UDC: 10.00.00 – Philological Sciences

INTERACTIVE AI-BASED DICTIONARIES: NEW HORIZONS IN LEXICOGRAPHY AND LANGUAGE LEARNING

Khalilova Maftuna Azamat qizi
Master's student of Shakhrisabz State Pedagogical Institute
xalilovamaftuna53@gmail.com

ABSTRACT

This article explores the extensive capabilities of artificial intelligence (AI)-based interactive dictionaries in enhancing vocabulary learning and lexicography. It discusses various AI applications, including chatbots, gamification, virtual and augmented reality, corpus linguistics, and social media platforms, which collectively contribute to a more engaging and effective language acquisition process. The integration of AI technologies enables real-time interaction, contextual understanding, pronunciation assistance, and error correction, thus offering personalized learning experiences.

Key words: artificial intelligence, interactive dictionaries, lexicography, vocabulary learning, chatbots, gamification, virtual reality, augmented reality, corpus linguistics, language technology, pronunciation assistance, language education

In the digital age, dictionaries are no longer static books on a shelf. The integration of artificial intelligence (AI) into lexicographic tools gives rise to *interactive AI-based dictionaries*—systems that not only provide definitions but engage users through context, adaptation, and interactivity. These systems open up novel possibilities for vocabulary acquisition, error correction, and user engagement in language learning, particularly in mother tongue (L1) education and second language settings. This article explores the capabilities, challenges, and future prospects of AI-driven interactive dictionaries, comparing them with traditional lexicographic resources and outlining directions for further development

1. Capabilities and Features of AI-Driven Interactive Dictionaries

Below are key capabilities and how they transform the learning and reference experience.

1.1 Conversational Agents and AI Helpers: Modern chatbots and AI language assistants (e.g. ChatGPT, Duolingo's conversational modules) can hold dialogues with learners, exposing them to new lexical items in an interactive manner. These agents can correct mistakes, suggest synonyms, and provide usage examples dynamically.

1.2 Mobile Apps and Flashcard Systems: Applications such as Quizlet and Anki

implement **spaced repetition systems (SRS)** which schedule reviews of words at optimal intervals, improving long-term retention. When enhanced with AI, these systems can adapt to a learner's error pattern, difficulty and context.

1.3 Speech and Voice Technologies: AI assistants like Siri, Google Assistant, or integrated speech modules in apps allow learners to hear correct pronunciations, test their own pronunciation, and query word meanings vocally. This supports phonetic learning and listening comprehension.

1.4 Gamification and Motivational Design: Interactive word games, quizzes, and point-reward systems give learners tangible incentives. Gamification helps maintain engagement, and AI can dynamically adjust difficulty to match the learner's progress.

1.5 Virtual Reality / Augmented Reality (VR/AR): Immersive VR or AR tools can place learners in virtual environments where objects are labeled, or vocabulary is embedded in scenes. For example, users might walk in a virtual city and click on objects to learn names and contexts. This contextual immersion solidifies vocabulary.

1.6 Corpus Linguistics and Data Analytics: AI can mine corpora (massive text collections) to detect frequently used words, collocations, idioms, and trends. Learners' written or spoken input can also be analyzed to highlight over- or underused words. Tools such as AntConc or Sketch Engine operate in this domain. In lexicographic research, corpus studies help ensure dictionary entries reflect actual usage.[1]

1.7 Interactive and Social Platforms: Learners can maintain blogs, post micro-essays, or record podcasts using new vocabulary, fostering active use. Social forums (Telegram, language learning communities) allow for peer feedback and real-world application of new lexemes.

1.8 Machine Translation and Correction Tools: AI translation tools (e.g. Google Translate, DeepL) help users understand lexical meaning in multilingual contexts. Writing assistants like Grammarly or LanguageTool detect lexical, grammatical, and stylistic errors and suggest alternative phrasings.

1.9 Real-Time Adaptivity and Updates: AI systems can update their lexicons in real time, ingesting new words, slang, or domain-specific vocabulary. They can also predict user queries (auto-completion), generate example sentences, and tailor suggestions to the user's history and preferences.

1.10 Contextual and Semantic Disambiguation: AI models, especially large language models (LLMs), can infer which meaning of a polysemous word is intended in context. This capability surpasses simple dictionary senses listing by adapting to nuance, idiomatic usage, and metaphorical extensions.

---

[1] De Schryver, Gilles-Maurice. *Generative AI and Lexicography: The Current State of the Art Using ChatGPT*, *International Journal of Lexicography*, 2023. lib.ugent.be+1

1.11 Multi-modal Support: Modern interactive dictionaries support multimedia: definitions accompanied by images, videos, usage animations, or audio. This helps users internalize meaning across sensory modes

2. Strengths, Limitations, and Comparative Perspectives

2.1 Strengths and Advantages

- **Personalization and adaptivity:** AI can tailor learning paths to individual needs, skipping familiar words and focusing on weak areas.

- **Rich contextualization:** Beyond bare definitions, interactive dictionaries provide example contexts, usage notes, and collocations dynamically.

- **Scalability and coverage:** AI models can handle huge vocabularies and rare senses without the labor demands of traditional lexicographers.

- **Real-time feedback:** Users can instantly query, revise, and receive corrections.

- **Continuous updates:** Lexicons remain dynamic, accommodating neologisms and domain-specific terms.

- **Engagement and retention:** Gamified approach and multi-sensory input improve motivation and memory retention.

2.2 Limitations, Risks, and Challenges

- **Bias and unfairness:** Word embedding models may reproduce semantic biases present in source corpora (e.g. gender, ethnicity).[2]

- **Lack of lexical authority:** AI outputs may lack editorial control, leading to errors or invented senses.

- **Data sparsity for minority languages:** AI models often underperform in low-resource languages or dialects, where curated lexicons remain essential.

- **Interpretability issues:** Complex AI reasoning is often opaque; users may not know why certain suggestions are made.

- **Overreliance on AI:** Learners might neglect deeper language study, relying on AI crutches rather than building active competence.

- **Cost and maintenance burden:** High computational and infrastructural demands; also updating models and domain adaptation is expensive.

- **Copyright and licensing constraints:** Integrating proprietary dictionaries or corpora into AI systems faces legal constraints.

- **Quality assurance:** Human lexicographers are still needed to vet AI outputs, maintain consistency, correct mistakes, and manage rare senses.

2.3 Comparative Views from Research

- De Schryver's study *"Generative AI and Lexicography"* critically evaluates how ChatGPT has been applied in dictionary writing, noting trends and limitations in

---

[2] Chayanon Phoodai & Richárd Rikk. "Exploring the Capabilities of ChatGPT for Lexicographical Purposes: A Comparison with Oxford Advanced Learner's Dictionary." *Electronic Lexicography in the 21st Century*. 2023. elex.link

automatically generating lexical entries. Another comparative study contrasts ChatGPT's microstructural lexicographic output to that of the Oxford Advanced Learner's Dictionary, highlighting divergences in lexical items and coverage.[3]

- The theoretical article *"Theoretical Considerations on AI-based Business Models for Lexicography"* discusses possible commercial models and curation strategies for AI lexicon services.
- The paper *"Dictionaries and lexicography in the AI era"* examines how AI and machine translation challenge traditional lexicography but asserts the continuing relevance of curated dictionaries, especially for under-resourced languages.

3. Implications for L1 (Native Language) Teaching and Lexicographic Practice

3.1 Enhancing Mother-Tongue Vocabulary Education

In native language instruction, interactive dictionaries can:

- Expose learners to less common, domain-specific vocabulary (e.g. scientific, literary) in context.
- Provide morphological breakdowns, etymologies, and usage examples tailored to learners.
- Adapt to each learner's pace, focusing on neglected lexical zones.
- Combine with reading materials (e.g. AR labels in textbooks) to create immersive learning.

3.2 Transforming Lexicographic Workflow

Lexicographers must evolve from sole compilers to *curators and trainers* of AI systems. Key roles include:

- Data curation: selecting high-quality corpora, balancing domains, removing noise.
- Prompt engineering: designing effective input prompts to steer AI outputs.
- Quality control: verifying AI-generated definitions, usage notes, and senses.
- Business modeling: determining sustainable models for AI dictionary services (subscription, institutional licensing, freemium, etc.).
- Ethical oversight: monitoring biases, privacy, and transparency in AI decisions.

3.3 Interdisciplinary Collaboration

Successful AI lexicography demands cooperation between linguistics, computational linguistics, machine learning, UX design, and education. For example:

- Linguists define criteria for sense distinctions, collocations, and registers.
- ML engineers optimize models, embeddings, and fine-tuning.
- Educators test pedagogical efficacy in real classrooms.

---

[3] Robert Lew. *Dictionaries and Lexicography in the AI Era*. *Palgrave Communications*, 2024. IDEAS/RePEc+2DOAJ+2

• UX designers ensure intuitive interfaces and feedback loops.

3.4 Localization and Low-Resource Languages

For many languages (e.g. Uzbek, minority languages), interactive AI lexicons can help preserve and develop vocabulary. Yet, low data volume poses a challenge. Hybrid models—combining AI suggestions with expert lexicographer intervention—are essential. The article *"Lexicography, artificial intelligence, and dictionary users"* argues for sensitive design when dealing with user communities.

AI-based interactive dictionaries represent a promising frontier in lexicography and language education. They offer interactivity, personalization, and dynamism far beyond what static print dictionaries can provide. However, they are not replacements for human expertise. Human lexicographers, educators, and language communities remain indispensable for quality control, addressing bias, and tailoring resources to specific linguistic contexts.

In practice, the most effective systems will likely combine AI automation with human oversight—a **hybrid paradigm**. Future work should emphasize:

• Better bias mitigation in word embeddings and context models.

• More robust support for low-resource languages.

• Transparent, explainable AI outputs.

• Empirical studies on learning outcomes in classrooms using interactive dictionaries.

By balancing innovation with rigor, AI-enabled lexicography can enrich language learning, preserve linguistic diversity, and expand the expressive capacity of interaction between users and words.

References:

1. De Schryver, Gilles-Maurice. *Generative AI and Lexicography: The Current State of the Art Using ChatGPT*, *International Journal of Lexicography*, 2023. lib.ugent.be+1

2. Chayanon Phoodai & Richárd Rikk. "Exploring the Capabilities of ChatGPT for Lexicographical Purposes: A Comparison with Oxford Advanced Learner's Dictionary." *Electronic Lexicography in the 21st Century.* 2023. elex.link

3. Robert Lew. *Dictionaries and Lexicography in the AI Era*. *Palgrave Communications*, 2024. IDEAS/RePEc+2DOAJ+2

4. Henrik Køhler Simonsen. *Theoretical Considerations on AI-based Business Models for Lexicography*. *Lexicographica*, 2023. De Gruyter Brill

5. Latipova, Gulasal B. *Corpus Study in Lexicographic Research*. International Journal of Artificial Intelligence, 2024. academicpublishers.org

6. Aylin Caliskan, Joanna J. Bryson, Arvind Narayanan. "Semantics derived automatically from language corpora contain human-like biases." arXiv, 2016. arXiv

7. "Lexicography, Artificial Intelligence, and Dictionary Users." *Lexicography (Journal)*, introduction volume.