Tasks, utilizing specialized terms from the Oppenheimer script, such as disloyal and statement, to create exercises where learners predict changes in grammatical categories and semantic nuances by manipulating affixes. Furthermore, Compounding "Deconstruction-Reconstruction" Workshops allow students to analyze how complex scientific constructs like wormhole and air force are encoded, practicing the decomposition of compound words to master technical lexical density. These can be supplemented by Scene-Based "Morphological Mapping", where students extract vocabulary from specific clips and engage in critical discussions on how these forms establish professional authority and technical credibility within high-stakes professional discourse. Such Corpus-Informed Exercise Design supports the development of morphological awareness and academic vocabulary in EFL/ESL contexts through the use of authentic, high-stakes materials.

Despite these pedagogical contributions, the study's limitations must be thoroughly examined to guide future linguistic inquiry. A primary Genre Restriction exists because the analysis is limited to a single scientific-historical script, preventing the generalizability of this specific "morphological fingerprint" to informal or hybrid film genres. Additionally, there is a Lack of Functional Integration, as the study focuses primarily on formal classification and

frequency while leaving the complex interactions between word-formation and semantic, pragmatic, or sociolinguistic factors largely unexplored. The dataset Scale, relying on a relatively small, specialized corpus, further constrains the ability to identify broader linguistic patterns. Future research should address these gaps by conducting comparative, larger-scale corpus-based studies across diverse film genres and integrating additional linguistic perspectives to provide a more comprehensive understanding of meaning-making in context.

Future research should address these gaps by conducting comparative, larger-scale corpus-based studies across diverse film genres. Integrating semantic and discourse analysis would further provide a comprehensive understanding of meaning-making in context. Pedagogically, future work should focus on evaluating structured instructional models that utilize corpus-informed film data through evaluated task-based learning activities.

References

- Audring, J. (2014). Booij, Geert: The Grammar of Words. An Introduction to Linguistic Morphology. *Morphology*, 24(4), 433–434.
- Bauer, L. (1983). *English Word-Formation*. Cambridge: Cambridge University Press.
- Chaudhari, A. M. (2023). The Global Language: English's Impact on the World. *International Journal of Research in Humanities & Soc. Sciences*, 11(01), 47-49.
- Delahunty, G. P. & Garvey, J. J. (2010). *The English Language: From Sound to Sense*. Fort Collins: Parlor Press.
- Fromkin, V., Rodman, R., & Hyams, N. (2018). *An Introduction to Language* (11th ed.). Boston: Cengage Learning.
- Gleason, H. A. (1955). *An Introduction to Descriptive Linguistics*. New York: Henry Holt and Company.
- Jelimun, M. O. (2024). The Application of Word Formation in Students' Essay Writing: Morphology. *Journal of English Language Teaching and Linguistics*, 9(2), 147–156.
- Jenkins, J. (2007). *English as a Lingua Franca: Attitude and Identity*. Great Clarendon Street: OUP Oxford.
- Katamba, F. (1993). *Morphology*. London: Macmillan Press.
- Mahmood, R. Q. & Ahmed, M. O. (2023). Morphology Awareness in Vocabulary Construction. *International Journal of Social Sciences & Educational Studies*, 10(3), 151–160.
- McIntosh, S. (2024). Oscars 2024: Oppenheimer sweeps awards as Cillian Murphy wins best actor. Retrieved April 1, 2025, from <https://www.bbc.com/news/entertainment-arts-68530809>
- Nindyakirana, S. A. (2022). An analysis of word formation process in Coco movie (Unpublished undergraduate thesis Department of English Literature). Semarang, Indonesia: Diponegoro University.
- Nugraha, D. S. (2025). A decade of research on word formation in foreign language education: A scientometric mapping (2014–2024). *Journal on English as a Foreign Language*, 15(2), 485–516.
- Plag, I. (2003). *Word-Formation in English*. Cambridge: Cambridge University Press.

- Prachumchai, S. & Humna, N. (2022). An Analysis of Language Use in Films on Social Status Classification and Racism: A Case Study of The Help and Hidden Figures. *Journal of Roi Kaensarn Academi*, 7(7), 177–193.
- Safitri, A. W. (2022). An analysis of word formation processes in *Incredibles 2* movie (Unpublished undergraduate thesis English Education Department). Indonesia: State Institute of Islamic Studies Ponorogo.
- Solaja, O., & Crepaldi, D. (2024). The role of morphology in novel word learning: A registered report. *Royal Society Open Science*, 11(6), 1-30.
- Xu, R. & Wijitsopon, R. (2023). Corpus Linguistics and Cinematic Discourse: Lexical Bundles in Mainstream Film Scripts. *LEARN Journal: Language Education and Acquisition Research Network*, 16(1), 545–574.
- Yule, G. (2010). *The Study of Language* (4th ed.). Cambridge: Cambridge University Press.