https://lexikos.journals.ac.za; https://doi.org/10.5788/36-1-2197 (Article)

Investigating the Features of Definitions from a General-Purpose Dictionary and AI-Generative Source: A Quantitative Comparative Study

Yuran Luo, *Center for Linguistics and Applied Linguistics, Guangdong University of Foreign Studies, Guangzhou, China*
(luoyuran0711@outlook.com)
(https://orcid.org/0009-0009-7599-4330)
and
Wenzhi Wu, *College of Liberal Arts, Jinan University, Guangzhou, China*
(Corresponding Author, wintenwu@gmail.com)
(https://orcid.org/0009-0007-9132-1860)

Abstract: Employing a multidimensional quantitative framework, this study undertakes a systematic comparison of definition features of mathematical terms from the *Modern Chinese Dictionary* (7th edition, *MCD*), a traditional Chinese general-purpose dictionary, and their AI-generated counterparts across four dimensions: lexical differences in definitions, textual similarity, semantic features, and syntactic complexity. The results reveal fundamental divergence between the two sources. (1) *MCD* definitions exhibit a stronger tendency toward direct definitions, whereas AI-generated ones are more explanatory. (2) In textual similarity analysis, the average similarity between the two types is quantified at 0.28, indicating a low level of similarity. (3) Regarding semantic features, *MCD* definitions prioritize essential features, with other features supplemented, while AI-generated ones exhibit a greater variety and a more balanced distribution of semantic features. (4) In terms of syntactic complexity, AI outputs manifest higher syntactic complexity. Based on these analyses, the study further discusses the implications for the cognitive load of definitions for users, the knowledge representation and organization in definitions, and reflections on dictionary compilation, thereby providing empirical insights for future comparative studies in traditional lexicography and AI-generated/assisted lexicography.

Keywords: *MODERN CHINESE DICTIONARY*, AI-GENERATED DEFINITIONS, DEFINITION FEATURES, MATHEMATICAL TERMS, QUANTITATIVE ANALYSIS

Opsomming: 'n Ondersoek na die kenmerke van definisies uit 'n algemene woordeboek en 'n KI-gegenereerde bron: 'n Kwantitatiewe vergelykende studie. Met behulp van 'n multidimensionele kwantitatiewe raamwerk onderneem hierdie studie 'n sistematiese vergelyking van die definisiekenmerke van wiskundige terme uit die *Moderne Chinese Woorde-*

https://lexikos.journals.ac.za; https://doi.org/10.5788/36-1-2197 (Article)

boek (7de uitgawe, *MCW*), 'n tradisionele Chinese algemene woordeboek, en hul KI-gegenereerde ekwivalente oor vier dimensies: leksikale verskille in definisies, tekstuele ooreenstemming, semantiese kenmerke en sintaktiese kompleksiteit. Die resultate toon fundamentele verskille tussen die twee bronne. (1) *MCW*-definisies neig meer na direkte definisies, terwyl KI-gegenereerde definisies meer verklarend van aard is. (2) In die tekstuele ooreenstemmingsanalise word die gemiddelde ooreenstemming tussen die twee tipes op 0.28 gekwantifiseer, wat op 'n lae vlak van ooreenstemming dui. (3) Wat semantiese kenmerke betref, prioritiseer die *MCW*-definisies noodsaaklike kenmerke, met ander kenmerke wat as aanvulling dien, terwyl KI-gegenereerde definisies 'n groter verskeidenheid en 'n meer gebalanseerde verspreiding van semantiese kenmerke toon. (4) Wat sintaktiese kompleksiteit betref, toon KI-uitsette hoër sintaktiese kompleksiteit. Gebaseer op hierdie ontledings, word die implikasies vir die kognitiewe lading van definisies vir gebruikers, die kennisuitbeelding en -struktuur in definisies, en oorwegings vir woordeboeksamestelling verder bespreek om sodoende empiriese insigte vir toekomstige vergelykende studies in die tradisionele leksikografie en KI-gegenereerde/-ondersteunde leksikografie te verskaf.

Sleutelwoorde: *MODERNE CHINESE WOORDEBOEK*, KI-GEGENEREERDE DEFINISIES, DEFINISIEKENMERKE, WISKUNDIGE TERME, KWANTITATIEWE ANALISE

1. Introduction

The rapid rise and significant breakthroughs of Large Language Models (LLMs) have posed challenges and opportunities for language research. In the field of lexicography, De Schryver (2023: 355) noted that "a new age that of the successful application of generative AI in lexicography has dawned." Many researchers have engaged in conducting exploratory studies on dictionary compilation in the AI era, encompassing the following topics. Early research focused on the evaluation of the capabilities of AI and LLMs in lexicography. For instance, Tarp and Gouws (2023) highlighted that AI has a huge impact on the compilation process and the processing of lexicographical data, while Jakubíček and Rundell (2023) raised concerns regarding its non-deterministic characteristics, stability, and reproducibility. Along with technological advancements, research on human-AI collaboration in the compilation process has advanced, covering discussions on the influence of AI in the lexicographical process (e.g., Klosa-Kückelhaus and Tiberius 2025), structured model propositions (e.g., Mandour 2024), automatic compilation processes (e.g., Liu 2025), and the integration of AI tools for different lexicographical needs (Bothma and Gouws 2025). Another critical topic is evaluating the quality of AI-generated outputs (e.g., De Schryver 2025a; De Schryver 2025b). Moreover, a growing body of research explores this from the user perspective, specifically comparing users' performance on AI against traditional dictionaries, such as the *Longman Dictionary of Contemporary English* (Ptasznik et al. 2024) and the *Cambridge Dictionary* (Albalawi et al. 2026), alongside investigating users' perception of AI tools (e.g., Ptasznik and Lew 2025; Liu et al. 2025; Han and De Schryver 2026). Furthermore, some research also explores the broader applica-

https://lexikos.journals.ac.za; https://doi.org/10.5788/36-1-2197 (Article)

tion of AI tools across different task scenarios, such as writing tasks (e.g., Huete-García and Tarp 2024; Li and Tarp 2024). These advancing models and research trends indicate that AI is reconstructing traditional lexicographical data, paradigms, and even compilation methods.

However, current research on dictionary definitions and AI-generated definitions is still in the exploratory phase. Dictionary definitions constitute a crucial component of a dictionary and are a focal point in lexicographical research. A large body of research focuses on the content structure of definitions, compilation methods, and characteristics of definitions. Structurally, traditional dictionary definitions typically include formats such as definitional, explanatory, synonymous, and antonymous definitions (Huang 1987). The "genus + differentia" structure is a prevalent format in definitions, which highlights taxonomic classification through a genus attribute and specifies distinctive features. In terms of compilation methods, traditional dictionary definitions are predominantly compiled manually. With the advancement of compilation technology, lexicographic practices have diversified to include corpus-assisted definition writing (Kilgarriff et al. 2004; 2014) and multimodal definitions (Zhang 2024). As for the definition characteristics, traditional definitions are characterized by essence prioritization, substitutability, grammatical function reflection, simplicity, brevity, and avoidance of ambiguity (Landau 2001: 163-170). However, the rise of LLMs increasingly challenges the structure, compilation methods, and features of definitions.

Since 2023, AI-generated or AI-assisted definitions have become a focal point in lexicographical research. Many studies have investigated AI-generated and AI-assisted definition writing, such as exploring DeepSeek in generating definitions and exemplifications (Chen et al. 2025), and evaluating the performance of AI in the retrieval and translation of English idioms into Chinese (Zheng and Chen 2026). The quality assessment of generated definitions has also attracted crucial research attention. For example, Lew (2023) evaluated the ability of ChatGPT to generate accurate and high-quality sense definitions and example sentences. Similarly, Zhang and Liu (2024) examined user satisfaction with AI-generated definitions. Beyond quality assessment, some studies have focused on the strategies for optimizing generated definitions; for example, Rundell (2023: 13-16) identified optimal prompting words in generated definitions, and Zhong (2025) developed prompt guides for users. From the users' perspective, some research has emphasized the empirical validation of AI-generated definitions, comparing user performance between traditional and AI-generated definitions (e.g., Rees and Lew 2024; Chen 2025). Much research, such as Lew (2024), has pointed out that AI-generated definitions exhibit contextual dynamics, human-intelligence interaction, and data-driven characteristics. Nonetheless, AI-generated definitions also present certain shortcomings, including insufficient definitional accuracy, stylistic inconsistencies, and knowledge hallucinations (e.g., Rundell 2023: 16; De Schryver and Joffe 2023).

These studies undertake a deep investigation into a fundamental question: how to effectively analyze and evaluate AI-generated and traditional definitions.

https://lexikos.journals.ac.za; https://doi.org/10.5788/36-1-2197 (Article)

In traditional, particularly corpus lexicography, dictionary definition compilation has incorporated multiple quantitative and computational methods for definition evaluations. These methods span the mutual information score (Church and Hanks 1990), Key Word in Context (Sinclair 1991), WordNet-based computation (Miller 1995), LogDice (Rychlý 2008), corpus pattern analysis (Hanks 2004), and the vector spaces framework (Mikolov et al. 2013); all of which have laid the analytical foundation for measuring definition features. Concurrently, comparative studies on definitions have drawn upon quantitative linguistic methods to analyze definition features. For example, Luk and Chan (1995) applied several quantitative methods to analyze dictionary definitions, including word length distribution, frequency, and lexical currency. Recently, studies have employed quantitative methods to examine definition features from traditional dictionaries and AI tools. For instance, Almeman et al. (2024) utilized automated parameters to quantitatively analyze the examples from WordNet, the *Oxford Dictionary*, and ChatGPT. Arias-Arias et al. (2024) employed several natural language processing text similarity algorithms, such as Jaccard, Dice, and ROUGE-N, to conduct a comparative analysis of AI-generated definitions. Their findings revealed that definitions generated by AI exhibit a lack of stability and diverge significantly from those in authoritative dictionaries. Crucially, regarding the dynamic and non-deterministic nature of generated definitions, these quantitative methods provide insightful statistical references for comparative analysis of AI outputs. Meanwhile, the quantitative linguistic methods in quantitative stylistics and stylometry literary studies provide mature theoretical and methodological frameworks for the quantitative analysis of definition texts, offering valuable methodological insights for examining generated definitions.

However, while quantitative methods have been applied to examining dictionary definitions and AI-generated ones, many aspects remain fragmented. Beyond textual similarity, definitions can be explored at the lexical, semantic, and syntactic levels. For instance, word-frequency statistics at the lexical level can reveal differences in word choice between the two sources; semantic feature analysis can reveal underlying differences in their definition strategies; and syntactic complexity metrics (such as dependency distances) can provide critical data regarding how accessible the definitions are. Crucially, such quantitative methods are suited for accommodating the non-deterministic and dynamic generation characteristic of AI outputs. Nevertheless, current comparative research on AI-generated definitions lacks a comprehensive analysis of these metrics.

To address this gap, this study adopts a multidimensional quantitative framework to compare definitions from the *Modern Chinese Dictionary* (7th edition, *MCD*), a traditional monolingual general-purpose dictionary, with their counterparts from AI across lexical, textual, semantic, and syntactic levels. This study first details the study design, then analyzes the empirical findings and their theoretical implications for dictionary compilation. Finally, it concludes the main findings with the limitations.

https://lexikos.journals.ac.za; https://doi.org/10.5788/36-1-2197 (Article)

2. The study design

2.1 Research questions

Based on the aforementioned research background, this study poses the following two research questions:

RQ1: What are the distinctions in the linguistic features between *MCD* definitions and AI-generated definitions across lexical, textual, semantic, and syntactic dimensions?

RQ2: What implications do these distinctions hold for dictionary definition compilation?

2.2 Data collecting and processing

2.2.1 Data collecting

In this study, the mathematical terms in *MCD* are selected as the subjects and provided as the control group, while the corresponding definitions generated by DeepSeek-R1 Pro are utilized as the experimental group for comparative analysis.

The selection of the dictionary, entries, and AI model discussed in this study is based on the following rationale. As a foundational reference for contemporary Chinese general-purpose dictionaries, *MCD* is highly representative in Chinese lexicography, and its treatment of definitions provides reference value for compilation practice. Additionally, the treatment of specialized terms in general-purpose dictionaries is one of the intractable challenges. As Yu and Tan mentioned, the elements of specialized terms in a general-purpose dictionary operate not as a binary "yes-or-no" decision, but rather a "more-or-less" continuum of representational adequacy (Yu and Tan 2021). The mathematical terms in *MCD* demonstrate significant representativeness due to the pedagogical gradient distribution, as the math subject progressively appears at different educational stages for the target users of *MCD*. Furthermore, *MCD* incorporates the mathematical terms standardized in the textbooks (Li 1993), reflecting a systematic structure. To ensure both data balance and structural consistency during processing, this study selects mathematical terms as the study subject. The DeepSeek-R1 Pro model is selected as the AI-generated source, as it serves as a generation engine predominantly trained on Chinese corpora. Based on this, the detailed process is as follows:

(1) Mathematical lemmas and their definitions in *MCD* underwent rigorous manual selection. Since specialized terms in the 7th edition of *MCD* are not marked with explicit subject labels, identifying the mathematical terms in it poses a challenge. To address the implicit subject labeling in it, we adopted a multi-source and cross-verification approach for selection. First, the trial edition of *MCD*, wherein terms are explicitly labeled, was referenced. Also, adhering to the inclusion principles for textbook terms within *MCD* (see Li 1993), references were

https://lexikos.journals.ac.za; https://doi.org/10.5788/36-1-2197 (Article)

made to the official Mathematics Curriculum Standards for primary and secondary education in China. Furthermore, to address the evolution of the term inclusion in the 7th edition, *Termonline*[1], the official platform operated by the China National Committee for Terms in Sciences and Technologies and aggregating standardized terms across disciplines authorized by the state, was adopted as the authoritative reference. In addition, given that general-purpose dictionaries encompass terminological variations (common forms, translated names, etc.) of mathematical terms that are embedded in some of the definitions and cross-references, these related terms were manually screened in *MCD* and also included. Consequently, a total of 228 mathematics-related terms were selected; the complete list of them, along with their English translation, has been compiled into the Appendix.

(2) DeepSeek-R1 Pro is configured as the AI model, adhering to the principle of "one lemma per dialogue session" to avoid contextual interference. The prompt is set in a consistent format in simplified Chinese: "现编写中型单语语文词典中数学学科相关词条的释义，请写出'XX'的定义。" (Please write the definition of 'XX' for mathematical terms in a medium-sized monolingual general-purpose dictionary.) The generated definitions for all mathematical terms in this study were produced by DeepSeek-R1 Pro during the operational window of 23–28 April 2025.

(3) An original experimental database was established to implement a paired sample design, where each term matched definitions from the dictionary definition ("def_mcd") and the AI-generated definition ("def_ai"). The profile of the experimental database is specified in Table 1.

Table 1: Experimental database profile

Attribute of database	Specification
Sampling source	*MCD* and DeepSeek
Target domain	Mathematics
Data size	228 mathematical terms 456 definitions (228 from *MCD*, 228 from AI)

Data structure & storage	Stored in JSON format. Each entry is structured hierarchically as follows:
	{
	"term": "XXX",
	"definition": {
	"def_mcd": "XXX",
	"def_ai": "XXX"
	}
	}

https://lexikos.journals.ac.za; https://doi.org/10.5788/36-1-2197 (Article)

2.2.2 Data processing

Before data analysis, preprocessing the raw text is essential to optimize computational performance and ensure data consistency. The preprocessing stage encompasses data cleaning, word segmentation, and stop word filtering.

(1) Data cleaning. In this stage, the processing of definitions, which mostly consist of Chinese characters but also incorporate non-Chinese elements such as punctuation marks, Arabic numerals, formulas, charts, Latin, and Greek letters, is implemented through Python (version 3.13.5). It is essential to employ different strategies to ensure text consistency. The treatment of irrelevant and special elements is crucial for ensuring the accuracy and standardization of subsequent calculations and analyses.

(2) Word segmentation and stop word filtering. In this stage, we first implemented the "precision mode" of Jieba (version 0.42.1) through Python (version 3.13.5) for initial word segmentation. Subsequently, a stop word list was applied to filter out high-frequency functional words. The filtration adheres to a conservative strategy, with the aim of retaining key terms in domain-specific lexicons. This approach ensures that the analysis remains focused on mathematical terms, enhancing the feature relevance of the present study.

2.3 Multidimensional quantitative framework

To comprehensively analyze *MCD* and AI-generated definitions across various aspects, this study frames its multidimensional quantitative analysis. Instead of relying on a single analytical dimension, this framework integrates selective quantitative linguistic indicators to examine the similarities and differences in definitions from the lexical, textual, and syntactic dimensions. Given that traditional lexicography and automated language models employ divergent modes of text generation and organizational logic, this framework also incorporates semantic features analysis in definitions to explore the definition strategies. Consequently, this section systematically introduces the specific methods across four dimensions of the framework: lexical analysis of definitions, textual similarity analysis, semantic features analysis, and syntactic complexity analysis.

2.3.1 Lexical analysis

After the data processing, a word frequency analysis was first performed, and the term frequency-inverse document frequency (TF-IDF) algorithm was employed to quantitatively assess the significance and differences of keywords in their definitions. The formula is shown below for calculating the weight of a word t within a document d. As the paired-sample corpus was established for analysis, this study sets each document to correspond to the definitions from different sources for a single lemma. Therefore, in the formula, $f_{t,d}$ indicates the frequency of a word t in an individual definition, which also refers to a document d; N represents the

https://lexikos.journals.ac.za; https://doi.org/10.5788/36-1-2197 (Article)

total number of definition documents, with N = 456 (derived from 228 entries plus 2 sources); and n_t signifies the number of documents containing word t.[2]

$$\mathrm{TF-IDF}(t,d) \;=\; \log_{10}(1 + f_{t,d}) \times \left[\log_{10}\left(\frac{N}{n_t + 1}\right) + 1\right]$$

2.3.2 Textual similarity analysis

Building upon the aforementioned word frequency analysis, a further investigation was conducted into the overall divergence between the two types of definitions at the textual level. Based on the vectorized results from TF-IDF weights at the lexical level, the cosine similarity approach was employed to calculate the degree of similarity for each paired definition. Specifically, the definition for each lemma was represented as a vector in a high-dimensional space, with its components constituted by the respective TF-IDF weights. The similarity between two definitions was then assessed by measuring the cosine of the angle between their corresponding vectors, using the following formula:

$$Similarity \;=\; \cos(\theta) \;=\; \frac{\mathrm{V_{dic}} \cdot \mathrm{V_{ai}}}{\|\mathrm{V_{dic}}\| \times \|\mathrm{V_{ai}}\|}$$

2.3.3 Semantic feature analysis

However, an exclusive focus on the lexical and textual levels is insufficient for capturing how these two differ in the definition strategies. Hence, this section expands to include a semantic level by annotating the definition mode of each definition according to the types of semantic features manifested in the specific content. For instance, a mathematical term's definition might include semantic features such as genus, formula, and example, as illustrated by the term "几何体" (geometric solid). Each annotated definition thus yields a sequence of labels, with all occurrences retained, and the resultant data on how semantic features are distributed enables a comparative analysis between dictionaries and AI.

【几何体】空间的有限部分[定义]，由平面和曲面围成[性质]。如棱柱体、正方体、圆柱体、球体[实例]。也叫立体[语言信息]。（《现汉》）([geometric solid] A finite portion of space[definition], bounded by planes and curved surfaces [property]; e.g., prisms, cubes, cylinders, and spheres[example]. It is also referred to as a solid [linguistic information]. *MCD* 7th edition)

2.3.4 Syntactic complexity analysis

In addition to lexical differences, the definitions in *MCD* and AI also exhibit distinct syntactic characteristics, the complexity of which influences text comprehensibility. In quantitative linguistics, key indicators for analyzing syntactic complexity include definition length, mean dependency distance, and mean tree depth. This analysis, therefore, compares these three indicators to assess syn-

https://lexikos.journals.ac.za; https://doi.org/10.5788/36-1-2197 (Article)

tactic complexity, thereby providing insights into the differential structures of defining terms.

3. Results

3.1 Lexical analysis

A comparative word frequency analysis was first conducted on the definitions from *MCD* and AI, with the representation of the respective high-frequency words in *MCD* and AI presented in Table 2.

Table 2: High-frequency words (Top15) in *MCD* and AI

Source	Words*(Gloss)
MCD	如(such as), 数(number), 叫作(be named), 表示(denote), 平面(flat), 运算(operation), 叫(be called), 直线(straight line), 相等(equal), 数学(math), 另(another), 这个(this), 量(magnitude/size), 围成(surround), 都(both/all)
AI	数学(math), 如(such ass), 表示(denote), 运算(operation), 关系(relation), 数值(numerical value), 例如(for example), 平面(flat), 指(refer to), 即(namely), 通过(through), 符号(sign), 构成(constitute), 直线(straight line), 几何(geometry)

The word frequency analysis further highlights the definition styles between the two types of sources. The high-frequency words in *MCD* definitions can be categorized into several categories. The first category relates to linguistic information that refers to direct naming or defining of the headword, such as "叫作" (be named), "表示" (denote). The second category comprises fundamental mathematical objects, such as "数" (number), "量" (magnitude/size), and "平面" (flat). The third one consists of terms associated with mathematical operations, such as "运算" (operation), "围成" (surround). The fourth one includes logical connectives, such as "如" (to be like) to introduce examples, "另" (another) for introducing additional information, "这个" (this) and "都" (all of which) serve specification or cohesive functions. In general, the lexical style of *MCD* favors a concise, direct defining and nomenclature orientation approach. The definition content in *MCD* centers more on the essential denotation of the headword.

In contrast, the high-frequency words in AI-generated definitions reveal a distinct function orientation. One prominent category acts as a disciplinary identifier, with the term "数学" (mathematics) exhibiting a high frequency. Another major category is devoted to linking expressions, featuring in "表示" (denote), "指" (refer to), "即" (namely), and "例如" (for example). A third category can be described as terms that explain mathematical content, including "运算" (opera-

https://lexikos.journals.ac.za; https://doi.org/10.5788/36-1-2197 (Article)

tion), "数值" (numerical value), "平面" (flat), and "符号" (sign), etc. Yet, unlike the definition words in *MCD*, AI-generated definitions within this category place more emphasis on abstract relationships and structural composition in math, as seen in words like "关系" (relation), "通过" (through), and "构成" (constitute). In summary, the lexical choices in AI-generated definitions suggest a content strategy that is explanatory and associative, employing more diverse words in definitions. This indicates that AI not only focuses on the core concepts but also on the interrelation with other concepts, even other knowledge elements.

Overall, the high-frequency words indicate distinct characteristics: *MCD* is predominantly definitional, whereas AI-generated ones tend to be more explanatory.

3.2 Textual similarity analysis

At the textual level, before we begin the similarity analysis, the average word tokens and average character count for both *MCD* and AI were first calculated. The result suggests that definitions in *MCD* were notably shorter (M_{word_token} = 12.70, $M_{character}$ = 34.83) than their AI-generated counterparts (M_{word_token} = 52.35, $M_{character}$ = 151.17). Given this substantial disparity in length, the similarity analysis is computed for each definition pair based on their TF-IDF vectorized results. This calculation reveals a mean similarity score of 0.28, demonstrating a low similarity between these two sources. Based on the similarity calculation, the top ten headwords with the highest similarity and the top ten with the lowest similarity can be identified, as presented in Table 3.

Table 3: Top 10 / Bottom 10 lemmas in textual similarity

Group	Headwords(translation)
High similarity	象限(quadrant), 排列(permutation), 相反数(additive inverse), 棱台(frustum), 通分(reduction of fraction to a common denominator), 平行四边形(parallelogram), 余数(remainder), 钝角(obtuse angle), 倍数(multiple), 不定方程(Diophantine equation)
Low similarity	比率(ratio), 棱角(edge angle), 畸零(literally "odd amount"), 几何(geometry), 维度(dimension), 方程式(equation), 代数(algebra), 横断面(cross-section), 平均(average), 代数方程(algebraic equation)

The clusters of headwords reveal that the group with high similarity exhibits the following characteristics: the headwords consist largely of monosemic and foundational terms with low contextual dependence. Their definitions tend to be more precise and exhibit clear conceptual boundaries, showing fewer poly-

https://lexikos.journals.ac.za; https://doi.org/10.5788/36-1-2197 (Article)

semous variations within or outside the discipline. Since their core definitions are precise and standardized, both *MCD* and AI converge in the definition content. Regarding the term position and relationship within the discipline, although the high-similarity group spans various subfields such as geometry and number theory, most of them are concrete concepts, representing either specific entities (e.g., "棱台", frustum) or operations (e.g., "通分", reduction of fraction to a common denominator). Also, they represent fundamental concepts in mathematics learning, which are included in the China Compulsory Education Mathematics Curriculum Standards (2022). Furthermore, according to the stability and low context dependence of definitions, the AI model is able to produce outputs that align more closely with dictionary definitions.

However, the low similarity group is characterized by the following aspects: polysemy in senses, conceptual generalization, high abstraction of concepts, and divergent presentation formats. The polysemy in senses or conceptual generalization refers to the headwords that carry multiple senses, such as "维度" (dimension), "棱角" (edge angle). *MCD* and AI represent distinct paradigms of definition processing. *MCD*, constrained by compilation principles and storage space, provides definitions that are highly normative, accurate, and relatively stable, thus presenting a selective treatment of polysemous senses. In contrast, AI provides contextualized explanations tailored to the different contexts in which a headword might appear, such as definitions from various subdisciplines (e.g., "代数方程", algebraic equation). Furthermore, compared to the concrete concepts in the high similarity group, this group contains a greater proportion of abstract concepts, including sub-branch names (e.g., "几何", geometry, as a branch of mathematics). For these, AI generates more expansive and informative definitions, thereby reducing similarity. The presentational forms are the third factor, which manifests in two scenarios. The first involves internal cross-reference entries within dictionaries, where detailed definitions are reserved for main entries (e.g., equation "方程式" as the main entry and "方程" as the reference entry). AI-generated definitions, being produced in an isolation mission, lack this systematic internal reference, leading to lower similarity. The second scenario stems from how concisely definitions are formulated. This divergence is particularly evident in fundamental concepts, such as "比率" (ratio) and "平均" (average); *MCD* prioritizes concise and core definitions, whereas AI tends to provide comprehensive explanations by distinguishing related concepts and offering specific examples, focusing more on informative details.

In conclusion, the textual similarity analysis reveals clear findings. The monosemous, concrete, and foundational terms exhibit greater similarity between *MCD* and AI, with their definitions showing stronger convergence in the core features. Conversely, polysemous and abstract terms, or those exhibiting conceptual generalizations, demonstrate more divergence. Within this low similarity group, definitions in *MCD* remain highly standardized and stable, while AI definitions tend to be descriptive.

https://lexikos.journals.ac.za; https://doi.org/10.5788/36-1-2197 (Article)

3.3 Semantic feature analysis

Beyond the quantitative analysis of lexical and textual analysis, it is essential to conduct a statistical analysis of the semantic features of the definitions to identify the different definition strategies of the two sources. All definitions are manually annotated for semantic features, yielding 14 label categories: *genus*, *definition* (core), *property*, *composition* (including subclasses and components), *context* (including applications across various subfields), *function* (covering the purpose and utility of the concept or term), *example*, *relation* (including connections and distinction with other concepts), *symbol*, *formula*, *nomenclature*, *linguistic information* (including common names, aliases), *historical development*, and *illustration*. Based on the frequency count of each label within the definitions, a comparative distribution chart of semantic features can be observed (Figure 1).

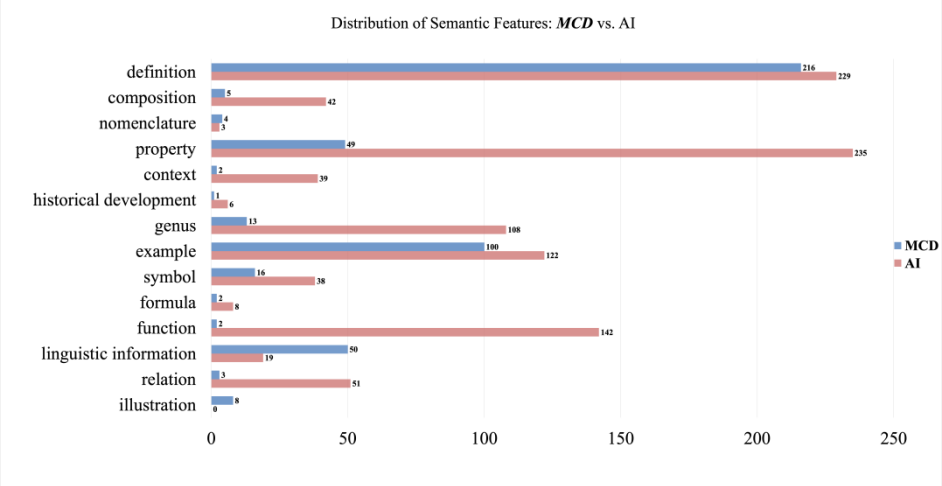

Figure 1: Distribution of semantic features (*MCD* vs. AI)

The comparative figures show a systematic difference in the distribution of semantic features in definitions. *MCD* exhibits its focus on *definition*, with other features serving as supplements, among which the features *example*, *linguistic information*, and *property* are the most prevalent. This reflects a traditional prototype for normative definition, which combines the "*genus* + *differentia*" pattern with examples. Also, the focus of *MCD* is a linguistic resource that further explains the significant weight given to *linguistic information*. Notably, *illustration* accounts for a certain proportion in *MCD*, indicating conscious choices in the presentational strategy in dictionary definitions.

In comparison, AI-generated definitions present a more balanced distribution of the semantic features, emphasizing the functional paradigm. Although *definition*

https://lexikos.journals.ac.za; https://doi.org/10.5788/36-1-2197 (Article)

persists as the focal point in AI-generated definitions, features such as *property*, *function*, *example*, and *genus* also constitute high proportions. This implies that AI-generated definitions favor an extension pattern, emphasizing the *function* and *property* of the headwords. Two notable features are worth noting: one is the *relation*, and the other is the *illustration*. The former shows that AI definitions stress the connections and distinctions with other concepts. This pattern may reflect a broader objective of AI to express knowledge representation beyond simple definitions. In the *illustration* analysis, in the present tested AI model, the *illustration* is an item requiring human intervention and must be explicitly stated in the prompts; otherwise, the illustrations are not inherently included in the outputs.

In summary, the semantic feature analysis shows that *MCD* definitions tend towards normative characteristics, prioritizing systemic organization, whereas AI-generated definitions are more data-driven, emphasizing practical application and functional knowledge representations.

3.4 Syntactic complexity analysis

Having examined the lexical and textual differences between *MCD* and AI-generated definitions, the preceding analysis has revealed differences at both formal and semantic levels. This section proceeds to a quantitative examination of their syntactic characteristics. Since Chinese syntax lacks extensive morphological variations, we employ three parameters to investigate the syntactic complexity of these two types of definitions: the clause (sentence) number, dependency distance, and syntactic tree depth. The findings are as follows:

(1) Clause number of definitions

Considering the overall definitions, we take the punctuation mark "。" as the statistical indicator for counting a complete sentence. A total of 1471 sentences were counted, consisting of 409 sentences from *MCD* and as many as 1062 sentences from AI-generated definitions. The distribution of all entries across *MCD* and AI is presented in Table 4 and Figure 2.

Table 4: Sentence count statistics for *MCD* and AI-generated definitions

Group	Sample Size (*n*)	Total	*M* ± *SD*	Max	Min	Median	Mode
MCD	228	409	1.79±0.92	7	1	2	1
AI	228	1062	4.66±1.51	9	1[3]	4	4

https://lexikos.journals.ac.za; https://doi.org/10.5788/36-1-2197 (Article)

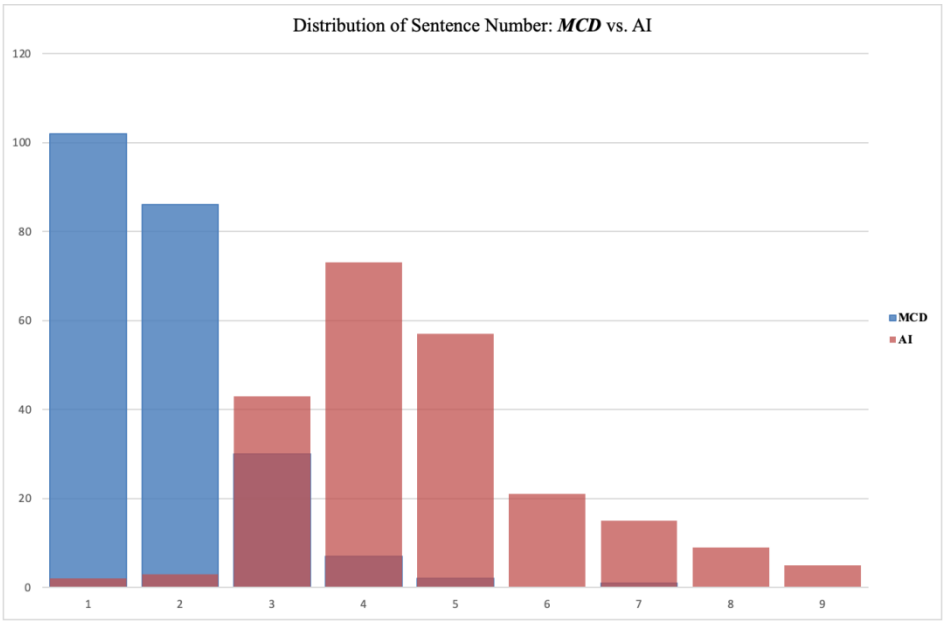

Figure 2: Distribution of sentence number: *MCD* vs. AI

The depicted chart clearly illustrates that most *MCD* definitions consist of 1 to 2 sentences, whereas AI-generated definitions are more commonly composed of 4 to 5 sentences. The Kolmogorov-Smirnov normality test indicates that neither group follows a normal distribution: *MCD*: $D(228) = .25$, $p < .001$; AI: $D(228) = .20$, $p < .001$, and both exhibit somewhat skewed distributions. This implies that a substantial proportion of entries consist of a relatively low number of sentences, with a few extreme outliers. Also, the result of the Wilcoxon signed-rank test, $Z = 12.67$, $p < .001$, suggests a significant difference in the sentence count between *MCD* and AI.

(2) Mean dependency distance

Dependency distance is formally defined as the linear distance between two syntactically related words within a sentence (the governor and the dependent). The dependency distance calculation can be used to measure the complexity of sentence parsing and comprehension (Liu 2022), where a shorter average distance typically correlates with simpler syntactic structures and reduced cognitive load for readers (Liu and Lin 2018). This section conducts a dependency distance analysis from both *MCD* and AI to figure out the differences, and the results are presented in Figure 3.

https://lexikos.journals.ac.za; https://doi.org/10.5788/36-1-2197 (Article)

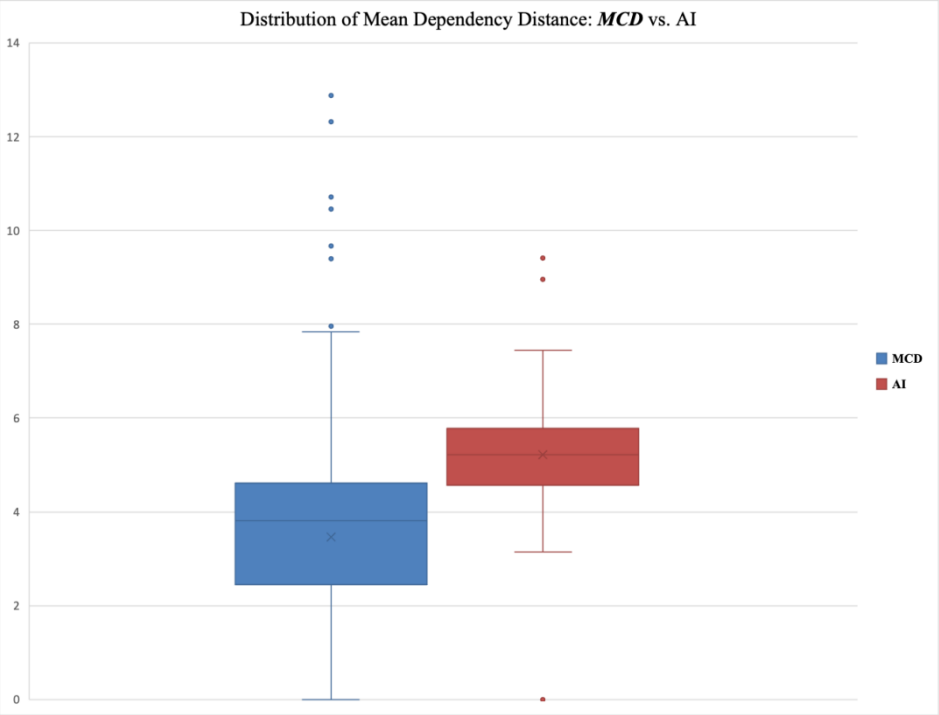

Figure 3: Distribution of mean dependency distance: *MCD* vs. AI

The results indicate that the mean and standard deviation of dependency distance for *MCD* are 3.47 ± 2.43, while for AI-generated definitions are 5.22 ± 1.06. On the whole, the syntactic structure of dictionary definitions is more concise compared to AI-generated definitions, but the complexity of syntactic structures in AI-generated definitions shows less variation across entries. According to the boxplot, both types highlight left skew, with the outliers in *MCD* being more dispersed. The results of Wilcoxon signed-rank test show a significant difference between the two, $Z = 9.40$, $p < .001$, indicating that *MCD* definitions tend to be concise with shorter modifiers. In contrast, AI definitions exhibit a stronger tendency toward structural construction with more modifications.

(3) Mean syntactic tree depth

By comparing the syntactic tree depths of definitions from *MCD* and AI, we can observe differences in the degree of embedding, recursion, and complexity of logical and modification relations in sentences. The mean syntactic tree depth of *MCD* and AI is shown below (Figure 4).

https://lexikos.journals.ac.za; https://doi.org/10.5788/36-1-2197 (Article)

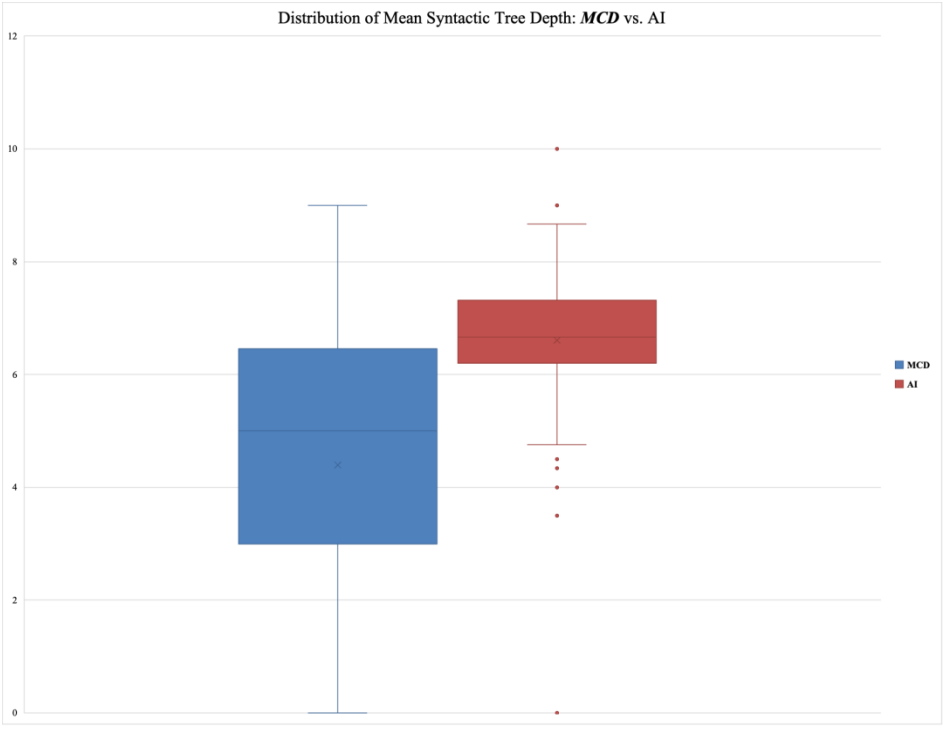

Figure 4: Distribution of mean syntactic tree depth: *MCD* vs. AI

We find that there is a significant difference, $Z = -9.93$, $p < .001$ (Wilcoxon signed-rank test), in the mean syntactic tree depth between *MCD* definitions ($M = 4.39$, $SD = 2.72$) and AI-generated ones ($M = 6.60$, $SD = 1.07$). This suggests that AI-generated definitions generally exhibit greater syntactic tree depth and more embedded structures, accompanied by lower internal variability. As shown in Figure 4, the distribution for *MCD* is left-skewed, indicating that a few of the entries with substantial depth influenced the mean depth. In contrast, AI-generated definitions are left-skewed and contain several outliers representing simpler structures. These contrasting distribution patterns reflect divergent structural representations in definitions. While *MCD* employs a syntactically concise approach tailored for efficient lookup and comprehension, AI-generated definitions are characterized by higher clausal embedding and increased modification density. Crucially, such metric variances indicate distinct text structuring orientations of syntactic organization.

Overall, the aforementioned three syntactic parameters demonstrate significant differences between *MCD* and AI, and AI-generated definitions are higher than *MCD* in three aspects, revealing that AI outputs are relatively more complex at the syntactic level. Furthermore, examining their internal distribution,

https://lexikos.journals.ac.za; https://doi.org/10.5788/36-1-2197 (Article)

AI-generated definitions show limited variation across individual entries, whereas *MCD* definitions are constrained by compilation conditions such as internal reference and entry complexity, which produces greater syntactic diversity within *MCD*. However, these three indicators of syntactic complexity cannot fully measure the informational load; combining them with semantic features would allow a further analysis of their implications for users.

4. Discussion and suggestions

4.1 Cognitive load of definitions for users

Text reception, text production, and translation are the three core communicative situations engaged by dictionary users (Bergenholtz and Tarp 2003). The information conveyed in a definition and its syntactic complexity directly influence user comprehension. Excessively dense information or overly complex syntax can diminish readability and increase the user's cognitive load. Within the comparative framework of the four-level analysis, the syntactic dimension, in particular, demands attention regarding the cognitive load of traditional versus AI-generated definitions. For instance, the length of definitions (sentence number) affects comprehension time, dependency distance impacts the working memory load (Liu et al. 2017), and syntax tree depth influences the user's efficiency of structural integration.

Upon the results, on the one hand, AI-generated definitions have higher averages in sentence count, dependency distance, and syntactic tree depth. *MCD*, with its lower averages in dependency distance and tree depth, along with fewer sentence and word counts, implies that it conveys information through more succinct syntactic structures, thus imposing a relatively lower working memory load for users. On the other hand, when combined with semantic feature analysis, AI-generated definitions suggest a greater informational content. However, their complex syntactic structures mean that users may need to process deeply embedded constructions, which could potentially exceed their working memory capacity. Nonetheless, the superior efficiency of *MCD* does not inherently facilitate efficient information acquisition by users. Considering both syntactic and semantic analysis, several definitions in *MCD* feature notably long dependency distances, particularly in entries with lengthy modifiers, such as the definitions for "十进制" (decimal system) and "阿拉伯数字" (Arabic numerals), which use constructs like "[example A, example B] of [a superordinate concept]". These extended modifiers frequently serve to illustrate the extensional forms through examples. Conversely, AI-generated definitions often provide detailed examples related to *context*, *function*, and *example*, which reflects a syntactic consequence of incorporating more semantic features. Overall, while the lower syntactic complexity of dictionary definitions may facilitate quicker comprehension for the user, the limited content and simpler syntactic structures can lead to high

https://lexikos.journals.ac.za; https://doi.org/10.5788/36-1-2197 (Article)

information density or missing details. Consequently, users might spend more time compensating for these deficiencies by seeking additional information, thereby reducing the efficiency of information retrieval. AI-generated definitions show a marked increase in syntactic complexity, which can heighten the challenge of understanding them for users. However, their more detailed explanations, contextual information, and logical connections may reduce the cognitive effort required during comprehension.

Hence, instead of relying solely on data analysis to make assumptions about their treatment of cognitive load in definitions, a critical distinction must be made between the intrinsic cognitive load imposed by the conceptual difficulty of a term itself and the extrinsic cognitive load generated by the manner of information presentation. The distinct structural patterns manifested by these two definition types offer meaningful insights for the future compilation of general-purpose dictionaries. Traditional dictionaries, especially in printed ones, exhibit text structures optimized for informational compression. When evaluating definition strategies for terms, compilers should carefully balance the intrinsic conceptual difficulty of a term against its decoding efficiency at the syntactic level. The relationship between these two variables is synergistic and complementary. For highly specialized terms, definitions should be organized using relatively simpler syntactic structures to lower the users' cognitive load in linguistic decoding. Conversely, the high syntactic complexity of AI-generated definitions warrants caution. If basic terms are presented through overly complex structures, it not only inflates the users' working memory load but also leads to informational redundancy and wastes the dictionary volume. Consequently, this issue remains intertwined with the knowledge representation construction and presentation in the dictionary.

4.2 Knowledge representation construction and presentation of definitions

Term definitions play a crucial role in the representation of specialized knowledge. It is regarded that a term definition constitutes its minimal knowledge representation (Faber and Sánchez 2001). The definition from *MCD* and AI can also be viewed as knowledge representations of terms, though they differ in construction and presentation. Collectively, regarding similarity, the definitional knowledge representation for fundamental and monosemous terms tends to be similar, whereas abstract and polysemous ones exhibit significant differences. In terms of the lexical style perspective, *MCD* definitions present a tendency toward concise, systematic, constructed, and static definitions, while AI-generated definitions emphasize descriptions about relations within concepts of terms. Therefore, these characteristics are also reflected in the semantic features of definitions. Echoing the characteristics in the lexical style perspective, *MCD* definitions align more closely with classical category theory, frequently employing a "genus + differentia" structure; AI-generated definitions are more in line with prototype category theory, presenting a distinctive "attribute + function" expan-

https://lexikos.journals.ac.za; https://doi.org/10.5788/36-1-2197 (Article)

sion structure. This divergence stems from their fundamentally different logics in organizing knowledge representation. The compilation of traditional dictionary definitions is a systematic endeavor, aiming for authoritative and essential definitions (Landau 2001: 63). AI-generated definitions, however, rely on the probabilistic generation mechanism of large language models (Brown et al. 2020), with training data (such as encyclopedias and textbooks) favoring function-oriented explanatory characteristics.

Despite these differences in how knowledge representations are constructed and presented, and the technical backing available to AI-generated definitions, traditional dictionary compilation can draw valuable insights from AI, particularly in the construction of internal knowledge systems for definition strategies. Within the framework of Frame-Based Terminology (FBT), a specialized term can be embedded within a domain-specific event frame and activate a series of conceptual relationships through its constituent elements. Typically, the constituent elements of a term include its concepts, relationships, attributes, and context (Faber 2012). Traditional dictionary definitions, constrained by space, focus on depicting the core concept. AI, to some extent, accommodates various constituent elements of a term, which in turn can activate different parts of its conceptual frame. As lexicography undergoes a digital and intelligent transformation that enables the storage and application of vastly larger datasets (Klosa-Kückelhaus 2024), it is imperative to reassess the traditional organizational frameworks of lexicographical data. Crucially, we need to consider whether the current presentation of the knowledge representation of term definitions in traditional dictionaries meets users' needs. As noted by Yu and Du (2025), in general-purpose dictionaries, terms often require explanations rather than rigid direct definitions; the former implies more accessible, user-friendly explanations tailored to users' cognitive load and communicative context, whereas the latter reflects systematic definitions rooted in terminology; in *MCD*, a substantial portion of terms remains toward the latter. We suggest that intelligent lexicography requires a stronger focus on knowledge representations within general-purpose dictionaries. Moreover, the optimization of *MCD* definitions could restructure and encode terminological knowledge of terms, thereby better accommodating users' multifaceted demands for cognitive acquisition and linguistic application.

4.3 Reflections on the dictionary compilation

The comparative analysis of *MCD* definitions and AI-generated definitions compels a critical reflection on the opportunities and challenges faced by traditional lexicography amid the rapid development of AI.

From the analysis, it becomes clear that AI-generated definitions have certain shortcomings. The quality of these definitions is constrained by the limitations of their training data. Consequently, AI models frequently fail to capture linguistic information, misidentify specific terms (e.g., the diachronic terminological variants

https://lexikos.journals.ac.za; https://doi.org/10.5788/36-1-2197 (Article)

"奇零" and "畸零", both literally "odd amount"), and reduce definition reliability. These limitations underscore that employing AI for generating terminological definitions still necessitates expert intervention and rigorous verification. Furthermore, although AI-generated definitions are generally more detailed, the efficiency of processing such content requires further evaluation. While traditional dictionaries facilitate rapid consultation, AI supports deeper, open-ended exploration. With the advancement of intelligent dictionaries, future general-purpose dictionaries should incorporate user profiles to dynamically tailor content, ensuring that the definition strategies of terms are appropriately adapted to the comprehension levels of diverse user groups.

As Lew (2024: 7) notes, in the future, lexicographers will be curators and analysts of linguistic data, guiding AI to better understand and process less-represented languages and complex linguistic phenomena. AI serves as an auxiliary tool in the lexicographic process, the output of which requires careful selection, supervision, and verification by human compilers, particularly when processing terminological units within specialized knowledge domains. Crucially, in the future practice of AI-generated or AI-assisted lexicography, relying solely on post-hoc manual comparison is insufficient due to the highly dynamic and probabilistic nature of AI outputs. Instead, we propose an integration practice in AI-generated lexicography: utilizing existing lexicographic data as structured inputs for language models, as the lexicographical data serve as the gold standard in the planning and compilation of lexicographic products (Bothma and Gouws 2025: 293) and represent a sustainable and economic product (Fuertes-Olivera 2024); meanwhile, the multidimensional quantitative metrics in this study could offer a systematic framework to evaluate the outputs, which be systematically adapted to other definition tasks.

Nonetheless, we should caution against overreliance on the quantitative data alone; these metrics should serve as comparative references rather than absolute qualitative judgments. As Tarp (2025) discusses, artificial intelligence can assist in dictionary compilation and enhance efficiency, yet it can never replace the deep thinking of humans. Thus, despite differing technical backups, the emergence of human-AI synergy is a prevailing trend. The development of AI is not the end for lexicography but a transformative shift; it prompts us to reflect on not only the format, compilation methods, and presentation of dictionaries, but also the assessment method of these definitions. Traditional assessment relying on human supervision is inadequate to match the characteristics of AI-generated or AI-assisted definitions, which implies that it necessitates an integration of data assessment and human supervision for deeper evaluation.

5. Conclusion

This study employs a multidimensional quantitative framework to systematically compare *MCD* definitions and AI-generated definitions across four dimensions: lexical difference in definitions, text similarity, semantic features, and syn-

https://lexikos.journals.ac.za; https://doi.org/10.5788/36-1-2197 (Article)

tactic complexity. In terms of lexical differences, *MCD* definitions tend to be more concise in essential definitions, whereas AI-generated definitions lean towards explanatory content. Regarding textual similarity, we find that the average similarity between the two is relatively low (M = 0.28). Entry clustering based on definitional similarity reveals distinct patterns across lexical categories. High similarity tends to cluster around monosemous and concrete terms representing basic concepts. Conversely, lower similarity is more frequently observed among polysemous, abstract, and generalized conceptual terms. In semantic feature distribution, *MCD* definitions center on the *definition* feature, using others as supplements, thus emphasizing the essential definition. AI-generated definitions display a wider variety of semantic features, demonstrating a functional inclination. As for syntactic complexity, AI-generated definitions surpass *MCD* definitions in the number of defining sentences, average dependency distance, and average syntactic tree depth, indicating greater syntactic complexity. However, a comparison of internal syntactic complexity reveals distinct patterns. While the structural variance within AI-generated definitions remains insignificant, *MCD* entries display substantial differences. This contrast suggests that traditional dictionaries involve tailored structural organizations for different word types, whereas AI outputs tend toward a relatively uniform syntactic complexity. The analysis across these four dimensions further explores the distinct manifestations of user cognitive load of the two definition sources. We propose that treatment of terms in general-purpose dictionaries should distinguish between intrinsic and extrinsic cognitive load of definitions. We should adopt differentiated definition strategies for the cognitive difficulty of various terms. Furthermore, the semantic feature organization analysis between *MCD* and AI definitions points to different patterns of knowledge representation construction and presentation. This empirical variance suggests that future optimization of term definitions in general-purpose dictionaries should look beyond static text generation to consider how knowledge can be more dynamically structured to better accommodate the multifaceted user needs in the intelligent era. Based on this multidimensional quantitative framework, we propose that this approach can serve as a systematic evaluation for AI-generated or AI-assisted definitions, which also necessitates human collaboration for future lexicographical practices.

While this study has constructed a four-dimensional framework to analyze the linguistic features of *MCD* and AI-generated definitions, the study still has its limitations. First, although the categorization of semantic features was grounded in definitional content, the annotation process inherently involves a degree of subjectivity. Future research could incorporate automated identification and annotation to enhance accuracy and objectivity. Second, despite exploring systematic differences between traditional dictionary definitions and AI-generated ones through quantitative metrics, the study falls short of assessing the actual impact of these differences on user comprehension through empirical validation. In subsequent research, we will further incorporate user feedback to achieve more practical implications.

https://lexikos.journals.ac.za; https://doi.org/10.5788/36-1-2197 (Article)

Acknowledgements

The authors would like to express their sincere gratitude to Prof. Elsabé Taljard and the two anonymous adjudicators for their insightful comments and suggestions. Special thanks are also due to Prof. Pingfang Yu from the Faculty of Chinese Language and Culture, Guangdong University of Foreign Studies, and Prof. Suen Caesar Lun from the Department of Linguistics and Translation, City University of Hong Kong, for their valuable suggestions on earlier drafts of this study.

Endnotes

1. https://termonline.cn/
2. Specifically, the interpretations for the three instances of "+1" in the formula are as follows: The primary purpose of the "+1" in $\log_{10}(1 + f_{t,d})$ is to avoid the mathematically undefined condition of $\log_{10}(0)$, thereby ensuring the continuity and validity of the calculation. The "+1" in $\log_{10}\left(\frac{N}{n_t + 1}\right)$ is mainly introduced to prevent division by zero. The "+1" at the end of the formula ensures that the entire IDF term does not drop to zero when a specific term is extremely common, thereby preserving the fundamental semantic contribution of that term.
3. There are extreme values (terms "畸零" and "奇零" were not recognized by the AI) that are unsuitable for statistical analysis; after their removal, the minimum value is 1.

References

A. Dictionaries

Institute of Linguistics of the Chinese Academy of Social Sciences. 2016. *Xiandai hanyu cidian* [Modern Chinese Dictionary]. Seventh edition. Beijing: The Commercial Press.

B. Other literature

Albalawi, A., S. Alzeer, N. Alharbi, N. Alsharidi, N. Alqahtani and M. Asiri. 2026. Comparing the Effectiveness of ChatGPT and Dictionaries in Second Language Vocabulary Learning. *International Journal of Lexicography* 39: ecag014.

Almeman, F.Y., S. Schockaert and L. Espinosa Anke. 2024. WordNet under Scrutiny: Dictionary Examples in the Era of Large Language Models. Calzolari, N., M.-Y. Kan, V. Hoste, A. Lenci, S. Sakti and N. Xue (Eds.). 2024. *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024), May 2024, Turin, Italy*: 17683-17695. Turin: ELRA and ICCL.

Arias-Arias, I., M.J. Domínguez Vázquez and C. Valcárcel Riveiro. 2024. Der Effizienz- und Intelligenzbegriff in der Lexikographie und künstlichen Intelligenz: kann ChatGPT die lexikographische Textsorte nachbilden? *Lexikos* 34: 51-76.

https://lexikos.journals.ac.za; https://doi.org/10.5788/36-1-2197 (Article)

Bergenholtz, H. and S. Tarp. 2003. Two Opposing Theories. On H.E. Wiegand's Recent Discovery of Lexicographic Functions. *HERMES — Journal of Language and Communication in Business* 31: 171-196.

Bothma, T.J. and R.H. Gouws. 2025. Generative Artificial Intelligence (GenAI) as Information Tool for Lexicographic Information Needs. *Lexikos* 35(2): 276-295.

Brown, T.B., B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D.M. Ziegler, J. Wu, C. Winter, … and D. Amodei. 2020. Language Models are Few-shot Learners. *Advances in Neural Information Processing Systems* 33: 1877-1901.

Chen, X., Y. Pan and Y. Yang. 2025. A Preliminary Study of the Large Language Model-assisted Compilation of Learner-oriented Chinese Dictionaries. *Journal of Yunnan Normal University (Teaching and Researching on Chinese as a Foreign Language)* 23(2): 11-20.

Chen, Y. 2025. A Comparative Study on the Effectiveness of AI Chatbots and Dictionary Apps for Lexical Tasks and Retention. *Lexikos* 35: 157-182.

Church, K.W. and P. Hanks. 1990. Word Association Norms, Mutual Information, and Lexicography. *Computational Linguistics* 16(1): 22-29.

De Schryver, G.-M. 2023. Generative AI and Lexicography: The Current State of the Art Using ChatGPT. *International Journal of Lexicography* 36(4): 355-387.

De Schryver, G.-M. 2025a. Out-of-the-Box GPTs for Lexicography. *Lexikos* 35(2): 362-396.

De Schryver, G.-M. 2025b. Customised GPTs for Lexicography. *Lexikos* 35(2): 397-422.

De Schryver, G.-M. and D. Joffe. 2023. The End of Lexicography, Welcome to the Machine: On How ChatGPT Can Already Take over All of the Dictionary Maker's Tasks. *20th CODH Seminar, Center for Open Data in the Humanities, Research Organization of Information and Systems, National Institute of Informatics, Tokyo, Japan, 27 February 2023.* http://codh.rois.ac.jp/seminar/lexicography-chatgpt-20230227/

Faber, P. (Ed.). 2012. *A Cognitive Linguistics View of Terminology and Specialized Language.* Berlin: De Gruyter Mouton.

Faber, P. and M. Tercedor Sánchez. 2001. Codifying Conceptual Information in Descriptive Terminology Management. *Meta* 46(1): 192-204.

Fuertes-Olivera, P.A. 2024. Making Lexicography Sustainable: Using ChatGPT and Reusing Data for Lexicographic Purposes. *Lexikos* 34: 123-140.

Han, W. and Gilles-Maurice de Schryver. 2026. Perceptions of Language Service Technologies and Reference Works by Chinese Users in the Era of Generative AI. *International Journal of Lexicography* 39: ecag002.

Hanks, P. 2004. Corpus Pattern Analysis. Williams, G. and S. Vessier (Eds.). 2004. *Proceedings of the 11th EURALEX International Congress*, EURALEX 2004, *Lorient, France, 6–10 July 2004*: 87-97. Lorient: Faculté des lettres et des sciences humaines, Université de Bretagne-Sud.

Huang, J. 1987. *Studies in Lexicography*. Shanghai: Shanghai Lexicographical Publishing House.

Huete-García, Á. and S. Tarp. 2024. Training an AI-based Writing Assistant for Spanish Learners: The Usefulness of Chatbots and the Indispensability of Human-assisted Intelligence. *Lexikos* 34: 21-40.

Jakubíček, M. and M. Rundell. 2023. The End of Lexicography? Can ChatGPT Outperform Current Tools for Post-editing Lexicography? Medveď, Marek, Michal Měchura, Carole Tiberius, Iztok Kosem, Jelena Kallas, Miloš Jakubíček and S. Krek (Eds.). 2023. *Electronic Lexicography in*

https://lexikos.journals.ac.za; https://doi.org/10.5788/36-1-2197 (Article)

the 21st Century (eLex 2023): Invisible Lexicography. Proceedings of the eLex 2023 Conference, Brno, 27–29 June 2023: 518-533. Brno: Lexical Computing CZ s.r.o.

Kilgarriff, A., V. Baisa, J. Bušta, M. Jakubíček, V. Kovář, J. Michelfeit, P. Rychlý and V. Suchomel. 2014. The Sketch Engine: Ten Years On. *Lexicography* 1(1): 7-36.

Kilgarriff, A., P. Rychlý, P. Smrz and D. Tugwell. 2004. The Sketch Engine. Williams, G. and S. Vessier (Eds.). 2004. *Proceedings of the 11th EURALEX International Congress, EURALEX 2004, Lorient, France, 6–10 July 2004*: 105-115. Lorient: Faculté des lettres et des sciences humaines, Université de Bretagne-Sud.

Klosa-Kückelhaus, A. 2024. *Internet Lexicography: An Introduction*. Berlin: De Gruyter.

Klosa-Kückelhaus, A. and C. Tiberius. 2025. The Lexicographic Process Revisited. *International Journal of Lexicography* 38(1): 1-12.

Landau, S.I. 2001. *Dictionaries: The Art and Craft of Lexicography*. Second edition. Cambridge: Cambridge University Press.

Lew, R. 2023. ChatGPT as a COBUILD Lexicographer. *Humanities and Social Sciences Communications* 10(1): 1-10.

Lew, R. 2024. Dictionaries and Lexicography in the AI Era. *Humanities and Social Sciences Communications* 11(1): 426.

Li, Q. and S. Tarp. 2024. Using Generative AI to Provide High-Quality Lexicographic Assistance to Chinese Learners of English. *Lexikos* 34: 397-418.

Li, Z. 1993. Lun *Xiandai hanyu cidian* de baike tiaomu [On Specialized Entries in *Modern Chinese Dictionary*]. *Lexicographic Studies* 5: 37-45.

Liu, C. 2025. An Experimental Study and Reflections on Automated Lexicographic Compilation Using Large Language Models. *Chinese Journal of Language Policy and Planning* 10(2): 61-72.

Liu, H. 2022. *Dependency Relation and Language Networks*. Beijing: China Science Publishing and Media.

Liu, H. and Y. Lin. 2018. Methodology and Trends of Linguistic Research in the Era of Big Data. *Journal of Xinjiang Normal University (Philosophy and Social Science)* 39(1): 72-83.

Liu, H., C. Xu and J. Liang. 2017. Dependency Distance: A New Perspective on Syntactic Patterns in Natural Languages. *Physics of Life Reviews* 21: 171-193.

Liu, R., X. Chen and Y. Xu. 2025**.** Beyond Replacement: How Large Language Models Influence Dictionary Usage Patterns Among Chinese English Learners. *International Journal of Lexicography* 38(4): 342-364.

Luk, R.W.P. and V.M.K. Chan. 1995. A Quantitative Analysis of Word-definition in a Machine-readable Dictionary. T'sou, B.K. and T.B.Y. Lai. 1995. *Proceedings of the 10th Pacific Asia Conference on Language, Information and Computation, Hong Kong, 27–29 December 1995*: 133-138. Hong Kong: City University of Hong Kong.

Mandour, W. 2024. Empowering Bilingual Lexicography with AI: The Role of AI-Model Triangulation Approach (AMTA) in Dictionary Design. Inoue, A., N. Kawamoto and M. Sumiyoshi. 2024. *Proceedings of the 17th International Conference of the Asian Association for Lexicography (AsiaLex 2024 Proceedings): Asian Lexicography — Merging Cutting-edge and Established Approaches, 12–14 September 2024, Tokyo, Japan: Asian Lexicography-Merging Cutting-Edge and Established Approaches*: 70-83. Tokyo: Toyo University.

Mikolov, T., I. Sutskever, K. Chen, G.S. Corrado and J. Dean. 2013. Distributed Representations of Words and Phrases and their Compositionality. *Advances in Neural Information Processing Systems* 26: 3111-3119.

https://lexikos.journals.ac.za; https://doi.org/10.5788/36-1-2197 (Article)

Miller, G.A. 1995. WordNet: A Lexical Database for English. *Communications of the ACM* 38(11): 39-41.

Ptasznik, B. and R. Lew. 2025. Dictionaries versus AI Tools through the Eyes of English Majors. *International Journal of Lexicography* 38(2): 140-158.

Ptasznik, B., S. Wolfer and R. Lew. 2024. A Learners' Dictionary Versus ChatGPT in Receptive and Productive Lexical Tasks. *International Journal of Lexicography* 37(3): 322-336.

Rees, G.P. and R. Lew. 2024. The Effectiveness of OpenAI GPT-generated Definitions Versus Definitions from an English Learners' Dictionary in a Lexically Orientated Reading Task. *International Journal of Lexicography* 37(1): 50-74.

Rundell, M. 2023. Automating the Creation of Dictionaries: Are We Nearly There? *Proceedings of the 16th International Conference of the Asian Association for Lexicography (AsiaLex 2023 Proceedings), 22–24 June 2023, Seoul, Korea: Lexicography, Artificial Intelligence, and Dictionary Users:* 9-17. Seoul: Yonsei University.

Rychlý, P. 2008. A Lexicographer-Friendly Association Score. Sojka, P. and A. Horák (Eds.). 2008. *Proceedings of Recent Advances in Slavonic Natural Language Processing, RASLAN 2008. 2. Vyd, 5–7 December 2008:* 6-9. Brno: Masaryk University.

Sinclair, J.M. 1991. *Corpus, Concordance, Collocation.* Oxford: Oxford University Press.

Tarp, S. (Translated by Jia, H. and Y. Ji). 2025. Towards a Symbiosis of Human and Digital Intelligence: A Lexicographic Perspective. *Lexicographical Studies* 4: 1-15, 125.

Tarp, S. and R.H. Gouws. 2023. A Necessary Redefinition of Lexicography in the Digital Age: Glossography, Dictionography and Implications for the Future. *Lexikos* 33: 425-447.

Yu, P. and J. Du. 2025. *Tongyong yuwencidian zhong shuyu chuli de duowei yanjiu* [A Multidimensional Study of Specialized Term Treatment in General-purpose Dictionaries]. Beijing: The Commercial Press.

Yu, P. and J. Tan. 2021. *A Comparative Study on the Microstructures of Chinese and English Monolingual General-purpose Dictionaries.* Beijing: Science Press.

Zhang, Y. 2024. The Theoretical Model of Multimodal Dictionaries and their Definitions within the Framework of Media Convergence. *Foreign Language Teaching and Research (bimonthly)* 56(4): 519-531.

Zhang, Y. and T. Liu. 2024. A Survey on Automatic Compilation of Chinese Dictionary Entries: A Case Study of ChatGPT. *Lexicographical Studies* 5: 20-33.

Zheng, X. and X. Chen. 2026. Evaluating Generative AI as a Bilingual Dictionary: Performance and Prompt Optimization in Idiom Retrieval. *Lexikos* 36: 127-144.

Zhong, A. 2025. Prompts for Language Learners: A Practical Guide to Using DeepSeek as a Dictionary. *Lexikos* 35(1): 274-285.

https://lexikos.journals.ac.za; https://doi.org/10.5788/36-1-2197 (Article)

Appendix 1: Mathematical lemmas in *MCD* with their English translations

阿拉伯数字 (Arabic numerals)
百 (hundred)
百分比 (percentage)
百分点 (percentage point)
百分号 (percent sign)
百分率 (percent rate)
百分数 (percent)
半圆 (semicircle)
倍式 (multiple [of a polynomial])
倍数 (multiple [of an integer])
比 (ratio)
比率 (ratio)
比例 (proportion)
比值 (ratio value)
边 (side)
扁率 (oblateness)
变量 (variable)
变数 (variable)
补角 (supplementary angle)
不等号 (inequality sign)
不等式 (inequality)
不定方程 (indeterminate equation)
不名数 (unnamed number)
参变量 (parameter variable)
参数 (parameter)
差 (difference)
长方体 (cuboid)
常量 (constant)
常数 (constant)
常用对数 (common logarithm)
乘 (multiply)
乘除 (multiplication and division)
乘法 (multiplication)
乘方 (exponentiation)
乘积 (product)
乘幂 (power)
除 (divide)
除法 (division)
代数 (algebra)
代数方程 (algebraic equation)
代数式 (algebraic expression)
代数学 (algebra)
等号 (equal sign)
等式 (equation)
等于 (equal to)
底数 (base)
对数 (logarithm)
钝角 (obtuse angle)
多边形 (polygon)
多面角 (polyhedral angle)
多面体 (polyhedron)
二面角 (dihedral angle)
反比 (inverse ratio)
反比例 (inverse proportion)
反函数 (inverse function)
方 (square)
方程 (equation)
方程式 (equation)
方根 (root)
分 (fraction)
分式 (fractional expression)
分数 (fraction)
负号 (negative sign)
负数 (negative number)
负增长 (negative growth)
概率 (probability)
高 (height)
高度 (height)
根号 (radical sign)
根式 (radical expression)
公比 (common ratio)
公差 (common difference)
公理 (axiom)
公式 (formula)
函数 (function)
和 (sum)

https://lexikos.journals.ac.za; https://doi.org/10.5788/36-1-2197 (Article)

恒等式 (identity equation)
横断面 (cross-section)
横剖面 (cross-section)
横坐标 (abscissa)
弧度 (radian)
黄金分割 (golden ratio)
几率 (probability)
奇零 (literally "odd amount")
畸零 (literally "odd amount")
奇数 (odd number)
基期 (base period)
集合 (set)
几何 (geometry)
几何体 (geometric solid)
几何图形 (geometric figure)
几何学 (geometry)
计算 (calculation)
加法 (addition)
减法 (subtraction)
交点 (intersection point)
角 (angle)
阶乘 (factorial)
解 (solution)
解析几何 (analytic geometry)
近似值 (approximate value)
弳 (radian)
绝对值 (absolute value)
括号 (parentheses)
累进 (progressive)
棱角 (edge angle)
棱柱 (prism)
棱台 (frustum of a prism)
棱锥 (pyramid)
立方 (cube)
立方体 (cube)
立体 (solid)
立体几何 (solid geometry)
立体角 (solid angle)
立体图 (three-dimensional diagram)
零数 (odd amount)
面 (face)
逆定理 (converse theorem)
逆运算 (inverse operation)
排列 (permutation)
抛物线 (parabola)
配方 (completing the square)
平方 (square)
平角 (straight angle)
平均 (average)
平面 (plane)
平面几何 (plane geometry)
平面角 (plane angle)
平行 (parallel)
平行四边形 (parallelogram)
平行线 (parallel lines)
千分点 (permillage point)
千分号 (permille sign)
千分数 (permille)
切 (tangent)
切点 (point of tangency)
切线 (tangent line)
曲率 (curvature)
曲线 (curve)
三角函数 (trigonometric function)
三角形 (triangle)
三角学 (trigonometry)
十进对数 (common logarithm)
十进制 (decimal system)
数 (number)
数表 (numerical table)
数理逻辑 (mathematical logic)
数量级 (order of magnitude)
数列 (sequence)
数论 (number theory)
数位 (digit)
数学 (mathematics)
数值 (numerical value)
数制 (numeral system)
数轴 (number line)
四边形 (quadrilateral)
四舍五入 (rounding)
四维空间 (four-dimensional space)

https://lexikos.journals.ac.za; https://doi.org/10.5788/36-1-2197 (Article)

素数 (prime number)
速算 (quick calculation)
算术 (arithmetic)
算术根 (arithmetic root)
算学 (arithmetic)
梯形 (trapezoid)
通分 (common denominator)
统计学 (statistics)
椭圆 (ellipse)
椭圆体 (ellipsoid)
拓扑学 (topology)
微分 (differential)
微积分 (calculus)
维 (dimension)
维度 (dimension)
尾数 (mantissa)
未知数 (unknown number)
位 (place)
无理式 (irrational expression)
无理数 (irrational number)
无穷大 (infinity)
无穷小 (infinitesimal)
系数 (coefficient)
相反数 (opposite number)
相交 (intersect)
相似形 (similar figure)
象限 (quadrant)
小数 (decimal)
小数点 (decimal point)
斜率 (slope)
验算 (verification)
杨辉三角 (Pascal's triangle)
因式 (factor)
因式分解 (factorization)
因子 (factor)
有理式 (rational expression)
有理数 (rational number)
余数 (remainder)
圆 (circle)
圆台 (frustum of a cone)
圆心 (center of a circle)
圆心角 (central angle)
圆周 (circumference)
圆周角 (inscribed angle)
圆周率 (pi)
圆柱 (cylinder)
圆锥 (cone)
约分 (reduction of a fraction)
约数 (divisor)
运算 (operation)
整式 (integral expression)
整数 (integer)
正比 (direct ratio)
正比例 (direct proportion)
正多边形 (regular polygon)
正方体 (cube)
正方形 (square)
正弦 (sine)
直径 (diameter)
直角 (right angle)
直线 (straight line)
指数 (exponent)
质数 (prime number)
自变量 (independent variable)
自乘 (self-multiplication)
自然对数 (natural logarithm)
自然数 (natural number)
组合 (combination)
祖率 (Zu's ratio/pi)
坐标 (coordinate)