Review Article

ENGLISH COMPOUND WORDS: ENDOCENTRIC AND EXOCENTRIC COMPOUNDS — A REVIEW OF STRUCTURE, SEMANTICS, PRODUCTIVITY, AND THEORETICAL PERSPECTIVES

Morphology

Keywords: English morphology; compounding; endocentric compounds; exocentric compounds; headedness; semantic compositionality; lexicalization; productivity; word formation; Construction Morphology.

Ajdin Bejtula — State University of Tetova, North Macedonia.

Abstract

Compounding represents one of the most productive mechanisms of lexical expansion in English and occupies a central position in morphological research. By combining two or more lexical bases, English speakers create expressions that refer to objects, individuals, activities, concepts, properties, and increasingly complex technological and social phenomena. One of the most influential classifications of English compounds distinguishes between endocentric and exocentric compounds. Endocentric compounds contain an internal constituent that functions as the semantic head of the compound, whereas exocentric compounds lack an internal semantic head and typically require an external interpretive relation. This review paper examines the morphological, syntactic, semantic, cognitive, and theoretical properties of English endocentric and exocentric compounds. It synthesizes major approaches to compound classification and critically examines the relationship between headedness, semantic compositionality, lexicalization, productivity, and contextual interpretation. We pay particular attention to the limitations of equating endocentricity with semantic transparency and exocentricity with complete semantic opacity. The review further considers recent developments in corpus linguistics, construction morphology, cognitive linguistics, psycholinguistics, computational linguistics, and English language pedagogy. The analysis demonstrates that headedness remains an important diagnostic for the classification of compounds but cannot independently account for the complexity of compound meaning. A multidimensional approach incorporating headedness, semantic transparency, lexicalization, productivity, contextual knowledge, and conceptual interpretation provide a more comprehensive account of English compounding. The paper concludes by identifying research opportunities concerning emerging digital compounds, psycholinguistic processing, corpus-based productivity, computational interpretation, and the pedagogical treatment of compounds in English as a second language.

1. INTRODUCTION

Compounding is one of the most productive processes of word formation in English. Through the combination of lexical bases, speakers continually create new lexical units capable of expressing objects, concepts, activities, social roles, scientific phenomena, and technological innovations. Examples such as *"blackboard," "toothbrush," "classroom," "football," "greenhouse," and "smartphone"* demonstrate the importance of compounding in both everyday and specialised English.

The productivity of English compounding is particularly evident in contemporary technological and digital discourse. Expressions such as *cloud computing, machine learning, data mining, social media, cybersecurity,* and *AI-generated content* illustrate the capacity of English to create new terminology by combining already available lexical material.

Compounds have traditionally been analyzed according to morphological, syntactic, and semantic criteria. Among the most influential distinctions is that between endocentric and exocentric compounds. This distinction is fundamentally associated with the concept of headedness. In an endocentric compound, one constituent functions as the semantic head and

determines the compound's basic category. Thus, a *toothbrush* is a type of *brush*, a *classroom* is a type of *room*, and a *raincoat* is a type of *coat*. In an exocentric compound, by contrast, no constituent functions as the semantic head of the whole expression. A *redhead*, for example, refers to a person with red hair rather than to a type of *head*.

The endocentric–exocentric distinction is useful because it provides a relatively straightforward means of examining the relationship between morphological structure and lexical meaning. Nevertheless, the traditional distinction becomes more problematic when semantic opacity, lexicalisation, metaphor, metonymy, and contextual interpretation are taken into consideration.

An endocentric compound may possess a clear head while exhibiting a specialised lexical meaning. Conversely, an apparently exocentric compound may remain interpretable through conventional conceptual associations. Thus, headedness and semantic transparency should not automatically be treated as equivalent properties.

The purpose of this review paper is to examine the traditional distinction between endocentric and exocentric compounds while integrating contemporary perspectives on compound formation.

The review addresses the following questions:

1. What structural properties characterise English compounds?
2. How is headedness manifested in endocentric compounds?
3. What makes a compound exocentric?
4. What is the relationship between headedness and semantic compositionality?
5. How do lexicalisation and semantic opacity influence compound interpretation?
6. How have recent theoretical and empirical approaches expanded the study of English compounding?
7. What are the implications of compound structure for language teaching, lexicography, and computational linguistics?

2. REVIEW METHODOLOGY

This article employs a narrative review methodology to synthesize major theoretical and descriptive contributions to the study of English compounding. The review focuses on established research in morphology, word formation, lexical semantics, construction morphology, cognitive linguistics, psycholinguistics, corpus linguistics, and computational linguistics.

The literature was selected according to four principal criteria. First, priority was given to established academic books and peer-reviewed studies dealing directly with English morphology and compounding. Second, major theoretical contributions to headedness, semantic interpretation, and compound classification were considered. Third, more recent approaches incorporating corpus

evidence, cognition, language processing, and computational analysis were included. Finally, studies relevant to English language teaching and lexical acquisition were considered, as they helped to comprehend the interpretation of compounds.

The review is organized around five analytical dimensions:

- morphological structure;
- headedness;
- semantic compositionality;
- lexicalization;
- productivity.

This approach recognizes that compounds cannot always be adequately explained through a single binary classification. Instead, the relationship between form and meaning is treated as multidimensional.

3. THEORETICAL BACKGROUND: COMPOUNDING AND HEADEDNESS

3.1 Compounding as Word Formation

Compounding involves the combination of two or more lexical bases to create a new lexical unit. Examples include:

black + *bird* → *blackbird*
tooth + *brush* → *toothbrush*
class + *room* → *classroom*
book + *case* → *bookcase*
rain + *coat* → *raincoat*

Compounding differs from derivation because it generally combines lexical bases rather than attaching an affix to a base.

English compounds may appear as closed, hyphenated, or open formations:

- *blackbird*
- *well-being*
- *ice cream*

Orthographic representation, however, should not be considered sufficient for determining compound status. The linguistic analysis must also consider constituent structure, grammatical behavior, semantic relations, and lexicalization.

3.2 Headedness

The notion of the head is fundamental to compound analysis. The head determines the basic semantic category of the compound and frequently determines its grammatical category.

Consider: *tooth* + *brush* → *toothbrush*

A toothbrush is a kind of brush. Therefore, *brush* functions as the head.

Likewise: *class* + *room* → *classroom*

A classroom is a type of room.

This relationship can be represented as: [Modifier + HEAD] → Compound

English noun compounds are characteristically right-headed, although this tendency is not universal.

Examples include:

- *snowball* → ball
- *bedroom* → room
- *football* → ball
- *teacup* → cup
- *sunflower* → flower

The right-headed pattern provides an important interpretive cue for English speakers.

4. ENDOCENTRIC COMPOUNDS

4.1 Definition and Structural Characteristics

An endocentric compound contains an internal constituent that functions as the semantic head. The whole compound denotes a subtype of the category represented by that constituent.

Examples include:

Compound	**Head**	**Interpretation**
blackbird	bird	a type of bird
toothbrush	brush	a type of brush
bookcase	case	a type of case
classroom	room	a type of room
raincoat	coat	a type of coat
football	ball	a type of ball

The IS-A relationship is therefore a useful diagnostic:

A toothbrush is a brush.
A classroom is a room.
A raincoat is a coat.

However, the IS-A relationship should be considered a diagnostic rather than a complete theory of compound semantics.

4.2 Grammatical and Semantic Properties

The head commonly determines the grammatical category of the compound.

For example: *tooth* + *brush* → *toothbrush* (noun)

Plural marking also frequently appears on the right-hand element:

- *toothbrush* → *toothbrushes*
- *classroom* → *classrooms*
- *bookcase* → *bookcases*

Many endocentric compounds are semantically transparent. However, transparency is not guaranteed.

For instance, *blackbird* is endocentric because it denotes a kind of bird, but its conventional meaning can be more specific than the literal combination *black* + *bird*.

Thus: Endocentricity does not necessarily entail complete semantic transparency.

5. EXOCENTRIC COMPOUNDS

5.1 Definition

An exocentric compound lacks an internal constituent that functions as the semantic head of the whole expression.

Examples include:

- *redhead* — a person with red hair;
- *pickpocket* — a person who steals from pockets;
- *scarecrow* — an object intended to frighten birds;
- *blockhead* — a person regarded as stupid;
- *loudmouth* — a person who talks excessively;
- *lazybones* — a lazy person.

In each case, the referent of the compound is not simply a subtype of either constituent.

5.2 Bahuvrihi Compounds

Exocentric compounds are traditionally associated with the term bahuvrihi, borrowed from Sanskrit grammatical tradition.

A bahuvrihi compound characterizes an external referent through a property associated with the constituents.

For example: *redhead* → a person characterized by red hair.

The compound therefore establishes a relationship between its constituents and an external referent.

However, the traditional category of exocentric compounds has been questioned in modern morphology because the category may encompass formations that differ significantly in their semantic mechanisms.

6. SEMANTIC COMPOSITIONALITY AND COMPOUND MEANING

Semantic compositionality refers to the extent to which the meaning of a complex expression can be derived from the meanings of its components and their structural relationship.

Highly transparent compounds include:

- *coffee cup*
- *school bus*
- *raincoat*
- *bookcase*

However, English also contains compounds whose meanings are specialized:

- *blackmail*
- *butterfly*
- *honeymoon*

Semantic opacity therefore cannot be equated directly with exocentricity.

A compound may be:

1. endocentric and transparent;
2. endocentric and partially opaque;
3. exocentric but conceptually interpretable;
4. highly lexicalized and opaque.

This multidimensional perspective is more consistent with contemporary morphological theory.

7. LEXICALIZATION AND BORDERLINE CASES

Lexicalization plays an important role in the development of compound meaning. A compound can begin as a transparent combination and subsequently acquire a conventionalized meaning.

Consider: *blackboard*

A blackboard is a type of board, making it structurally endocentric. However, its meaning is not necessarily restricted to a board that is black.

greenhouse

A greenhouse is structurally related to *house*, but its lexical meaning denotes a specialized structure for growing plants.

These cases demonstrate the importance of distinguishing:

- formal headedness;
- semantic headedness;
- semantic transparency;
- lexicalization;
- conventional meaning.

The existence of such borderline cases suggests that compounds are better analyzed along continua rather than through rigid binary categories.

8. ENDOCENTRICITY, EXOCENTRICITY, AND FIGURATIVE MEANING

Exocentric compounds frequently illustrate the relationship between morphology and figurative language.

In *loudmouth*, for example, *mouth* is associated conceptually with speech. The compound therefore develops a meaning referring to a person characterized by excessive or indiscreet talking.

Similar mechanisms occur in:

- *blockhead*
- *redhead*
- *egghead*
- *skinflint*
- *loudmouth*

The interpretation of such compounds may involve metaphor, metonymy, functional association, or culturally conventionalized meaning.

This supports approaches that treat compound interpretation as a conceptual rather than purely formal process.

9. COMPOUND CLASSIFICATION: A REVIEW OF MAJOR THEORETICAL APPROACHES

Research on English compounding has developed through several major theoretical traditions.

Approach	Major representatives	Main emphasis
Structural morphology	Bloomfield; Marchand	Constituent structure and formal patterns
Lexical morphology	Bauer; Lieber	Word formation and lexical representation
Generative morphology	Fabb; Spencer	Relationship between morphology and grammatical structure
Construction Morphology	Booij	Form–meaning schemas
Lexical semantics	Lieber	Semantic relations among constituents
Typological morphology	Scalise & Bisetto	Cross-linguistic compound classification
Cognitive approaches	Various researchers	Metaphor, metonymy, conceptual interpretation
Psycholinguistic approaches	Various researchers	Compound processing and lexical access
Corpus approaches	Bauer, Lieber & Plag and others	Frequency, productivity, and usage

The historical development of research demonstrates a movement from primarily structural descriptions toward multidimensional approaches incorporating semantics, cognition, corpus evidence, and language processing.

10. RECENT DEVELOPMENTS IN THE STUDY OF ENGLISH COMPOUNDING

10.1 Corpus Linguistics

Corpus linguistics has made it possible to investigate compounds through large collections of naturally occurring language.

Researchers can examine:

- compound frequency;
- constituent frequency;
- productivity;
- semantic change;
- register variation;
- emergence of new compounds.

Corpus-based approaches are particularly useful for studying productive patterns in contemporary English.

10.2 Construction Morphology

Construction Morphology views morphological patterns as conventionalized form–meaning pairings. This approach is particularly useful for compounds because it allows researchers to represent patterns in which meaning is partly predictable but also includes conventionalized information.

A constructional schema can capture relationships such as: [X + Y] → a Y associated with X

This approach therefore accommodates both productive patterns and partially lexicalized formations.

10.3 Cognitive Linguistics

Cognitive approaches emphasize:

- metaphor;
- metonymy;
- conceptual categorization;
- encyclopedic knowledge;
- conceptual associations.

This perspective is especially useful for exocentric compounds because their interpretation often depends on relationships extending beyond the literal meanings of the constituents.

10.4 Psycholinguistics

Psycholinguistic research examines whether compounds are processed as whole lexical units, decomposed into constituents, or represented through both mechanisms.

Variables such as:

- compound frequency;
- constituent frequency;
- semantic transparency;
- predictability;
- context

may influence compound processing.

10.5 Computational Linguistics

Compounds are increasingly important in Natural Language Processing.

Computational systems must interpret compounds in:

- machine translation;
- terminology extraction;

- information retrieval;
- text classification;
- question answering;
- language modeling.

A literal interpretation of *redhead*, for example, may produce an incorrect semantic representation. Modern contextual language models therefore provide an important new research environment for compound interpretation.

10.6 Digital Lexical Innovation

Technological developments have accelerated the creation of new English compounds.

Examples include:

- *smartphone*
- *smartwatch*
- *cybersecurity*
- *cloud computing*
- *machine learning*
- *data mining*
- *online learning*
- *content creator*
- *AI-generated content*

These formations demonstrate that compounding remains highly productive in contemporary English.

11. PEDAGOGICAL AND APPLIED IMPLICATIONS

11.1 English as a Second Language

Knowledge of compound structure can help learners infer unfamiliar vocabulary.

For example:

class + *room* → *classroom*
rain + *coat* → *raincoat*

However, literal interpretation becomes problematic with exocentric compounds:

redhead ≠ type of head
pickpocket ≠ type of pocket

Explicit instruction in headedness, semantic relationships, lexicalization, and figurative meaning can therefore improve morphological awareness.

11.2 Lexicography

The endocentric–exocentric distinction also has implications for dictionary compilation.

An endocentric compound can frequently be defined in relation to its head:

toothbrush: a brush used for cleaning teeth.

An exocentric compound requires a definition that explicitly identifies the external referent:

redhead: a person with red hair.

Lexicographers must therefore consider both internal structure and conventional lexical meaning.

11.3 Computational Applications

Understanding compound structure is relevant to computational terminology extraction and semantic analysis.

A system capable of distinguishing transparent compounds from lexicalized or exocentric formations may produce more accurate semantic representations.

This has applications in:

- machine translation;
- terminology management;
- information retrieval;
- educational technology;
- language modeling.

12. DISCUSSION

The literature reviewed demonstrates that the endocentric–exocentric distinction remains a valuable analytical framework but cannot provide a complete account of English compounding.

The distinction successfully explains basic contrasts such as: *toothbrush* → a type of brush

and: *redhead* → a person with red hair.

Nevertheless, the analysis becomes more complicated when semantic specialization, lexicalization, metaphor, and contextual knowledge are considered.

An endocentric compound may possess a clear head while developing a specialized meaning. Conversely, an exocentric compound may remain interpretable through conventional conceptual associations.

The evidence therefore supports a multidimensional approach:

Compound interpretation = morphological structure + headedness + semantic transparency + lexicalization + context + conceptual knowledge

Such an approach provides a more comprehensive account than a strict binary distinction between transparent endocentric and opaque exocentric compounds.

The literature further demonstrates that compounding intersects with several areas of linguistic research, including morphology, semantics, syntax, pragmatics, cognition, psycholinguistics, lexicography, computational linguistics, and language acquisition.

13. RESEARCH GAPS AND FUTURE DIRECTIONS

Several areas deserve further investigation.

13.1 Corpus-Based Research

Large contemporary corpora can provide evidence concerning the frequency and productivity of compound patterns.

13.2 Psycholinguistic Research

Experimental studies can compare the processing of transparent endocentric, opaque endocentric, and exocentric compounds.

13.3 Computational Modeling

Future NLP systems should be evaluated according to their ability to distinguish literal, compositional, metaphorical, and lexicalized compound meanings.

13.4 Cross-Linguistic Research

Comparative studies can investigate whether patterns of headedness and exocentricity found in English occur similarly in typologically different languages.

13.5 Language Teaching

Experimental pedagogical research could determine whether explicit instruction in compound structure improves vocabulary acquisition among learners of English.

13.6 Digital Lexical Innovation

Web and social-media corpora could be used to document the emergence and conventionalization of new compounds.

13.7 Constructional Analysis

Further research should investigate whether construction-based models provide a more comprehensive explanation of compound formation than traditional binary classifications.

14. CONCLUSION

Compounding is one of the most productive mechanisms of English word formation and plays a central role in the continuous expansion of the English lexicon. The distinction between endocentric and exocentric compounds provides an important starting point for understanding the relationship between compound structure and meaning.

Endocentric compounds contain an internal semantic head and generally denote subtypes of the category represented by that head. Exocentric compounds lack an internal semantic head and frequently depend on metaphorical, metonymic, functional, or culturally conventionalized interpretation.

The review demonstrates, however, that headedness should not be equated with semantic transparency. Endocentric compounds can become lexicalized and semantically specialized, while exocentric compounds can remain interpretable through conceptual associations. Consequently, the traditional binary classification should be supplemented by attention to semantic transparency, lexicalization, productivity, context, and conceptual knowledge.

Recent developments in corpus linguistics, Construction Morphology, cognitive linguistics, psycholinguistics, computational linguistics, and language teaching have broadened the study of English compounds. These perspectives demonstrate that compounds are dynamic form–meaning constructions that interact with lexical knowledge, cognition, context, and language use.

The continuing emergence of compounds associated with artificial intelligence, social media, cybersecurity, digital education, and other technological developments provides particularly promising opportunities for future research.

Overall, the evidence reviewed in this article supports a multidimensional approach to English compounding, in which headedness, semantic transparency, lexicalization, productivity, contextual interpretation, and conceptual knowledge are treated as interacting properties. Such an approach retains the descriptive value of the endocentric–exocentric distinction while providing a more comprehensive account of the complexity, flexibility, and productivity of English word formation.

References

Adams, V. (2001). *Complex Words in English*. Pearson Education.

Bauer, L. (1983). *English Word-Formation*. Cambridge University Press.

Bauer, L. (2003). *Introducing Linguistic Morphology* (2nd ed.). Edinburgh University Press.

Bauer, L. (2008). Exocentric compounds. *Morphology, 18*(1), 51–74. https://doi.org/10.1007/s11525-008-9122-5

Bauer, L. (2017). *Compounds and Compounding*. Cambridge University Press.

Bauer, L., Lieber, R., & Plag, I. (2013). *The Oxford Reference Guide to English Morphology*. Oxford University Press.

Bisetto, A., & Scalise, S. (2005). The classification of compounds. *Lingue e Linguaggio, 4*(2), 319–332.

Bloomfield, L. (1933). *Language*. Henry Holt.

Booij, G. (2005). *The Grammar of Words: An Introduction to Linguistic Morphology*. Oxford University Press.

Booij, G. (2010). *Construction Morphology*. Oxford University Press.

Booij, G. (2010). Construction morphology. *Language and Linguistics Compass, 4*(7), 543–555. https://doi.org/10.1111/j.1749-818X.2010.00213.x

Fabb, N. (1998). Compounding. In A. Spencer & A. M. Zwicky (Eds.), *The Handbook of Morphology* (pp. 66–83). Blackwell.

Lieber, R. (2004). *Morphology and Lexical Semantics*. Cambridge University Press.

Lieber, R. (2009). A lexical semantic approach to compounding. In R. Lieber & P. Štekauer (Eds.), *The Oxford Handbook of Compounding*. Oxford University Press.

Lieber, R. (2010). *Introducing Morphology*. Cambridge University Press.

Lieber, R., & Štekauer, P. (Eds.). (2009). *The Oxford Handbook of Compounding*. Oxford University Press.

Marchand, H. (1969). *The Categories and Types of Present-Day English Word-Formation: A Synchronic-Diachronic Approach* (2nd ed.). C. H. Beck.

Plag, I. (2003). *Word-Formation in English*. Cambridge University Press.

Scalise, S., & Bisetto, A. (2009). The classification of compounds. In R. Lieber & P. Štekauer (Eds.), *The Oxford Handbook of Compounding* (pp. 34–53). Oxford University Press.

Spencer, A. (1991). *Morphological Theory: An Introduction to Word Structure in Generative Grammar*. Blackwell.

Štekauer, P. (2012). Meaning predictability of novel context-free compounds. In R. Lieber & P. Štekauer (Eds.), *The Oxford Handbook of Compounding*. Oxford University Press.

Štekauer, P., & Lieber, R. (Eds.). (2005). *Handbook of Word-Formation*. Springer.