JURNAL MULTIDISIPLIN MADANI (MUDIMA)

Homepage: https://journalmudima.my.id/index.php/mudima
ISSN: 2808-5639 (Online)
Research Article

Vol. 6, No. 4, April (2026)	DOI: https://doi.org/10.55927/mudima.v6i4.36	Page: 412-417

Compounding as a Word-Formation Strategy in Modern English: A CWLM-Based Analysis

Usmonaliyeva Mohizarbonu[1], Karimjonova Shakhlo Ravshanjonovna[2*]

Fergana State University

Corresponding Author: Karimjonova Shakhlo Ravshanjonovna shahloxonkarimjonova@gmail.com

ARTICLE INFO

Keywords: Compounding, Word Formation, Neologisms, CWLM Model, Semantic Transparency, Discourse Analysis, Cognitive Linguistics, English Morphology, Second Language Acquisition, Translation

Received : 5 February
Revised : 22 March
Accepted : 25 April

ABSTRACT

Compounding is one of the most productive word-formation mechanisms in contemporary English. This paper examines how compound words are structured, classified, and used across different types of discourse, including scientific writing, media language, and everyday speech. Using the CWLM (Components-Word Formation-Lexical Meaning-Meaning Output) framework as the main analytical tool, and the Semantic Transparency Index (STI) as a classification instrument, this study analyzes twenty compound words drawn from three major academic sources. Results indicate that noun+noun structures are the most frequent type, while low-transparency compounds appear more commonly in informal contexts. The findings have implications for translation studies and second language vocabulary teaching

INTRODUCTION

Language changes constantly. Every year, new words enter common use, and existing words sometimes shift in meaning. One of the main forces behind this process is the need people have to name new things, like new technologies, social phenomena, and ideas that did not exist before. In English, one of the most common and reliable ways to create new words is compounding: combining two or more existing words to form a new unit with its own distinct meaning [1].

Compound words appear in almost every area of modern life. In digital technology, terms like smartphone, cloud storage, and data privacy have become standard. In public discourse, expressions such as fake news and social distancing entered everyday speech rapidly. In academic writing, compounds like gene expression and machine learning are essential tools for precise communication. This wide distribution across different domains shows that compounding is not simply a grammatical rule but it is a flexible, living process through which speakers create meaning in response to real communicative needs [2].

Researchers have long recognized compounding as a central feature of English word formation. Bauer identifies it as one of the most frequently used morphological processes available to English speakers, noting that it allows new concepts to be named quickly and with relative structural simplicity [3]. Plag builds on this view, arguing that compounding is strongly influenced by frequency of use: the patterns that appear most often in language become the easiest to apply when new words are needed [4]. Lieber and Štekauer provide a broader cross-linguistic perspective, confirming that compounding is especially productive in English compared to many other languages [5].

Despite this body of research, few studies have attempted to connect the structural, semantic, and cognitive dimensions of compounding within a single analytical model. This study addresses that gap by applying the CWLM framework that is a four-stage model that tracks the full process of compound formation from component selection to meaning stabilization. The research has three main objectives:

1. To analyze the structural patterns of compound words found in key academic sources.
2. To classify those compounds using the Semantic Transparency Index (STI).
3. To discuss the implications of these findings for translation and language teaching.

METHODS

This study uses a qualitative analytical approach. Rather than collecting primary data from speakers or texts, it draws on compound words already documented and discussed in established academic literature. This approach is appropriate for a study focused on theoretical classification and analysis, as it allows conclusions to be grounded in peer-reviewed sources [6].

Twenty compound words were selected from three major academic works in English morphology:

1. Lieber R. & Štekauer P. The Oxford Handbook of Compounding. Oxford University Press, 2010.
2. Plag I. Word-Formation in English. Cambridge University Press, 2003.
3. Bauer L. Morphological Productivity. Cambridge University Press, 2001.

These sources were chosen because they are widely cited in the field and contain extensive discussions of compound word types with concrete examples. The selected compounds represent a range of structural categories and levels of semantic transparency.

The CWLM Analytical Framework

The CWLM model is used in this study as a tool for describing the formation process of each compound. It breaks compounding down into four stages:

Table 1. CWLM Analytical Framework

Stage	Name	Description
C	Components	Identifying the individual lexical elements used
W	Word Formation	Combining elements through productive morphological patterns
L	Lexical Meaning	Constructing an initial semantic interpretation
M	Meaning Output	Stabilizing meaning through repeated contextual use

This model is useful because it shows that compounding is not a single act but a gradual process from the moment a new word is first created to the point where it becomes a fully established part of the language.

Semantic transparency index (STI)

To classify compounds by how easily their meaning can be understood from their parts, this study uses the Semantic Transparency Index (STI). Three levels are distinguished:

1. High STI - meaning follows directly from the components, such as notebook, sunlight.
2. Medium STI - meaning is partly predictable, like stepchild, highway.
3. Low STI - meaning cannot be reliably guessed from the parts, for example blackbird, daredevil.

This classification draws on Aitchison's distinction between transparent and opaque items in mental lexicon organization [7], and on Plag's discussion of semantic compositionality in compounds [8].

RESULTS AND DISCUSSION

Compound word analysis

The following table presents all twenty compounds analyzed in this study, along with their structural type, STI level, and source reference.

Table 2. Compound word analysis

N	Compound word	Structure	STI	Source
1	blackboard	Adj+N	Medium	Plag (2003)
2	toothbrush	N+N	High	Lieber & Štekauer (2010)
3	notebook	N+N	High	Plag (2003)
4	classroom	N+N	High	Bauer (2001)
5	football	N+N	Medium	Lieber & Štekauer (2010)
6	raincoat	N+N	High	Plag (2003)
7	handbag	N+N	High	Bauer (2001)
8	sunlight	N+N	High	Lieber & Štekauer (2010)
9	keyboard	N+N	High	Plag (2003)
10	headphone	N+N	High	Bauer (2001)
11	greenhouse	Adj+N	Low	Plag (2003)
12	blackbird	Adj+N	Low	Lieber & Štekauer (2010)
13	blueprint	Adj+N	Low	Bauer (2001)
14	full-time	Adj+N	High	Plag (2003)
15	highway	Adj+n	Medium	Lieber & Štekauer (2010)
16	pickpocket	V+N	Low	Bauer (2001)
17	scarecrow	V+N	Low	Plag (2003)
18	daredevil	V+N	Low	Lieber & Štekauer (2010)
19	washing-machine	V+N	High	Bauer (2001)
20	workplace	N+N	High	Plag (2003)

Structural Patterns

Analysis of the twenty compounds shows that noun+noun (N+N) is the most frequent structural pattern, accounting for 10 out of 20 items (50%). This finding is consistent with Plag's observation that N+N compounding is the default and most productive pattern in English, largely because it places minimal morphological demands on the speaker [9].

Adjective+noun (Adj+N) compounds make up 25% of the sample (5 items), while verb+noun (V+N) compounds account for 20% (4 items). One compound (full-time) uses an adjective+adjective structure, suggesting that while less common, other patterns do occur.

STI Distribution

Of the twenty compounds analyzed:

1. High STI: 11 items (55%) — toothbrush, notebook, washing-machine
2. Medium STI: 4 items (20%) — blackboard, football, highway
3. Low STI: 5 items (25%) — greenhouse, daredevil, blackbird

High STI compounds are the most frequent across all structural categories. However, it is notable that all four V+N compounds in the sample, except washing-machine, fall into the Low STI category. This suggests that verb-based compounds are more likely to develop opaque, idiomatic meanings over time that is a pattern also noted by Lieber and Štekauer in their discussion of exocentric compounds [10].

CWLM Analysis of Selected Items

Applying the CWLM model to selected examples illustrates how different compounds occupy different stages of the formation process:

1. Notebook (N+N, High STI): Components note and book are clearly identifiable (C-stage). The combination follows a standard N+N pattern (W-stage). Meaning is directly compositional: a book for notes (L-stage). Through widespread use, the meaning has become fully stable (M-stage complete).
2. Greenhouse (Adj+N, Low STI): Components green and house are identifiable (C-stage). The Adj+N pattern is productive (W-stage). However, the lexical meaning has shifted away from a literal "greenhouse" to refer specifically to a glass structure for growing plants (L-stage deviation). This shift is the result of long-term semantic specialization through use (M-stage).
3. Daredevil (V+N, Low STI): The components dare and devil are combined in an unusual way (C and W stages). The resulting meaning, a reckless or bold person, is not predictable from the parts (L-stage opacity). This compound has completed the M-stage fully, meaning it functions as a single stored unit in the mental lexicon [11].

Discussion

Structural Productivity and Cognitive Economy

The dominance of N+N compounds in this study reflects a broader principle in English morphology: speakers tend to favor structures that are easy to produce and recognize. Plag explains this in terms of cognitive economy that is the most frequent patterns become the most automatic, requiring less mental effort to apply [12]. This has a practical consequence: as new concepts emerge, especially in technology and science, English speakers naturally reach for N+N compounding as their first tool.

Within the CWLM framework, this tendency is most visible at the W-stage: the word formation step where productive patterns are selected. The regularity of N+N compounds suggests that this pattern has reached a high level of automaticity at the W-stage, making it the default choice for new lexical creation.

Semantic Transparency and Lexicalization

The STI results reveal an important tension in compound semantics. On one hand, most compounds in the sample are semantically transparent (High STI = 55%), which makes them easy to understand even for learners. On the other hand, a significant proportion (25%) are opaque, the meanings of which have drifted from the literal sum of their parts through a process of lexicalization.

Bauer describes lexicalization as the gradual process by which a word loses its compositional meaning and becomes stored as a fixed, unanalyzed unit [13]. This is clearly visible in cases like greenhouse and blackbird, where the original descriptive meaning no longer applies. In CWLM terms, these compounds have advanced through the full cycle: from C to W to L, and finally to a stable, non-compositional M-stage output.

Implications for Translation

The variation in semantic transparency creates real challenges in translation. Transparent compounds like notebook or sunlight can often be translated directly, since their meaning is close to the sum of their parts. However, opaque compounds like blueprint or daredevil require the translator to find a target-language equivalent that captures the established meaning rather than the literal components.

This is particularly relevant when translating between English and languages like Uzbek, where compounding works very differently. As Newmark argues, culturally embedded lexical items often require adaptive rather than literal translation strategies [14]. The STI classification developed in this study can serve as a practical guide for translators: High STI compounds are safer to translate directly, while Low STI compounds demand greater interpretive attention.

Implications for Language Teaching

For second language learners, the STI framework offers a useful pedagogical tool. Research has consistently shown that learners find transparent compounds easier to understand and remember than opaque ones. Nation emphasizes that building awareness of word structure is one of the most effective vocabulary learning strategies available to learners [15].

A CWLM-based teaching sequence could help learners engage with compounds more systematically:

1. C-stage: identify and understand each component word
2. W-stage: recognize the structural pattern (N+N, Adj+N)
3. L-stage: analyze the semantic relationship between parts
4. M-stage: practice using the compound in real contexts

This approach would be especially valuable for teaching Low STI compounds, where the meaning cannot be guessed from the parts and must be explicitly taught and practiced.

This study has several limitations. The sample of twenty compounds is relatively small and was drawn exclusively from academic sources. The analysis is qualitative rather than experimental, meaning that the STI classifications reflect analytical judgment rather than empirical measurement. Future studies could expand the corpus, include spoken language data, and test learners‘ actual comprehension of compounds at different STI levels.

CONCLUSION

This study has examined twenty compound words from three major academic sources, using the CWLM framework and Semantic Transparency Index to analyze their structure and meaning. The main findings are:

1. N+N compounding is the dominant structural pattern, consistent with its established status as the most productive pattern in English morphology
2. High STI compounds are the most common overall, though V+N compounds tend toward Low STI due to semantic drift
3. The CWLM model effectively captures the multi-stage nature of compound word formation, from component selection to meaning stabilization
4. The STI classification has direct practical value for both translation and language teaching

Taken together, these findings support the view that compounding in modern English is a dynamic, cognitively grounded process that responds to communicative and social needs. The CWLM model and STI framework provide complementary tools for understanding this process, and their application can be extended in future research to larger corpora and additional language pairs.

REFERENCES

Aitchison J. Words in the Mind: An Introduction to the Mental Lexicon. Oxford: Wiley-Blackwell, 2012. -P. 89-104.

Aitchison J. Words in the Mind: An Introduction to the Mental Lexicon. Oxford: Wiley-Blackwell, 2012. -P. 110-115.

Bauer L. Morphological Productivity. Cambridge: Cambridge University Press, 2001. -P. 12.

Bauer L. Morphological Productivity. Cambridge: Cambridge University Press, 2001. -P. 67-72.

Creswell J. W. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. Thousand Oaks: SAGE Publications, 2014. -P. 185-200.

Lieber R. & Štekauer P. (2010). The Oxford Handbook of Compounding. Oxford: Oxford University Press, 2010. -P. 56-62.

Lieber R. & Štekauer P. The Oxford Handbook of Compounding. Oxford: Oxford University Press, 2010. -P. 3-18.

Lieber R. & Štekauer P. The Oxford Handbook of Compounding. Oxford: Oxford University Press, 2010 -P. 7-9.

Nation I. S. P. Learning Vocabulary in Another Language. Cambridge: Cambridge University Press, 2001. -P. 270.

Newmark P. A Textbook of Translation. New York: Prentice Hall, 1988. -P. 83-96.

Plag I. Word-Formation in English. Cambridge: Cambridge University Press, 2003. Chapter 1.

Plag I. Word-Formation in English. Cambridge: Cambridge University Press, 2003. -P. 132-134.

Plag I. Word-Formation in English. Cambridge: Cambridge University Press, 2003. -P. 145-148.

Plag I. Word-Formation in English. Cambridge: Cambridge University Press, 2003 -P. 138.

Plag I. Word-Formation in English. Cambridge: Cambridge University Press, 2003. -P. 140-142.