Dictionary Analysis: Features and Lemmatisation

Joseph A Villarama*

Central Luzon State University, Philippines. *Corresponding Author's Email: villaramajoseph120294@clsu.edu.ph

Abstract

Language documentation in the Philippines includes comprehensive vocabulary compilations, where dictionaries evolved, leading to modernisation and standardisation. Selected based on institutional authority, lexicographical relevance, and educational significance, this study compares "UP Diksiyonaryong Filipino" (UPDF) and "KWF Diksiyonaryo ng Wikang Filipino" (KWF-DWF), in terms of features, lemmatisation approach, and modern relevance, through a qualitative descriptive content analysis anchored on CULU (coverage, user interface, lemmatisation, and usability) framework, without the utilisation of any external corpora and/or computational lemmatisers. The findings reveal that the UPDF offers descriptive research-based approach and emphasises linguistic diversity, historical depth, and cultural inclusivity with respect to variety whereas the KWF-DWF emphasises standardisation and accessibility, and closely aligns with language policy and educational needs. The UPDF provides thorough entries, including etymology, regional variants, and usage examples while the KWF-DWF has a practical function in education and governance, supporting linguistic consistency and further enabling formalised language education. Both dictionaries struggle to reconcile the traditional lexicographical practice with contemporary computation. Adopting digital innovations, allow dictionaries to both extend their relevance and expand their availability in contemporary lexicographical milieu. It is imperative in these collaborative efforts between linguists, educators, and technologists that both traditional and computational needs are met without sacrificing cultural and linguistic sustainability.

Keywords: Dictionary-Making, Language Documentation, Language Education, Lexicography, Lexicology.

Introduction

The history of language documentation in the Philippines includes comprehensive vocabulary compilations, where dictionaries of Filipino language have gradually evolved, leading to its modernisation and standardisation. Two of the most influential of these works are "UP Diksiyonaryong Filipino" (UPDF) and "KWF Diksiyonaryo ng Wikang Filipino" (KWF-DWF), products of two distinct approaches to thorough depiction of lexical items in vernacular, considering local realities and dedicated to language work. The UPDF, a project developed by the University of the Philippines (UP), focuses on research and linguistic scholarship, to document Filipino as a living language. KWF-DWF, being published by Komisyon sa Wikang Filipino (KWF), is consistent with agency's mandate of promoting and standardising national language, having elements of prescriptivism and inclusivity. With the evolution of Filipino, further relevance of such dictionaries stretch outside of academic setting and into realms of education, media, and even day-to-day communication in times of globalisation. Dictionary served as keystone of lexicography and language learning for many years, playing crucial role in normalising languages, reviving culture; thus, enabling communication (1-3). Lexicographical works played crucial role in the "evolution of academic research" and "language policy" in Filipino language (4). Some scholars also sought to delve into Filipino lexicography while employing socio-political perspective (5-7). Previous works addressed broad topics of language planning and policy, contributing to a dearth of comparative studies on central lexicographical resources such as "UP Diksiyonaryong Filipino" (UPDF) and "KWF Diksiyonaryo ng Wikang Filipino" (KWF-DWF). Despite studies that focus one dictionary, praising their additions towards linguistic research and education, not much research exists that contribute to systematic overview of structural features, lemmatisation strategies, and contemporary importance of these dictionaries (8-9). In the digital age, dictionaries are asked to play a crucial role in this context evolving to adapt to technological innovations and to capture the dynamics of ever-changing worlds (10-12).

Further, there are no empirical studies on the effect of these dictionaries on language education, specifically on language literacy and critical language awareness among Filipino language learners. This study addressed these gaps through examining the structural features of dictionaries, specifically "UP Diksiyonaryong Filipino" (UPDF) and "KWF Diksiyonaryo ng Wikang Filipino" (KWF-DWF) in terms of features, lemmatisation approach, and modern relevance. As this comparative study carries a continuing value for the advancement of discussion about lexicography, the results have important implications for educators, linguists, and policymakers, as they can shed light on the efficacy of these dictionaries in supporting language literacy and developing critical language awareness. Furthermore, this research recognises the changing demands of Filipino language speakers in the digital era, emphasising the necessity of adapting dictionary-making behaviours to modern-day linguistic contexts.

Methodology

A qualitative descriptive research design was used in this study to achieve a close descriptive content analysis and comparison of "UP Diksiyonaryong Filipino" (UPDF) and "KWF Diksiyonaryo ng Wikang Filipino" (KWF-DWF), which derived the interpretative and descriptive aspects of lexicographical practices by looking at the features, lemmatisation, and contextualisation gaps for each dictionary, making qualitative methodology the most suitable for studying the general trends between them; thus, did not use external corpora or computational lemmatisers (13). Focusing on dictionary entries, structural organisation, and the treatment of relevant linguistic concepts, the study systematically compared the similarities and differences to identify areas of comparison. It also viewed these dictionaries through the lenses of the socio-cultural and educational contexts in which they operate.

The bases for selecting "UP Diksiyonaryong Filipino" (UPDF) and "KWF Diksiyonaryo ng Wikang Filipino" (KWF-DWF) for this study were framed by three aspects: "institutional authority,"

"lexicographical relevance," and "educational significance" (14-16). Both dictionaries are products of reputable institutions—University of the Philippines and Komisyon sa Wikang Filipino—so they cannot be simply dismissed as uncredible or counter-productive to national language policies. These dictionaries are indispensable references for linguistic research, language planning, and cultural preservation, as they greatly influence Filipino lexicography. Prevalence of UPDF and KWF-DWF in educational contexts reflects their function in teaching languages and developing literacy.

In scrutinising the UPDF and KWF-DWF, this study used a "multiple-dimensional approach" with respect to coverage, user interface, lemmatisation, and usability (CULU) as shown in Figure 1 (17). The scope of lexical entries was evaluated through coverage, which examined the presence of standard, colloquial, and specialised terms, in addition to cultural references attesting to the dynamic quality of Filipino language. Given the way contemporary users engage with lexicographical resources, user interface was primarily concerned with how well-defined and how accessible the structural organisation of a dictionary is in digital form. Lemmatisation approaches were explored to identify the processes by which varying forms of a word can be treated as instances of a single canonical underlying form, and were tested for consistency, completeness, and linguistic correctness (18). Usability looked at how clickable and accessible each dictionary is, from how easy it is for both an academic and general user to flip around and find information. The analysed dictionaries did not observe the "TEI (Text Encoding Initiative) guidelines," as they were primarily structured for print and digital reference use without explicitly implementing TEI-compliant encoding.

This study was reviewed in accordance with the Ethics Research Committee (ERC) of the Central Luzon State University (CLSU), Philippines, dated January 6, 2025.

Figure 1 covers the scope of lexical entries evaluated through coverage, user's interface, lemmatisation techniques, and usability.
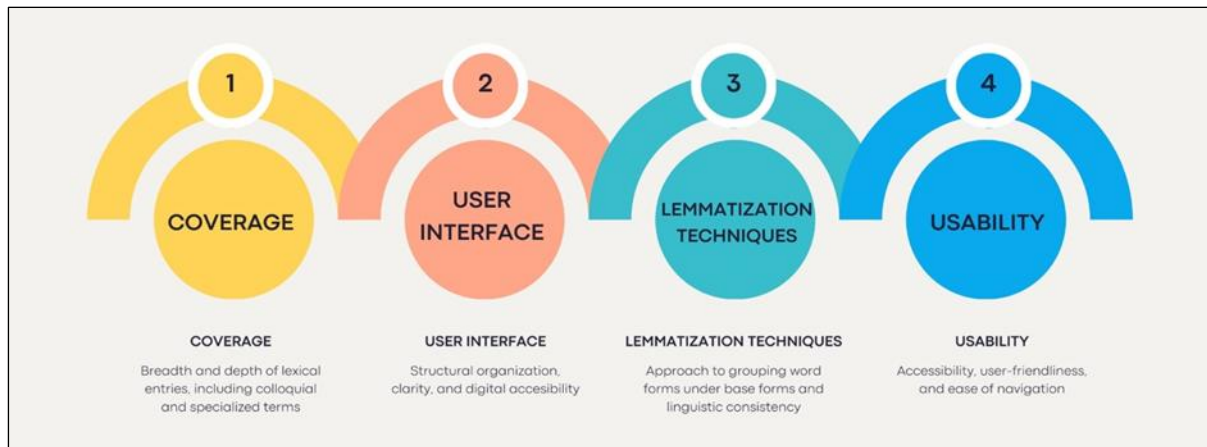
Figure 1: CULU Framework

Results

The "UP Diksiyonaryong Filipino" (UPDF) is an important reference for the study and expansion of language, specifically the Filipino language. It was published by the "Sentro ng Wikang Filipino" of the University of the Philippines (UP) with the aim of developing and enriching the national language while preserving the rich culture of the Philippines. There have been two editions of this, the first in 2001 and the latest in 2010, which was introduced as the "Deluxe Edition." One of its most important features is its multilingual scope because it is not limited to Tagalog words only, but also uses words from various languages and dialects in the country such as Cebuano, Ilokano, Hiligaynon, Kapampangan, Ifugao, and even foreign languages; thus, possesses the characteristic of the Filipino language of being inclusive. This is different from other dictionaries because of its descriptive way of presenting meanings that show the actual use of words in various contexts, where not only the traditional meaning of the word is provided but also its broader usage in everyday life. In some entries, equivalent meanings in other languages in the country are also shown. A notable characteristic of the UPDF is its incorporation of contemporary lexicon, encompassing terminology associated with technology, science, and popular culture. This dictionary notably accommodates diverse spelling variants of words from various areas.
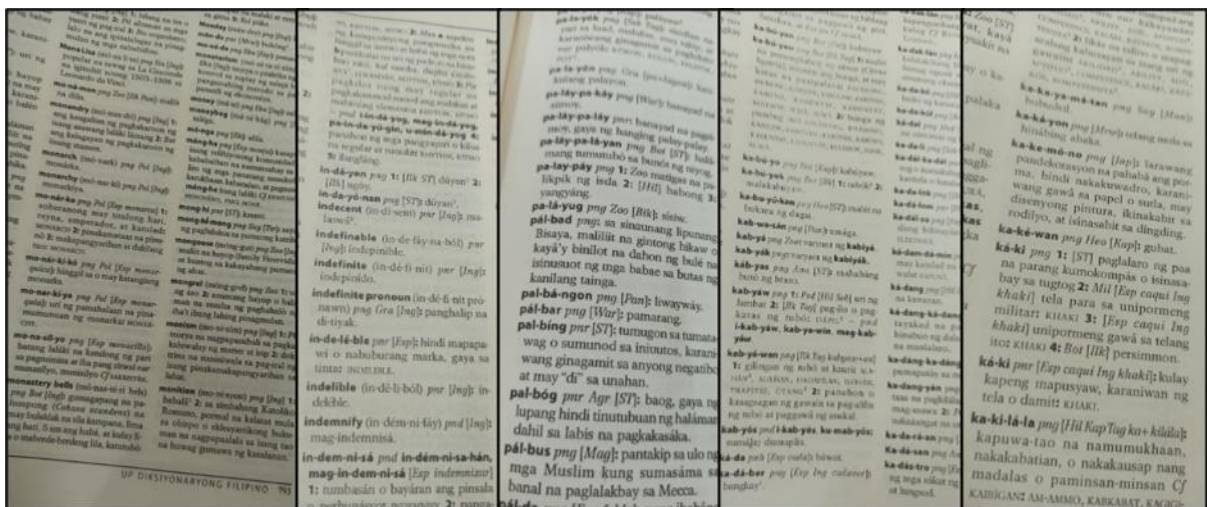

Figure 2: Multilingual UPDF

Figure 2 shows the "KWF Diksiyonaryo ng Wikang Filipino" (KWF-DWF) with a sample of words, their meanings and usage, and pronunciations and variances. In comparing UPDF and "KWF Diksiyonaryo ng Wikang Filipino" (KWF-DWF), their main differences in purpose and scope of language can be seen. The UPDF aims to expand and achieve the intellectualisation of Filipino by incorporating various languages in the Philippines, whereas the KWF-DWF focuses more

on the standard Filipino based on Tagalog, believing it to be the nucleus of Filipino. The first uses a descriptive definition, while the latter follows a prescriptive definition or the establishment of the correct use of words according to the rules set by the Komisyon sa Wikang Filipino (KWF). The UPDF has a broader scope, estimated to have over 100,000 entries and is more open to incorporating modern words, whereas the KWF Dictionary has only about 30,000 words and is more conservative. Meanwhile, the KWF Online Dictionary shows the use of the included words within sentences, correct pronunciation through phonetic transcription, and the correct pronunciation can also be heard.

Table 1: Select Features of UPDF and KWF-DWF

	UP Diksiyonaryong Filipino (UPDF)	KWF Diksiyonaryo ng Wikang Filipino (KWF-DWF)
Scope	Est. 100,000 entries	Est. 30,000 entries
Meaning/Word Senses	Descriptive Multiple definitions, etymologies, contextual examples, and regional variations	Prescriptive Meanings align with official usage and language policy
Morphological Rules	Derived forms (affixed forms) Root forms	Root forms Minimal inclusion of affixed variants
Usage	Detailed-Contextualised Grammatical Information	Standardised/Concise
Part of Speech/Grammar	Word classes Grammatical variations and dialectal influences	Grammatical Categorisation
Spelling	Traditional and Modernised	Orthographical
Pronunciation	Phonetic Transcription	Simplified (With Sounds)

Table 1 presents select features of "UP Diksiyonaryong Filipino" (UPDF) and "KWF Diksiyonaryo ng Wikang Filipino" (KWF-DWF) focusing on meaning, usage, grammar, spelling, and pronunciation. From a semantic perspective, UPDF takes the descriptive, while KWF-DWF takes the prescriptive stance; both corresponding not only to the directives established by the official license of our language but also within the dynamic and progressive characteristic of the language. Definitions tend to be short and contextually based around effective examples of modern usage. Similarly, UPDF and KWF-DWF include both formal and informal meanings in dictionary, embracing the variety of language (19). In terms of usage, UPDF offers in-depth use cases of words in different registers, calling attention to alternative usages, and encourages a deeper understanding of the language while KWF-DWF is more concerned with correct usage with respect to "state language policy", favouring so-called "formal" or "standardised location" that marks an effort to create "linguistic uniformity" across education and public communication (20-21). As for grammar UPDF provides more thorough grammatical details such as word class (e.g., noun, verb, adjective) while KWF-DWF focuses on the most basic grammatical categories, labelling a word as a part of speech. Further, the UPDF uses both standard and new orthography spellings while being aware of variant spellings, particularly those deriving from dialectal variation and historical orthography, which maximises inclusivity. However, KWF-DWF follows orthographic guidelines set forth by the "Komisyon sa Wikang Filipino," ensuring linguistic homogeneity. Furthermore, UPDF supplies phonetic transcriptions for selected entries, particularly for uncommon or difficult words. It also indicates regional pronunciation variants while KWF-DWF uses simplified pronunciation guides and its pronunciation guides focus on the standard pronunciation prescribed based on language policy. The findings reveal that their structures are designed with a traditional lexicographical approach, focusing on semantic depth (UPDF) and standardisation (KWF-DWF). The data structure directly impacts lemmatisation accuracy: UPDF's descriptive and detailed entries with

morphological richness and semantic nuance facilitate comprehensive lemmatisation by recognising various inflected and derived forms. Conversely, KWF-DWF's simpler, prescriptive structure enhances consistency but risks underrepresenting the morphological diversity and context-driven word senses crucial for accurate computational lemmatisation.
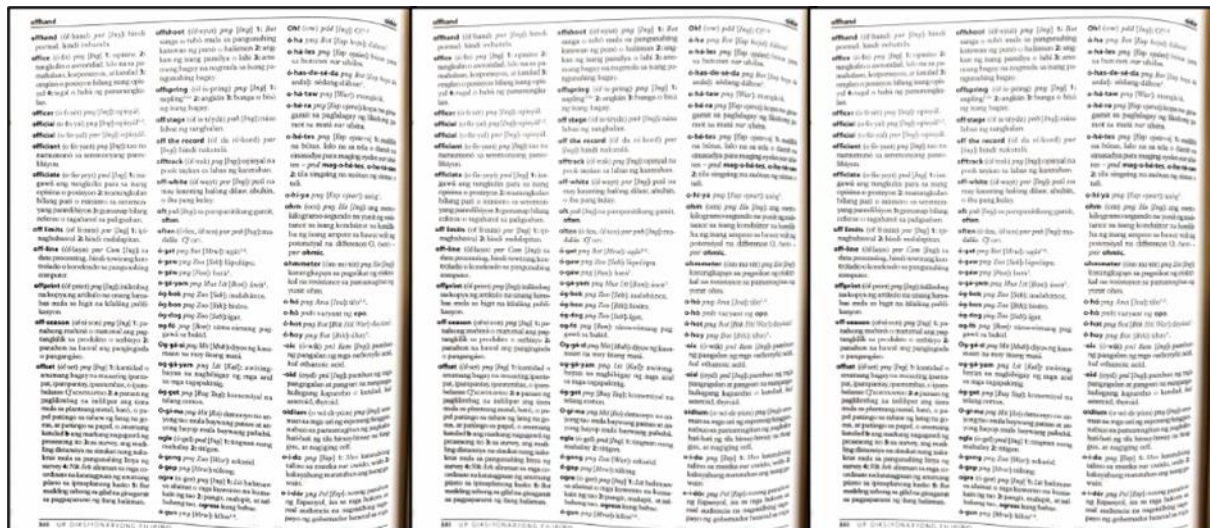

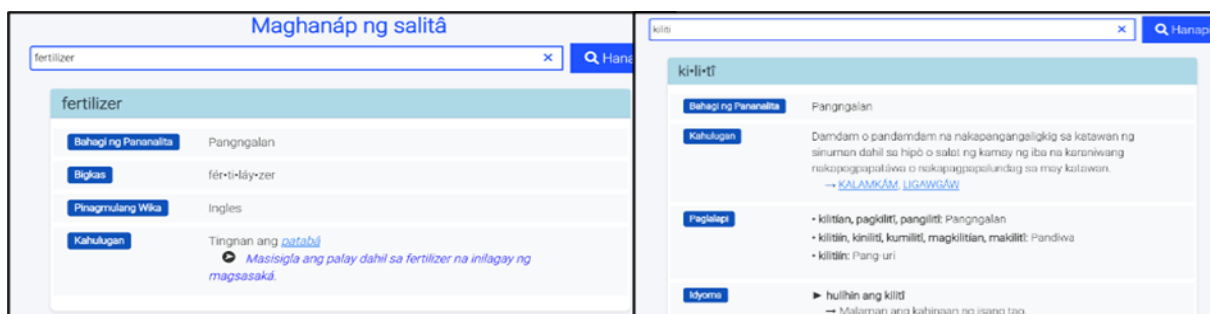**Figure 3:** Select Features of UPDF

Figure 4: Select Features of KWF-DWF

Figure 3 shows the select features of the "UP Diksiyonaryong Filipino" (UPDF) such its root, affix, and other forms. Figure 4 shows the select features of the "KWF Diksiyonaryo ng Wikang Filipino" (KWF-DWF) such its lexical data and other relevant details and forms.

"Lemmatisation"—the act of grouping together different inflected forms of the same word, including its base form (lemma)—are of utmost importance in lexicography, as it determines how users navigate between different meanings, usages and grammatical behaviours of each lexical item (22-24). The lemmatisation of both "UP Diksiyonaryong Filipino" (UPDF) and "KWF Diksiyonaryo ng Wikang Filipino" (KWF-DWF) explores the lexical data, which these dictionaries are based upon, as informed by a philosophy of descriptivism (UPDF) and prescriptivism (KWF-DWF). On one hand, the "UP Diksiyonaryong Filipino" (UPDF), which is both descriptive and linguistically sensitive in nature and reflective of the dynamic nature of the Filipino language sets out root words in the lemma, but often root words are accompanied by derived forms, which have different meanings. For example, root verbs may be listed alongside commonly used inflected or affixed forms, such as "*lakad*" (walk) and "*paglakad*" (walking), allowing users to understand grammatical shifts and semantic nuances. Further, UPDF lemmas typically notify up sampled phrases that illustrate distinctions of meaning and usage between compound and affixed presentation types, encouraging a more intuitive understanding of the functional and grammatical variety within the vocabulary. Regional variations, alternative versions, slang and dialectal influences, even informal usages widen the net of lemmatisation and inclusivity within the UPDF.

Figure 5: Lemmatisation of UPDF

Figure 5 shows the lemmatisation of the "UP Diksiyonaryong Filipino" (UPDF) showing the dynamic nature of the Filipino language with its root words in the lemma, often accompanied by derived forms. On the other hand, the "KWF Diksiyonaryo ng Wikang Filipino" (KWF-DWF) adopts a prescriptive and uniform lemmatisation strategy, fostering linguistic consistency and conformity to national language policy as shown in Figure 6. In KWF-DWF, lemma is listed as the base form or root form of a word, and users are primarily redirected to inflected or derived forms on the basis of standard morphological rules. For instance, "*takbo*" (run) is presented as the main entry, with minimal inclusion of its affixed variants. Only to ensure easier learning in primary and secondary education, KWF-DWF simplifies its lemmatisation, which is comparable to similar studies advocating for uniformity in writing, grammar, and usage based on official rules of language use (25-26). All entries in the dictionary abide by the canonical orthography prescribed by the "Komisyon sa Wikang Filipino." Though it has many usage examples compared to the UPDF, the KWF-DWF is more prescriptive than descriptive, standardising language instruction but may limit users' understanding of semantic flexibility and regional nuances.

Figure 6: Lemmatisation of KWF-DWF

Figure 6 shows the lemmatisation of the "KWF Diksiyonaryo ng Wikang Filipino" (KWF-DWF) showing listed lemma as the base form or root form of a word, and users redirected to inflected or derived forms.

In the present time, dictionaries keep coming up in the classroom and beyond, and they no longer stand on defining words alone, instead, dictionaries also serve as tools for preservation of language and culture, education, and communication in a digital and global society (27-30). The "UP Diksiyonaryong Filipino" (UPDF) and the "KWF Diksiyonaryo ng Wikang Filipino" (KWF-DWF) represent today's significant strides in Philippine lexicography by meeting distinct but equally invaluable needs of Filipino language users, educators, and learners in contemporary times as shown in Figures 5 and 6.

Table 2: Modern Relevance of UPDF and KWF-DWF

	UP Diksiyonaryong Filipino (UPDF)	KWF Diksiyonaryo ng Wikang Filipino (KWF-DWF)
Adaptability to Technology	Print-focused	Digitalised
Preface and Introduction	In-depth	Concise
Guidelines	Structured	Standardised

Table 2 shows the modern relevance of "UP Diksiyonaryong Filipino" (UPDF) and "KWF Diksiyonaryo ng Wikang Filipino" (KWF-DWF) focusing on adaptability to technology, preface and introduction, and other guidelines.

As a government-endorsed lexicographical tool, the "UP Diksiyonaryong Filipino" (UPDF) maintains a print-focused infrastructure as much as attempts have been made to share this lexicographical effort on official government web domains as shown in Table 2. Meanwhile, "KWF Diksiyonaryo ng Wikang Filipino" (KWF-DWF) has shown more adaptability by introducing digital formats as access to online platforms or e-books. This increases user engagement, particularly for students and researchers who opt for digital materials (31-34). Their thorough entries—including etymologies, usage examples and cross-references—are consistent with contemporary mandates for depth in linguistic information in academic as well as informal contexts.
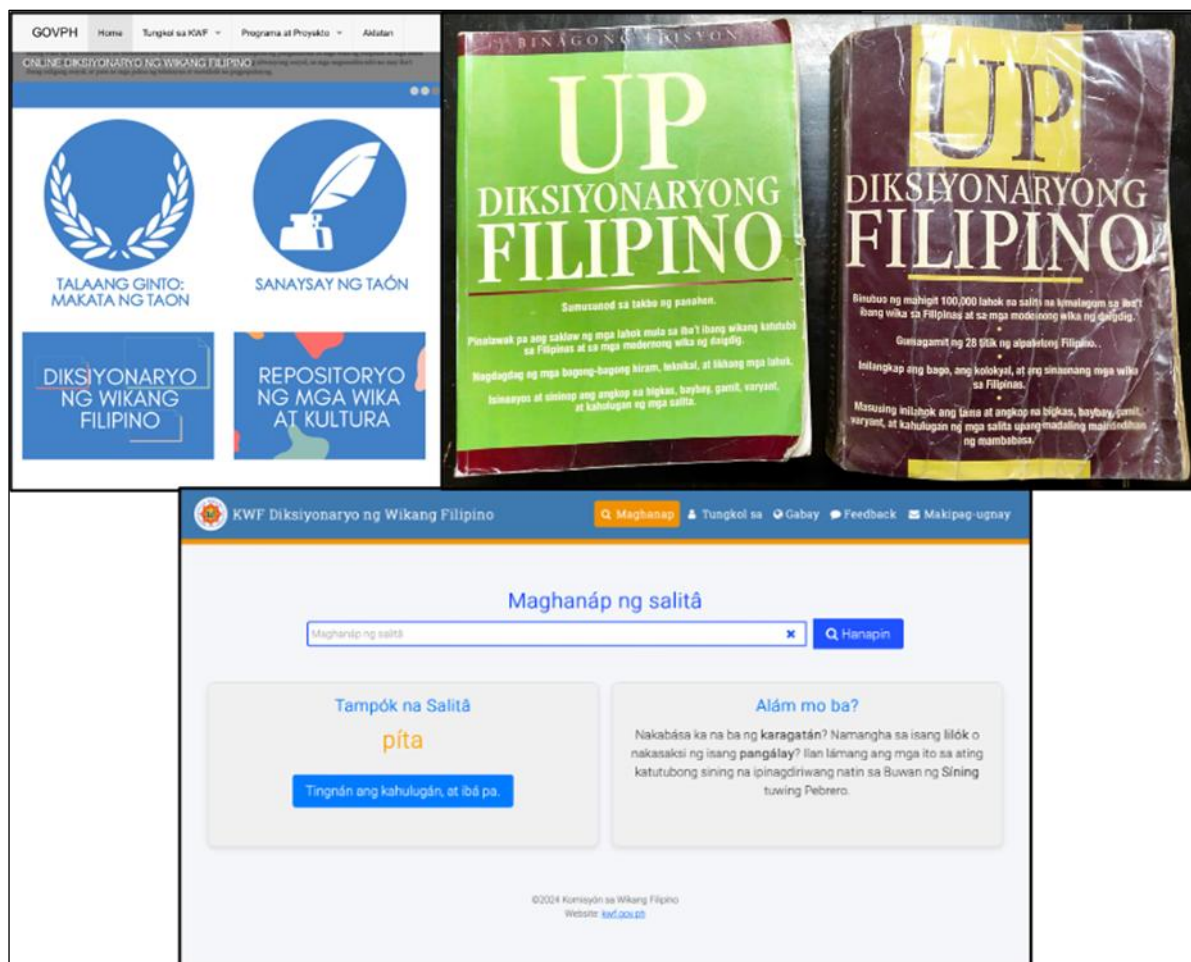
Figure 7: Cover—UPDF and KWF-DWF

Figure 7 shows the covers of the two editions of "UP Diksiyonaryong Filipino" (UPDF) in print and of the "KWF Diksiyonaryo ng Wikang Filipino" (KWF-DWF) in digital platform.

On top of the main entries for dictionaries, both UPDF and KWF-DWF added sections that extend their usefulness, inform about contextualised use of language, and live up to their different educational and cultural missions as shown in Figure 7; thus, each embeds its respective lexical philosophy and fulfills its language estate role in the development, education, and cultural use of Filipino as expressed in auxiliary dictionary components. On UPDF's preface, a comprehensive discussion of its lexicographical stance is made, highlighting descriptive quality of its content, and underlining its aim towards recognising the ever-changing and progressive condition of the Filipino language, including the selected usages of word formation, and adjustments of informal and regional usages. The introduction often emphasises the cultural and historical significance of Filipino, asserting the dictionary as a dual linguistic and cultural archive. At the same time, KWF-DWF showcases a brief foreword that highlights its purpose of supporting the national language policy and standardisation outlined by the "Komisyon sa Wikang Filipino."

Figure 8: Introductory—UPDF and KWF-DWF

Figure 8 shows the introductory phases of the two editions of "UP Diksiyonaryong Filipino" (UPDF) in print and of the "KWF Diksiyonaryo ng Wikang Filipino" (KWF-DWF) in digital platform.

The UPDF elucidates the development of spelling conventions, nature of diacritics, and common morphological processes, including affixation and reduplication as shown in Figure 8. The guidelines also offer a relative freedom in certain grammatical constructions, such as parallelism, that changes and expands the way the Filipino plays. As a prescriptive tool, KWF-DWF leans towards more orthographic and grammatical rules similar to an authoritative source based on official guidelines set by the "Komisyon sa Wikang Filipino," harmonising spelling, punctuation, and the structure of each sentence, from which educators, students, and government officials can reference. By emphasising codified norms, such efforts assist with language planning, and help in the standardisation of language through various institutions (35).
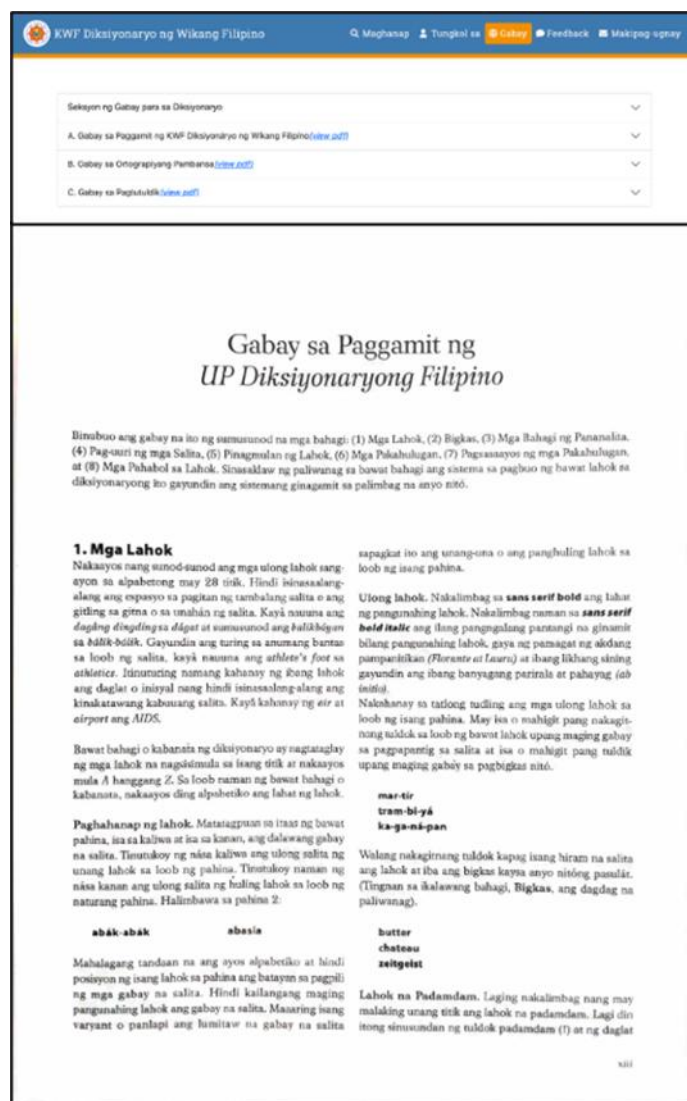
Figure 9: Guidelines—UPDF and KWF-DWF

Figure 9 shows the introductory phases focusing on guidelines on the use of the "UP Diksiyonaryong Filipino" (UPDF) in print and of the "KWF Diksiyonaryo ng Wikang Filipino" (KWF-DWF) in digital platform.

Discussion

The "UP Diksiyonaryong Filipino" (UPDF) serves as a useful resource for lexicographic-academic research, cultural studies, and advanced language learning while "KWF Diksiyonaryo ng Wikang Filipino" (KWF-DWF) takes a degree of uniformity to the guidelines set forth by language authorities. The latter appears online in a standardised format making it more accessible to educators, learners, and government entities. Although both dictionaries play pivotal roles in the evolution and purveyance of Filipino language, they serve different audiences and purposes. With extensive lemmatisation approaches, UPDF reflects to be helpful for advanced learners, researchers, and teachers looking to explore complex grammatical structure from an educational perspective. KWF-DWF's simplified lemmatisation provides students and non-experienced users access to a standardised version of the language, making it easier to learn basic vocabulary and grammatical rules. Both dictionaries are massively important, but serve diverging threads of contemporary lexicography. UPDF has a greater depth and diversity in its applications, whereas KWF-DWF is one of the main pillars with regard to the standard way of educating and communicating in formal settings.

The contrasting treatment of lemma forms between UPDF and KWF-DWF reflects an underlying bridge between descriptive and prescriptive lexicography in the Philippine

context. Although UPDF contributes to linguistic scholarship by documenting the living language, KWF-DWF serves the essential purpose of ensuring that language is standardised and taught. An ideal balance between descriptive richness and prescriptive clarity not only serves to bolster future lexicographical projects but also promotes linguistic diversity while maintaining some sense of consistency within language education. The fact that both dictionaries are relevant in this day and age emphasises the importance of having a back and forth in the field of lexicography moving forward—the progressive yet formalised and inclusive style of the UPDF contrasted with the standardisation and accessibility of the KWF-DWF. To remain relevant in an age of rapid change, both dictionaries must embrace digital tools, include user-generated content, and continuously update entries to capture the fluidity of language. The UPDF and KWF-DWF cater to various aspects of lexicography, and complement each other in terms of dictionary-making in the Philippines. Combined, these two dictionaries give a sense of the dual tasks that contemporary lexicography faces—documenting linguistic change, for one, but also encouraging and expanding linguistic consistency and accessibility. The subsequent sections of both texts transcend the mere definition of words as they help enrich language development, policy, and cultural identity formation in the Philippines.

The transition of lexicography from printed forms of data to computerised and digital platforms provides opportunities and challenges for "UP Diksiyonaryong Filipino" (UPDF) and "KWF Diksiyonaryo ng Wikang Filipino" (KWF-DWF). These works balance their foundational roles in the standardisation of language and the preservation of culture against the challenges of modern technology and face unique challenges in integrating traditional lexicographical methods with computational approaches, including data accessibility, structure and digitalisation, linguistic diversity and variation, updating and maintaining lexicographical relevance, cultural sensitivity and language evolution, and the traditional-computational divide. The solution to such challenges lies in a hybrid framework, borrowing descriptively inclusive structures from the UPDF, paired with the structural clarity of

KWF-DWF. Adopting digital innovations, like Artificial Intelligence (AI)-powered lexicography, crowdsourced content moderation, and other multimodal learning tools, allow a pair of dictionaries to both extend their relevance and expand their availability in the contemporary lexicographical milieu. It is imperative in these collaborative efforts between linguists, educators, and technologists that both traditional and computational needs are met without sacrificing cultural and linguistic sustainability.

Conclusion

This comparative analysis of the "UP Diksiyonaryong Filipino" (UPDF) and the "KWF Diksiyonaryo ng Wikang Filipino" (KWF-DWF) exposes some critical considerations regarding the divergent courses of Filipino lexicography evidenced through features, the ways lemmatisation is done, and the modern dynamics underpinning the choice of word entries. The findings show that the UPDF offers a descriptive research-based approach and emphasises linguistic diversity, historical depth, and cultural inclusivity with respect to variety whereas the KWF-DWF emphasises standardisation and accessibility, and closely aligns with language policy and educational needs. The UPDF provides thorough entries, including etymology, regional variants, and usage examples while the KWF-DWF has a practical function in education and governance, supporting linguistic consistency and further enabling formalised language education. Comparably, both dictionaries struggle to reconcile traditional lexicographical practice with contemporary computation. Although this qualitative approach allows for a thorough exploration, it limits external validity. This study only did content analysis with no use of user experience data or empirical testing of the dictionaries in educational or technological environments. Further, the study confined itself to only two principal dictionaries, with the exclusion of other regional or specialised lexicographical works that could have provided more comparative perspectives. Therefore, an area for future research could be conducting user-centered studies (e.g., surveys or usability tests) to assess the extent to which various types of end-users use these dictionaries, both in print and digital form. Looking into incorporating Filipino

lexicographical information to natural language processing (NLP) tools would further broaden the scope of the research, targeting regional dictionaries or multilingual lexicographical resources and filling out the picture of lexicography on a wider linguistic canvas. The findings have the potential to significantly enhance language learning tools, dictionaries, and lemmatisation engines by comparing the KWF-DWF and UPDF approaches, which yielded significantly different outcomes. On one hand, KWF-databases provide examples of high standardisation and may be beneficial for the development of automated spelling and grammar analysers, such as those for homographic cases. For example, the integration of UPDF's deep morphology and complex semantics into a NLP system would be a valuable contribution to the development of more precise lemmatisers and morphological analysers. Language instructors may capitalise on these findings to generate instructional materials that are sufficiently consistent and contextually relevant. Future dictionaries could be created using hybrid frameworks that are informed by the inclusiveness of UPDF and the clarity of KWF-DWF, and are further enhanced by AI and user-generated content to remain current with the evolution of language. This interdisciplinary framework is capable of achieving both educational and cultural objectives, as well as the development of robust NLP applications, including chatbots, translation applications, and educational platforms, that are tailored to the Filipino language and the multilingual context. This study is primarily a qualitative comparison of dictionary-based lemmatisation strategies and did not incorporate corpus-derived wordform patterns or include test cases, data samples, or error analysis. To address this gap, future research may contrast the lemmatisation performance of two dictionaries against corpus-based findings, providing empirical evidence of how entries correspond to actual language use. Incorporating test cases using corpus data or conducting an error analysis of dictionary lemmatisation may be considered to reveal patterns of inconsistency or areas for improvement. Such an approach would not only enhance the validity but also contribute to the development of more robust NLP tools, aligning computational lemmatisation with both theoretical models and real-world usage.

Abbreviations
AI: Artificial Intelligence, AY: Academic Year, CLSU: Central Luzon State University, CULU: Coverage, User Interface, Lemmatisation, Usability, ERC: Ethics Research Committee, KWF: Komisyon sa Wikang Filipino, KWF-DWF: KWF Diksiyonaryo ng Wikang Filipino, NLP: Natural Language Processing, TSU: Tarlac State University, TEI: Text Encoding Initiative Guidelines, UP: University of the Philippines, UPDF: UP Diksiyonaryong Filipino.

Author Contributions
Joseph A. Villarama: Conceptualisation, Introduction and Literature Review Writing, Methodology, Data Analyses, and Data Interpretation, Revision, Supervision.

Conflict of Interest
The author affirmed that there is no conflict of interest in this research article.

Ethics Approval
This study is conducted in accordance with the ethics protocol of the Ethics Research Committee (ERC) of Central Luzon State University (CLSU), Philippines, dated January 6, 2025.

References
1. Paglioni VP, Groth KM. Dependency definitions for quantitative human reliability analysis. Reliab Eng Syst Saf. 2022 Apr 1;220:108274.
2. Wierzbicka A, Wierzbicka A. "Semantic Primitives", fifty years later. Russian Journal of Linguistics. 2021 Jun 23;25(2):317–42.

3. Manzano DL. Development and Validation of a Monolingual Glossary of Terms Used by Cockfighters. 3L: Language, Linguistics, Literature® The Southeast Asian Journal of English Language Studies. 2018 Jun 29;24(2):16–28.

4. Manalastas NEL, Auxtero SVJ. Legalities of language use in linguistic landscaping. Language Problems and Language Planning. 2024 Jul 25;48(1):75–102.

5. Dimaculangan NG, Gustilo LE. Lexical Patterns in the Early 21st Century Philippine English Writing. Adv Sci Lett. 2017 Feb 1;23(2):1094–8.

6. Symaco LP. Education, language policy and language use in the Philippines. Language Problems and Language Planning. 2017 Jul 20;41(1):87–102.

7. Marquez FF, Bandril LT. Evelopment and Validation of a Wordlist Translated from English to Filipino of Academic-Administrative Terms: An Attempt at Formulating Guidelines for Policy Making. Procedia Soc Behav Sci. 2015 Jun 24;192:327–38.

8. Ge X, Zhang S, Xu H, Zhang X. A Scoping Review of Studies into Dictionary Use and Language Learning. Lexikos. 2024 Sep 25;34:331–54.

9. Zhang S, Xu H, Zhang X. The effects of dictionary use on second language vocabulary acquisition: A meta-analysis. International Journal of Lexicography. 2021 Mar 23;34(1):1–38.

10. Márkus KP, Dringó-Horváth I. Developing Dictionary Skills through Monolingual and Bilingual English Dictionaries at Tertiary-level Education in Hungary. Lexikos. 2023 Aug 30;33:350–81.

11. Alamri HR, Hakami HM. Exploring Perspectives of EFL Students on Using Electronic Dictionaries to Improve Vocabulary Learning: A Comparative Study. International Journal of Curriculum and Instruction. 2022;14(2):1578–99.

12. Cai W, Gao M, Jiang Y, et al. Hierarchical Domain Adaptation Projective Dictionary Pair Learning Model for EEG Classification in IoMT Systems. IEEE Trans Comput Soc Syst. 2023 Aug 1;10(4):1559–67.

13. López-Santana M, Tanca DE. Comparative Policy Analysis, Subnational Research, and Small-N Qualitative Research: Research Design Considerations. Journal of Comparative Policy Analysis: Research and Practice. 2024 Nov 1. https://www.tandfonline.com/doi/abs/10.1080/13876988.2024.2401809

14. Chen W. Towards a Discourse Approach to Critical Lexicography. International Journal of Lexicography. 2019 Sep 1;32(3):362–88.

15. Wiliński J. Conventional Knowledge, Pictorial Elucidation, Etymological Motivation, and Structural Elaboration in a Thematic Dictionary of Idioms. ANGLICA - An International Journal of English Studies. 2022;31(2):109–31.

16. Fuertes-Olivera PA. Lexicographical storing: a key lexicographical task in the era of the internet. Lexicographica. 2015 Nov 13;31(1):67–89.

17. Wang R, Shen M, Li Y, Gomes S. Multi-task Joint Sparse Representation Classification Based on Fisher Discrimination Dictionary Learning. Computers, Materials & Continua. 2018;57(1):25–48.

18. Kestemont M, de Pauw G, van Nie R, Daelemans W. Lemmatization for variation-rich languages using deep learning. Digital Scholarship in the Humanities. 2017 Dec 1;32(4):797–815.

19. Ooi VBY. Issues and prospects for incorporating English use in Japan into the dictionary. Asian Englishes. 2021 Jan 2;23(1):62–78.

20. Fattah BO, Salih SM. Colloquialism and the Community of Practice. Koya University Journal of Humanities and Social Sciences. 2022 Jun 30;5(1):77–84.

21. Nkomo D. Dictionary culture in African language communities: research, development, challenges and prospects. Lexicographica. 2020 Nov 1;36(2020):11–37.

22. Luo Y, Watts M. Exploration of university students' lived experiences of using smartphones for English language learning. Comput Assist Lang Learn. 2024;37(4):608–33.

23. Gupta R, Jivani AG. LemmaChase: A lemmatizer. International Journal on Emerging Technologies. 2020;11(2):817-24.

24. Nkomo D. Developing a dictionary culture through integrated dictionary pedagogy in the outer texts of South African school dictionaries: the case of 'Oxford Bilingual School Dictionary: IsiXhosa and English.' 2024 Nov 4;2(1 Special Issue):71–99.

25. Nielbo KL, Karsdorp F, Wevers M, et al. Quantitative text analysis. Nature Reviews Methods Primers 2024 4:1. 2024 Apr 11;4(1):1–16.

26. Tummers J, Speelman D, Heylen K, Geeraerts D. Lectal constraining of lexical collocations: How a word's company is influenced by the usage settings. Constructions and Frames. 2015 Jan 1;7(1):1–46.

27. Villarama JA, Barcelita KJ, Pilien RV, Crisanto WA. Padlet: Post-pandemic avenue for a more dynamic language and literature learning through enhanced technology integration. Issues in Language Studies. 2024 Jun;13(1):304-21.

28. Wang J, Zhao J. Revitalization and propagation approach of Chinese historical dictionaries in the digital age. International Communication of Chinese Culture 2024 11:4. 2024 May 2;11(4):499–513.

29. Knežević L, Halupka-Rešetar S, Miškeljin I, Milić M. Millennials as Dictionary Users: A Study of Dictionary Use Habits of Serbian EFL Students. Sage Open.2021;11(2). https://journals.sagepub.com/doi/10.1177/21582440211008422

30. Galla CK. Indigenous language revitalization, promotion, and education: function of digital technology. Comput Assist Lang Learn. 2016 Oct 2;29(7):1137–51.

31. Villarama JA. Demystifying Generative Artificial Intelligence in Academic Classrooms: Students' Attitude and Critical Thinking. Pakistan Journal of Life and Social Sciences (PJLSS). 2025;23(1). https://www.pjlss.edu.pk/pdf_files/2025_1/3970-3979.pdf

32. Hanif A, Deswita D, Budiarti M, Herman H, Mudinillah A, Putri LR. Development of a Digital Dictionary for Measuring Arabic Language Education Students Retention. Migration Letters. 2023 Aug 2;20(5):321–39.

33. Çolak F, Balaman U. The use of online dictionaries in video-mediated L2 interactions for the social accomplishment of Virtual Exchange tasks. System. 2022 Jun 1;106:102772.

34. Lew R. Research into the use of online dictionaries. International Journal of Lexicography. 2015 Jun 1;28(2):232–53.

35. Koreinik K. Extralinguistic arguments in 21st century language planning discourse: a "superdictionary" between language standardization from above and below. Eesti ja soome-ugri keeleteaduse ajakiri Journal of Estonian and Finno-Ugric Linguistics. 2023 Oct 19;14(1):117-148.