

# **Evaluating Consistency Across Beginner and Intermediate Filipino Textbooks for a Model of Word Level Identification**

**Florinda Amparo A. Palma Gil**

## **Abstract**

The Filipino language is taught as a second or foreign language globally, including in Asia, North America, and Europe. In Japan, Filipino classes began in 1989, but the persistent challenge remains the lack of adequate teaching and learning materials (Laranjo, 2020). Furthermore, the author observed inconsistencies in terminology—Filipino, Tagalog, or Tagalog/Filipino—across existing reference materials. Lower-level materials are variably labeled as “basic,” “beginner,” or “elementary,” while higher-level materials are broadly categorized as “intermediate.” This study aims (1) to evaluate the consistency of the language referenced

in beginner and intermediate Filipino textbooks and (2) to identify differences in vocabulary between these levels, contributing to a model of word level identification, to address the following questions: (a) Are Filipino language teachers and students worldwide engaging with the same target language? (b) Are “beginner” and “intermediate” learners across countries exposed to the same vocabulary at these levels?

The study compared and analyzed the glossaries of eight textbooks/reference materials used for teaching Filipino globally. Three textbooks were designated for “beginner” level, two for “intermediate,” and the remaining three for “beginner to intermediate” levels. Despite variations in how the target language is named, at least five textbooks refer to the national language of the Philippines, Filipino, as their target language. For the “beginner” level, 4,889 unique words were extracted, with 894 common across six textbooks. At the “intermediate” level, 5,767 unique words were identified, with 674 common across five textbooks. Notably, “intermediate” textbooks contained less number of nouns and verbs, but the themes under these categories were found to be more compared to those at the “beginner” level.

*Keywords:* Filipino language, word level identification, vocabulary list comparison, beginner and intermediate textbooks, textbook analysis

## 1 Introduction

The 1987 Constitution designates Filipino as the national language of the Philippines (Article XVI, Section 6), with a mandate for it to evolve by incorporating elements from Philippine and other languages. However, it took five decades for Filipino to be formally named and defined. In 1937, Tagalog was declared the basis of the national language, which led to confusion over whether Tagalog itself was the national language or merely a foundation. In 1959, the language was renamed Pilipino to distinguish it from Tagalog, and in 1987, it was renamed again as Filipino, now incorporating elements from other Philippine languages, as well as from Spanish and English (Almario & Kilates, 2014).

Even before discussions about the national language began, Tagalog held prestige as the language of the capital and was more widely used in literature and newspapers than any other Philippine language (Almario & Kilates, 2014; Tupas, 2014). During the long period when Tagalog was the sole basis for the national language, the language spread and expanded through mass media and education, despite opposition (Constantino, 2012). Then, with the revision of the constitution in 1987, which declared that the national language would no longer be based on a single language but would instead draw from all Philippine languages as well as Spanish and English, changes were made to aid its modernization and intellectualization. One of the first changes was in the alphabet: eight letters (C, F, J, Ñ, Q, V, X, Z) were added to the old *abakada* to better accommodate sounds from other Philippine languages, as well as from Spanish and English, enriching the vocabulary (Almario & Kilates, 2014). However, even to this day, many continue to refer to Tagalog as the national language, despite

the fact that the national language taught in schools is officially called Filipino.

## 1.1 On Filipino Language Teaching and Learning

Filipino language teaching and learning is not confined to the Philippines but has expanded its reach to countries worldwide, including the United States of America, Canada, United Kingdom, Russia, France, Germany, China, Japan, South Korea, Brunei, Malaysia, and Australia (Laranjo, 2020; Ramos & Mabanglo, 2012; Takahata, 2022; Yap, 2012). While the manner of Filipino language instruction varies across these countries, it is noteworthy that classes are offered at all educational levels—primary, secondary, and tertiary (Takahata, 2022; Yap, 2012). Diverse learners engage with the language, with heritage language learners predominating in the USA and various European nations, foreign language learners prevalent in East Asia, and a mix of heritage, foreign, and second language learners in Southeast Asia, where the Philippines is situated (Laranjo, 2020; Yap, 2012).

The rising Filipino diaspora has heightened the interest in learning Philippine language and culture, particularly evident in the USA, where Filipino ranks as the second most commonly spoken Asian language and the sixth most spoken non-English language, according to the 1990 United States Census (Ramos & Mabanglo, 2012). Despite this growing demand, there remains a deficiency in learning materials, compounded by inconsistent nomenclature—Filipino, Tagalog, or Tagalog/Filipino—in available reference materials. Additionally, a proficiency assessment tool for those seeking it is conspicuously absent. Furthermore, while textbooks for higher levels are uniformly labeled as

“intermediate,” those for lower levels lack consistent designations, being referred to as “basic,” “beginner,” or “elementary.”

This study is grounded on the premise that the existence of a reliable assessment tool would facilitate a seamless transition for students moving to another country or transitioning to a different online course. Such a tool would offer a transparent means for individuals to communicate their current language proficiency, aiding schools in accurately assessing incoming students. Conversely, the absence of a reliable assessment tool could lead to students inadvertently retaking a course or enrolling in a class beyond their proficiency level, hindering the smooth progression of their learning (Byrnes, 1990; Laranjo & Palma Gil, 2023).

This study aims (1) to evaluate the consistency of the language referenced in beginner and intermediate Filipino textbooks and (2) to identify differences in vocabulary between these levels, contributing to a model of word level identification, to address the following questions: (a) Are Filipino language teachers and students worldwide engaging with the same target language? (b) Are “beginner” and “intermediate” learners across countries exposed to the same vocabulary at these levels? The goal is to leverage the study’s findings in developing a model for a user-friendly word level checker. This tool, once created, can serve as a valuable reference for the construction of teaching materials and proficiency assessment tools in Filipino language education.

## **1.2 On Foreign Language Teaching and Learning Guidelines and Related Studies**

Globally, the American Council on Teaching Foreign Language (ACTFL) Language Proficiency Guidelines in the USA and the

Common European Framework of Reference for Languages (CEFR) by the Council of Europe wield significant influence on language proficiency tests and standardization. These guidelines categorize language textbooks, assessment tools, and materials based on proficiency levels—ACTFL Proficiency Guidelines for materials in the USA and CEFR Levels for materials in European countries.

The ACTFL Language Proficiency Guidelines, which were developed in 1986, delineate proficiency into five major levels across each skill (speaking, writing, listening, reading): Distinguished, Superior, Advanced, Intermediate, and Novice. The major levels Advanced, Intermediate, and Novice are further divided into High, Mid, and Low sublevels. This framework defines the continuum of proficiency, ranging from highly articulate, well-educated language users to those with minimal or no functional ability (ACTFL, 2023). Conversely, the CEFR which was introduced in 2001, consists of six proficiency levels—A1, A2, B1, B2, C1, and C2, organized into three overarching categories: Basic User, Independent User, and Proficient User. These can be further subdivided based on local context needs, and the levels are characterized through ‘can-do’ descriptors (Council of Europe, 2023). CEFR ratings assigned on ACTFL assessments is presented in Section 6.1 (ACTFL, 2022).

The ACTFL and CEFR have a profound impact, extending their influence on languages across continents. An illustrative case is The JF Standard for Japanese-Language Education, developed by the Japan Foundation, which aligns with CEFR. This standard not only draws inspiration from CEFR but also incorporates its Common Reference Levels and Can-do Descriptors (Japan Foundation, 2010). Similarly,

the Japanese Language Proficiency Test (JLPT) recently announced the addition of a CEFR Level indication to JLPT score reports starting from the 2025 JLPT results (The Japan Foundation & Japan Educational Exchanges and Services, 2023).

CEFR's influence also extends to the development of CEFR-J, an innovative framework for English language teaching in Japan. Distinguishing features of CEFR-J include (a) more refined sub-levels (Pre-A1, A1.1-1.3, A2.1-2.2, B1.1-1.2, B2.1-2.2) with newly created and scaled descriptors, (b) the preparation of grammar and vocabulary corresponding to each CEFR-J level, (c) a detailed analysis of text features representing CEFR-J levels, and (d) the creation of tasks and tests tailored to each CEFR-J descriptor (Tono, 2017 in Tono, 2019, pp. 5–6).

However, before ACTFL and CEFR there was the Interagency Language Roundtable (ILR) Language Proficiency Skill Level Descriptions and Scales, which was developed in the 1950s and formalized in the 1960s by the U.S. government to standardize language proficiency assessment. The ILR scale provides a framework for assessing functional language skills, with levels ranging from 0 (no proficiency) to 5 (native or bilingual proficiency) describing language skills in listening, reading, speaking, and writing. It evolved from earlier military and diplomatic language training programs and has since become a key tool for government agencies and language professionals (Interagency Language Roundtable, n.d.). The ILR scale focuses on real-world communicative ability and has influenced other proficiency scales, such as the CEFR and ACTFL. In fact, ACTFL was derived from the ILR scale.

In the realm of Filipino language education, the absence of proficiency tests and the lack of comprehensive teaching guidelines are notable. The author's investigation revealed only two available resources. The first is the Frequency Count of Filipino by McFarland (1989), designed as a foundational wordlist for language instruction in elementary schools and as a guide for textbook preparation based on written corpus from the period in which Pilipino was the dominant term for the national language (p. 5). The second is the Language Learning Framework for Teachers of Filipino (LLFTF) and its accompanying curriculum guide, tailored for Filipino heritage learners in the USA. Developed in 2012 by the Council of Teachers of Southeast Asian Languages (COTSEAL), this framework spans four teaching levels, equivalent to four years of instruction: Level 1 (101 and 102), Level 2 (201 and 202), Level 3 (301 and 302), and Level 4 (401 and 402). The LLFTF outlines objectives for each of the four macro language skills (writing, reading, listening, and speaking) at every level. Within each level, nine specific learning objectives are identified: Novice, Novice Mid, Novice High, Intermediate, Intermediate Low, High Intermediate (Intermediate High), Skilled (Advanced), Skilled Low (Advanced Low), and Skilled High (Advanced High). Notably, Level 1 caters to 0-Novice to Mid-levels, Level 2 to Novice High to Intermediate Mid-levels, Level 3 to Intermediate Mid to Intermediate High levels, and Level 4 to Intermediate High to Advanced levels (Ramos & Mabanglo, 2012). Given its development in the USA, it is reasonable to infer that the terminology employed aligns with the standards set by the ACTFL.



### 1.3 On Wordlists, Word Level Checker and Related Studies

The CEFR-J wordlist underwent meticulous development through a comprehensive frequency analysis of English textbooks employed in primary and secondary schools across neighboring Asian countries/regions, such as China, Korea, and Taiwan. The process involved scrutinizing the learning objectives outlined in the national curricula of the target textbooks and approximating the CEFR levels associated with each. Utilizing these textbooks, a CEFR-level textbook corpora, spanning Pre-A1 to B2 Level, was established. To analyze the CEFR-level textbook corpora, the texts underwent initial tagging for parts of speech (POS) using TreeTagger from the study of Schmidt (1994, as mentioned in Tono, 2019). Subsequently, frequency lists of lemmas with POS were generated for each textbook published in every country/region, as well as for each CEFR level. The identification of Pre-A1 words was achieved by selecting only those that appeared in textbooks classified at the Pre-A1 level across all three regions. Following the same methodology, A1-level words through B2-level words were systematically extracted. Finally, the word levels were aligned with the English Vocabulary Profile (<https://www.englishprofile.org/wordlists>) (Tono, 2019, p. 7). The methodology carried out in comparing the vocabulary lists from eight Filipino language textbooks was roughly based on the development of the CEFR-J wordlist.

The English Vocabulary Profile (EVP) is a reference tool developed as part of the larger English Profile Project aimed at aligning English language teaching and learning materials with the Common European Framework of Reference for Languages (CEFR). It provides detailed

information about which English words and phrases are typically known and used at each CEFR level (from A1 to C2). The EVP is especially useful for educators, learners, and test designers as it helps to identify the vocabulary learners are expected to know at different proficiency stages. It is part of the Cambridge English Corpus and is available as an online tool for vocabulary profiling (English Profile, n.d.).

The EVP classifies vocabulary by considering not just the words themselves, but also different meanings, forms, and uses of those words at various CEFR levels. For instance, the word “run” might be recognized as a verb at the beginner level (A1), while more complex uses of “run” as a noun or idiomatic expressions like “run into” might appear at more advanced levels. According to EVP, there is a significant progression from A1/A2 (basic user) to B1/B2 (independent user), both in the range of vocabulary and the types of words learners are expected to know and use. B level vocabulary builds on A level vocabulary in terms of complexity, depth, and the ability to handle more abstract concepts. A1/A2 learners acquire the basics, which provide a foundation for more sophisticated usage at B1/B2. B1/B2 learners are expected to extend their knowledge by adding more abstract, technical, and context-specific vocabulary, but they still rely on the basic structures and vocabulary learned at A levels. For example, an A2 learner might say: “I don’t like this job. It’s boring.” A B2 learner would expand on that by saying: “I find this job incredibly monotonous and uninspiring. I’m looking for something more challenging and fulfilling.” In this way, B level vocabulary continues to build upon A-level knowledge, moving from basic to more complex language as learners progress (English Profile, n.d.). In many other languages, there are CEFR-aligned wordlists and

language profiling tools that work similarly to EVP, often developed by national language institutions or educational bodies that focus on language learning. Many of these tools, like EVP, aim to standardize vocabulary acquisition and language assessment, ensuring that learners progress systematically across CEFR levels. However, for the Filipino language, while there are few resources and corpora that can provide a foundation for analyzing vocabulary, syntax, and usage in various contexts, currently, there is no exact equivalent of the EVP for the Filipino language.

The New Word Level Checker (NWLC), developed by Mizumoto et al. (2021) and accessible at <https://nwlc.pythonanywhere.com/>, is an innovative online tool designed specifically for vocabulary profiling in Japanese English as a Foreign Language (EFL) learning environments. This application analyzes English texts submitted by users, generating a coverage profile based on built-in and user-selected word lists. NWLC boasts five reputable, research-based word lists as of September 2021: SEWK-J, New JACET8000, SVL12000, the New General Service List (NGSL), and Common European Framework of Reference for Languages (CEFR-J). Consequently, NWLC enables the alignment of learners' actual vocabulary knowledge with the texts under analysis (Mizumoto et al., 2021, pp. 30–31).

The article introducing NWLC provides clarity on the definition of lexical units, including the rules governing word counting, such as case sensitivity, contractions, abbreviations with periods, hyphenated words, and compounds. The study introduces the concept of “flemmatization (family + lemma)” in word counting, where the base form of a head-word is counted as one with its inflected forms. For example, the base

headword *study* (noun) is counted as one with its inflected form like *studied* (verb). Another example is when the headword *study* (noun) is counted as one with the headword *study* (verb). Additionally, the study distinguishes “lemmatization,” where the headword is counted separately from its inflected forms. For example, the headword *study* (noun) is counted as a separate headword from *study* (verb). “Lemmatization” is employed when part-of-speech (POS) information is required. The CEFR-J Wordlist adopts lemma counting since the original CEFR word lists are also lemmatized. Thus, in the CEFR wordlists, the verb “study” is classified as A1 and the noun “study” is A2 (Mizumoto et al., 2021, p. 33).

In the process of counting “common beginner words” and “common intermediate” across the eight Filipino textbooks, a hybrid approach of “flemmatization” and “lemmatization” was adopted for a simpler and easier way of counting. For example, the headword, *bukas* ‘tomorrow’ (adverb) is counted as one with headword *bukas* ‘to open’ (verb). However, the base headword *kain* ‘to eat’ (verb root) is counted as a separate word from *kumain* ‘to eat’ (verb *-um-* infinitive). In general, the rules created for data cleaning and defining a “word” for the purpose of counting “common” words was patterned after the rules established for NWLC.

The 1989 Frequency Count of Filipino by McFarland (1989), one of the scant two available resources providing teaching guidelines in Filipino language education, remains relevant despite its age. Despite being dated, its rules and insightful treatment of lemmas and headwords in Filipino still hold significant applicability to contemporary challenges in Filipino orthography. Noteworthy examples include addressing variations in

spelling or pronunciation for words with foreign and Filipinized forms, such as *professor* and *propesor*, or the multiple spellings like *puwede*, *pwede*, and *puede*. McFarland (1989, pp. 16–22) also delves into the intricate verbal system of Filipino, elucidating aspects such as inflection for aspect, which involves a root word and its myriad possible affixes. More importantly, McFarland’s study includes three definitions of what a “word” is. (1) a word is any sequence of letters separated from other sequences by one or more spaces; (2) a word is any given sequence of letters which can be set off by spaces, and which can be understood to occur repeatedly within a text (this talks about frequency of a word); and (3) different sequences are grouped together on the basis that they share basic meaning and differ only with regard to tense, number, case, etc. as how words are presented in dictionaries, e.g. sequence and sequences are understood to be different forms of the same word (McFarland, 1989, p. 9). Thus, this material also served as a guide in creating the rules for counting the “common” words across the textbooks investigated.

## **2 Methodology**

In this study, eight of the few available textbooks/reference materials used in teaching Filipino language in universities inside and outside the Philippines were compared and analyzed. Five of the textbooks are being used in the USA (mostly as supplementary materials), two of them are being used in Japan and one is being used in a university in South Korea and in a university in the Philippines. Since the textbooks are used in different countries, their target learners vary as well. Table 1 presents a summary of the background of the textbooks investigated in this study.

Table 1. Textbooks Investigated in This Study

| Textbook   | Country-dominant                       | Target Learners as indicated in the textbooks <sup>1</sup>                       | Level or targeted proficiency as described in the textbooks |
|--|--|--|---|
| <b>BOOK #1:</b><br><i>Tagalog for Beginners</i><br>(Barrios, 2014)   | USA                                    | different types of learners, e.g. heritage learners and second-language learners | beginner  |
| <b>BOOK #2:</b><br><i>Conversational Tagalog</i> (Ramos, 1985)   | USA                                    | foreign/second language learners   | SI+ (or higher) by the end of the session                   |
| <b>BOOK #3:</b><br><i>Panimulang Pag-aaral ng Wikang Filipino/Introductory Study of the Filipino Language</i><br>(Peregrino et al., 2016) <sup>2</sup> | South Korea & Philippines <sup>3</sup> | Koreans  | basic; introductory   |
| <b>BOOK #4:</b><br><i>Intermediate Tagalog</i> (Ramos & Goulet, 1982)  | USA                                    | Americans and Filipino-Americans   | intermediate  |

## Evaluating Consistency Across Beginner and Intermediate Filipino Textbooks for a Model of Word Level Identification

| Textbook   | Country-dominant | Target Learners as indicated in the textbooks                                    | Level or targeted proficiency as described in the textbooks |
|--|------------------|--|---|
| <b>BOOK #5:</b><br><i>Intermediate Tagalog</i> (Barrios, 2015)   | USA              | different types of learners, e.g. heritage learners and second-language learners | intermediate  |
| <b>BOOK #6:</b> フィリピン語 ( <i>Firipin-go</i> ‘Filipino’) (Oue & Yoshizawa, 2017)   | Japan            | Japanese <sup>4</sup>  | beginner to intermediate                                    |
| <b>BOOK #7:</b> 大学のフィリピン語 ( <i>Daigaku no Firipino-go</i> ‘Komprehensibong Tekstbuk ng Filipino/University Filipino’ ‘A Comprehensive Filipino Textbook’) (Yamashita et al., 2018) | Japan            | Japanese <sup>5</sup>  | beginner to intermediate or advanced                        |

| Textbook  | Country-dominant | Target Learners as indicated in the textbooks | Level or targeted proficiency as described in the textbooks       |
|---|------------------|---|---|
| <b>BOOK #8:</b><br><i>Elementary Tagalog</i><br>(Domigpe & Domingo, 2014) | USA              | heritage and non-heritage language learners   | intermediate low-level to intermediate mid-level after completion |

BOOK #2 aims for learners to reach speaking proficiency at the SI+ (Elementary Proficiency, Plus) level, which is assumed to be based on the ILR scale given the book’s 1985 publication date, around the time when ACTFL was being developed from the ILR scale. At this level, speakers can handle basic conversations but may struggle with social conventions and language control. Although no solid reference was

<sup>1</sup>The definitions of terms like second language can vary depending on the context and academic framework. Some sources follow a strict geographical distinction: **second language learning** is used when the language is learned in a setting where it is spoken, while **foreign language learning** is used for languages learned outside that context (Saville-Troike & Barto, 2016). Other sources, particularly in psycholinguistics or applied linguistics, may use **second language** to refer to any language learned after the first, without emphasizing the environment (Ellis, 1997; Krashen, 1982).

<sup>2</sup>This textbook was only published as an interactive book application during the data gathering and cleaning. However, just recently, this textbook has been published as part of a series of physical textbooks on Filipino in South Korea.

<sup>3</sup>The textbook does not explicitly state its use in the Philippines; however, an interview with one of its authors revealed that it is also utilized for teaching Filipino to Korean students at the University of the Philippines, Diliman.

<sup>4</sup>The textbook does not explicitly state that its target users are Japanese, but it is only available in Japan, and both the script and language used are in Japanese.

<sup>5</sup>The textbook does not explicitly state that its target users are Japanese, but it is only available in Japan, and both the script and language used are in Japanese.

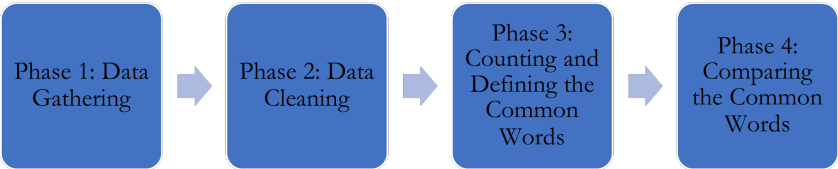


found regarding the correspondence between the ILR scale and the ACTFL or CEFR scales, a different approach was taken by comparing the expected functional and linguistic skills described in the textbook to the other textbooks. Upon close comparison of the learning objectives of the textbooks, it was found that the expected skills in BOOK #2 closely align with those in BOOK #3, both emphasizing foundational language abilities such as simple conversations, directions, and grammar for beginner learners. These texts likely correspond to the A1-A2 levels of the CEFR, focusing on practical, everyday interactions. For more on this comparison, see Section 6.2.

BOOK #3 which used the term “elementary” in its title, actually mentioned that the book covers a whole academic year and that the students are expected to achieve intermediate low level to intermediate mid-level after completion of the book. To achieve the two objectives of this study: (1) to evaluate the consistency of the language referenced in beginner and intermediate Filipino textbooks and (2) to identify differences in vocabulary between these levels, contributing to a model of word level identification, to address the following questions: (a) Are Filipino language teachers and students worldwide engaging with the same target language? (b) Are “beginner” and “intermediate” learners across countries exposed to the same vocabulary at these levels?, two approaches were employed. First, to identify the language referenced in the textbooks, a comparative content analysis was conducted across eight textbooks. This analysis examined how the target language is defined and presented throughout each textbook, with a focus on the introduction, terminology, grammatical frameworks, and pedagogical goals. Second, to differentiate vocabulary between beginner and intermediate levels,

the words from the glossaries or vocabulary lists in the textbooks were counted and compared using a simple R program to identify the “common beginner words” and “common intermediate words.” These words were further categorized by parts of speech and themes for additional differentiation.

In reference to the level or targeted proficiency indicated in each textbook, three textbooks were designated for “beginner” level [BOOK #1, BOOK #2, BOOK #3], two for “intermediate,” [BOOK #4, BOOK #5], and the remaining three for “beginner to intermediate” levels [BOOK #6, BOOK #7, BOOK #8]. To compare and analyze the eight textbooks and the frequency list, this study underwent four phases: [Phase 1: Data Gathering], [Phase 2: Data Cleaning], [Phase 3: Counting and Defining the Common Words] and [Phase 4: Comparing the Common Words]. Figure 1 shows an overview of the procedure applied in this study.



**Figure 1. Procedure Done in this Study**

In [Phase 1: Data Gathering], two types of data were gathered. These data were investigated to accomplish the two objectives and address the two questions of this study. The first set of data (named Data 1) is the “description/content” of each textbook which include its foreword, introduction, sample of its terminology, grammatical frameworks and pedagogical goals. The second set of data (named Data 2) is the “glossary or vocabulary list” of each textbook. The “description/content” were in-

## Evaluating Consistency Across Beginner and Intermediate Filipino Textbooks for a Model of Word Level Identification

vestigated to confirm whether the teachers and learners across countries are dealing with the same target language as presented in each textbook, while the “glossary or vocabulary list” were analyzed to describe and differentiate the “beginner” vocabulary from “intermediate” vocabulary.

Phases 2, 3 and 4 were done only for Data 2. In [Phase 2: Data Cleaning], rules for cleaning the data and rules for identifying and counting the common words among the textbooks and the frequency list were created and then these rules were implemented in a simple R program. Cleaning the data to identify and count the common words mean defining what a “word” is. The rules set in the development of the New Word Level Checker (<https://nwlc.pythonanywhere.com/>) for English Vocabulary by Mizumoto et al. (2021) in their study and the issues mentioned by McFarland (1989) regarding lemmas and headwords in Filipino were used as guides in creating the rules for cleaning the data of this study. As mentioned above, this study applied both “flemmatization” and “lemmatization.” As an example of “flemmatization,” the headword, *bukas* ‘tomorrow’ (adverb) is counted as one with headword *bukas* ‘to open’ (verb). However, the base headword *kain* ‘to eat’ (verb root) is counted as a separate word from *kumain* ‘to eat’ (verb -um-infinitive).

Some of the vocabulary lists contain borrowed words from English, such as *beer* for *mag-beer* and *text* for *i-text* and *mag-text*. However, to further simplify the counting process and due to their very low frequency, these words were eliminated from the list.

Table 2 shows the rules created for cleaning the data and for counting the common words.

Table 2. Rules Created to Clean the Data and Count Common Words

| Rules on cleaning data 2  | Rules on counting common words   |
|---|--|
| <ol style="list-style-type: none"><li>1. Separate sentences and phrases into words.</li><li>2. Separate words with slash.</li><li>3. Accept and retain words with hyphen and apostrophe.</li><li>4. Eliminate English words.</li><li>5. Do not include punctuation marks and special characters; Delete punctuation marks like period, question mark, exclamation mark, parentheses and special characters like tilde and asterisks.</li><li>6. Change vowels with accents marks to vowels without accent marks.</li><li>7. Change capital letters to lower case letters.</li><li>8. Make each word unique (no duplicates).</li></ol> | <p><i>Common Words in “Beginner”</i><br/><i>Textbooks:</i> “A word is considered common <b>if it appears in at least 3 books out of the 6 textbooks</b> indicated for counting ‘beginner’ words.” [BOOK #1, BOOK #2, BOOK #3, BOOK #6, BOOK #7, BOOK #8].</p> <p><i>Common Words in “Intermediate”</i><br/><i>Textbooks:</i> “A word is considered common <b>if it appears in at least 3 books out of the 5 textbooks</b> indicated for counting ‘intermediate’ words.”[BOOK #4, BOOK #5, BOOK #6, BOOK #7, BOOK #8]</p> |

After the data were cleaned and counted, in [Phase 3: Counting and Defining the Common Words], the common words that were found were counted. Words are called “common” among the six textbooks indicated for counting “beginner” words when they appear in three out of the six textbooks [BOOK #1, BOOK #2, BOOK #3, BOOK #6, BOOK #7, BOOK #8]. These six textbooks will be referred to as “beginner set.” Words are called “common” among the five textbooks indicated for counting “intermediate” words when they appear in three out of the five textbooks [BOOK #4, BOOK #5, BOOK #6, BOOK

#7, BOOK #8]. These five textbooks will be referred to as “intermediate set.”

Finally, in [Phase 4: Comparing the Common Words], the common words from the textbooks that contained “beginner” vocabulary were compared to textbooks which contain “intermediate” vocabulary in two ways. First, the number of overlapping “beginner” words within the “intermediate” was counted, and then the overlapping and non-overlapping words were identified. Second, the words at both levels were categorized by part of speech and themes to further distinguish them.

### **3 Findings and Analysis**

#### **3.1 Identifying the Language Referenced in Available Filipino Textbooks**

The first objective of this study is to evaluate the consistency of the language referenced in beginner and intermediate Filipino textbooks to answer the question: (a) Are Filipino language teachers and students worldwide engaging with the same target language? To address this objective and question, a comparative content analysis was conducted across eight textbooks. This analysis examined how the target language is defined and presented throughout each textbook, with a focus on the introduction, terminology, grammatical frameworks, and pedagogical goals.

Table 3 shows the target language indicated and described in each textbook. The eight textbooks refer to the target language in various names: Tagalog, Pilipino, Filipino.

Table 3. Target Language as Indicated in Each Textbook

| Textbook  | Target Language as indicated and described in each textbook               |
|---|---|
| <b>BOOK #1:</b> <i>Tagalog for Beginners</i><br>(Barrios, 2014)   | Tagalog (the basis of the national language of the Philippines, Filipino) |
| <b>BOOK #2:</b> <i>Conversational Tagalog</i><br>(Ramos, 1985)  | Tagalog   |
| <b>BOOK #3:</b> <i>Panimulang Pag-aaral ng Wikang Filipino/Introductory Study of the Filipino Language</i> (Peregrino et al., 2016)   | Filipino, pambansang wika (national language)                             |
| <b>BOOK #4:</b> <i>Intermediate Tagalog</i><br>(Ramos & Goulet, 1982)   | Tagalog   |
| <b>BOOK #5:</b> <i>Intermediate Tagalog</i><br>(Barrios, 2015)  | Tagalog/Filipino (the national language of the Philippines)               |
| <b>BOOK #6:</b> フィリピン語<br>( <i>Firipin-go</i> ‘Filipino’) (Oue & Yoshizawa, 2017)   | Filipino (national language/official language; Tagalog)                   |
| <b>BOOK #7:</b> 大学のフィリピン語<br>( <i>Daigaku no Firipino-go</i><br>‘Komprehensibong Tekstbuk ng Filipino/University Filipino’ ‘A Comprehensive Filipino Textbook’)<br>(Yamashita et al., 2018) | Filipino (the national language and one of the official languages)        |
| <b>BOOK #8:</b> <i>Elementary Tagalog</i><br>(Domigpe & Domingo, 2014)  | Tagalog (the basis of the national language of the Philippines, Filipino) |

Two of the textbooks, BOOK #1 and BOOK #5, written by the same author (Barrios, 2014; Barrios, 2015), use “Tagalog” in their titles but describe their target language as the national language of the Philippines,

## Evaluating Consistency Across Beginner and Intermediate Filipino Textbooks for a Model of Word Level Identification

referring to it as Filipino or Pilipino on the back cover and in the introduction. In BOOK #1, the author mentions that Filipino is based on Tagalog, while in BOOK #5, the same author clarifies that Filipino is not solely based on Tagalog (Barrios, 2015). BOOK #1 also includes an appendix titled “*Why Filipino and not Pilipino? – A Brief History of Making a Language the National Language*,” indicating the book acknowledges the distinctions between Tagalog, Pilipino, and Filipino (Barrios, 2014).

BOOK #2 and BOOK #4 use “Tagalog” in their titles and refer to the language as Tagalog in the content and structure. These textbooks, published in 1982 and 1985 before Filipino was officially recognized as the national language—are the oldest among the set. BOOK #2 states its purpose is to help students practice idiomatic Tagalog conversation (Ramos, 1985). BOOK #4 highlights *Tagalog Reference Grammar* by Paul Schacter and Fe T. Otones, as well as *Tagalog Structures* by one of the authors herself of BOOK #4, as its primary sources for grammar explanation (Ramos & Goulet, 1982).

BOOK #3, developed by the Departamento ng Filipino at Panitikan ng Pilipinas (DFPP)/Department of Filipino and Philippine Literature of the University of the Philippines, Diliman (UPD), and Busan University of Foreign Studies (BUFS), focuses on Filipino as its target language. The book provides a brief explanation of Filipino, the national language recognized in the Philippine constitution, and emphasizes its importance, along with what aspects of the Philippines can be learned through it. Lesson 1 features a short reading text on the composition of Filipino, derived from various Philippine and foreign

languages, followed by a list of Filipino words with foreign influences (Peregrino et al., 2016).

BOOK #6 and BOOK #7 are used at universities in Japan offering Philippine Studies programs. Both textbooks use “Filipino” in their titles and identify Filipino as the target language. While both provide a brief introduction to the relationship between the Philippines and Japan, BOOK #7 also includes a history of the national language, recognizing Filipino as its official name (Yamashita et al., 2018). BOOK #6, on the other hand, states that Tagalog is still more commonly used in the Philippines and argues that, aside from the alphabet, there are no significant differences between Filipino and Tagalog today. The author refers to the two terms as “nicknames” for the same language but notes that Filipino will be used in the book, as it is the term recognized by the Philippine government for the language based on Tagalog and enriched by various Philippine languages (Oue & Yoshizawa, 2017).

Finally, BOOK #8, which also uses “Tagalog” in its title, explicitly identifies Tagalog as its target language. The book provides a two-page description of Tagalog, stating it is the basis of the national language, Filipino. It explains the linguistic similarities between Tagalog and Filipino, noting they share the same syntax, morphology, and phonology, and describes the differences between the Pilipino alphabet, used by Tagalog, and the Filipino alphabet, used by Filipino. It also mentions that the term “Tagalog” is more commonly used by foreign-language learners and even by Filipinos themselves (Domigpe & Domingo, 2014).

In summary, despite the varied terminology used across the eight textbooks, at least five of them clearly refer to the national language of the Philippines, which is Filipino as their target language, while the



other textbooks recognize Tagalog as the basis of Filipino, and thus share the same structure and vocabulary. With this, it can be inferred that in the countries where these textbooks are utilized, teachers are teaching the same target language, and learners are being exposed to, or are currently studying, the same language.

### **3.2 Differentiating Vocabulary between “Beginner” and “Intermediate” Levels**

The next tables fulfill the second objective of this study which is to identify differences in vocabulary between these levels, contributing to a model of word level identification to answer the question: (b) are “beginner” and “intermediate” learners across countries exposed to the same vocabulary at these levels? by determining the “common beginner words” and the “common intermediate words.”

Table 4 shows the total number of words from the six textbooks which were tagged to be containing “beginner” vocabulary or the “beginner set.” Except for the BOOK #3, the total number words from each textbook exceeds 1000. BOOK #7 and BOOK #8 which claim to be for “beginner” to “intermediate” level learning, have almost twice the number of the other textbooks. The total number of unique words found in the six books is 4,889 words.

Table 4. Total Number of Words Gathered from the Glossary or Vocabulary List of “Beginner” Textbooks

| TEXTBOOK  | TOTAL NUMBER OF WORDS |
|---|-----------------------|
| <b>BOOK #1:</b> Tagalog for Beginners<br>(Barrios, 2014)  | 1,149                 |
| <b>BOOK #2:</b> Conversational Tagalog<br>(Ramos, 1985)   | 1,105                 |
| <b>BOOK #3:</b> Panimulang Pag-aaral ng<br>Wikang Filipino/Introductory Study<br>of the Filipino Language (Peregrino<br>et al., 2016)   | 455                   |
| <b>BOOK #6:</b> フィリピン語<br>( <i>Firipin-go</i> ‘Filipino’) (Oue &<br>Yoshizawa, 2017)  | 1,991                 |
| <b>BOOK #7:</b> 大学のフィリピン語<br>( <i>Daigaku no Firipino-go</i><br>‘Komprehensibong Tekstbuk ng<br>Filipino/University Filipino’ ‘A<br>Comprehensive Filipino Textbook’)<br>(Yamashita et al., 2018) | 2,321                 |
| <b>BOOK #8:</b> Elementary Tagalog<br>(Domigpe & Domingo, 2014)   | 1,369                 |
| <b>TOTAL NUMBER OF WORDS<br/>COMPARED</b>   | 4,889                 |
| <b>COMMON WORDS (appeared in 3<br/>out of 6 textbooks)</b>  | 894 (18.3%)           |

Upon comparing the words in the “beginner set,” a total of 894 words were identified as “common.” This represents only 18.3% of the overall

## Evaluating Consistency Across Beginner and Intermediate Filipino Textbooks for a Model of Word Level Identification

word count (4,889) derived from all six textbooks. In essence, when assessing the “common” words in relation to the total vocabulary of each textbook, as outlined in Table 5, the “common” words constitute approximately 32% to 71% of the respective total word counts. Even in the case of BOOK #3, which has the lowest total word count, it only encompasses 71% of the “common beginner words.” This suggests that “beginner” learners share only about 51% of their knowledge concerning “beginner” vocabulary, implying that the remainder of the vocabulary they encounter may have already surpassed the “beginner” level.

Table 5 shows the percentage of “common beginner” words in each textbook.

**Table 5. Percentage of “Common Beginner” Words in each Textbook of the “Beginner Set”**

| TEXTBOOK  | COMMON WORDS VS. TOTAL<br>NUMBER OF WORDS |
|---|---|
| <b>BOOK #1:</b> Tagalog for Beginners<br>(Barrios, 2014)  | 547 of 1,149 (48%)                        |
| <b>BOOK #2:</b> Conversational Tagalog<br>(Ramos, 1985)   | 627 of 1,105 (57%)                        |
| <b>BOOK #3:</b> Panimulang Pag-aaral ng<br>Wikang Filipino/Introductory Study<br>of the Filipino Language (Peregrino<br>et al., 2016) | 321 of 455 (71%)                          |
| <b>BOOK #6:</b> フィリピン語<br>( <i>Firipin-go</i> ‘Filipino’) (Que &<br>Yoshizawa, 2017)  | 697 of 1,991 (35%)                        |

| TEXTBOOK  | COMMON WORDS VS. TOTAL<br>NUMBER OF WORDS |
|---|---|
| <b>BOOK #7:</b> 大学のフィリピーノ語<br>( <i>Daigaku no Firpino-go</i><br>'Komprehensibong Tekstbuk ng<br>Filipino/University Filipino' 'A<br>Comprehensive Filipino Textbook')<br>(Yamashita et al., 2018) | 747 of 2,321 (32%)                        |
| <b>BOOK #8:</b> Elementary Tagalog<br>(Domigpe & Domingo, 2014)   | 668 of 1,369 (49%)                        |

While this observation aligns with the expected content of BOOKS #6, #7 and #8, which claim to cover “beginner” to “intermediate” level material, it raises questions about the significance and priority assigned to the non-“common beginner” vocabulary in other textbooks. Furthermore, it prompts an exploration into whether the inclusion of certain words and expressions in these textbooks was aligned with the needs of the target students during the design process.

Out of the 894 common words, 702 words appeared in all six textbooks. Table 6 shows the list of some of these words randomly chosen from the main list.

Table 6. Some of the Words which Appeared in All Six “Beginner Set” Textbooks

|                           |                               |                             |                           |                                |
|---------------------------|-------------------------------|-----------------------------|---------------------------|--------------------------------|
| <i>abril</i> ‘April’      | <i>daan</i> ‘road,’<br>‘path’ | <i>kapatid</i><br>‘sibling’ | <i>malinis</i><br>‘clean’ | <i>dagat</i> ‘sea’             |
| <i>agosto</i><br>‘August’ | <i>dalawa</i> ‘two’           | <i>kasi</i><br>‘because’    | <i>malungkot</i><br>‘sad’ | <i>huwebes</i><br>‘Thursday’   |
| <i>ako</i> ‘I’            | <i>dilaw</i><br>‘yellow’      | <i>ko</i> ‘my,’<br>‘mine’   | <i>mangga</i><br>‘mango’  | <i>lola</i> ‘grand-<br>mother’ |

## Evaluating Consistency Across Beginner and Intermediate Filipino Textbooks for a Model of Word Level Identification

|                                |  |                                   |                                 |                                     |
|--------------------------------|--|-----------------------------------|---------------------------------|-------------------------------------|
| <i>anim</i> 'six'              | <i>disyembre</i><br>'December'                     | <i>kotse</i> 'car'                | <i>manok</i><br>'chicken'       | <i>mamaya</i><br>'later'            |
| <i>apat</i> 'four'             | <i>dito</i> 'here'                                 | <i>kulay</i> 'color'              | <i>mantika</i><br>'cooking oil' | <i>marami</i> 'a lot'               |
| <i>araw</i> 'sun,'<br>'day'    | <i>doktor</i><br>'doctor'                          | <i>kuya</i> 'older<br>brother'    | <i>marso</i><br>'March'         | <i>martes</i><br>'Tuesday'          |
| <i>araw-araw</i><br>'everyday' | <i>enero</i><br>'January'                          | <i>libro</i> 'book'               | <i>masarap</i><br>'delicious'   | <i>mataas</i> 'tall,'<br>'elevated' |
| <i>aso</i> 'dog'               | <i>ng</i> 'object<br>focus<br>marker,'<br>'linker' | <i>lima</i> 'five'                | <i>matamis</i><br>'sweet'       | <i>mesa</i> 'table'                 |
| <i>asukal</i><br>'sugar'       | <i>gusto</i><br>'pseudo-verb<br>like or want'      | <i>linggo</i> 'week,'<br>'Sunday' | <i>mayo</i> 'May'               | <i>mga</i> 'plural<br>marker'       |
| <i>asul</i> 'blue'             | <i>hindi</i> 'no,'<br>'not'                        | <i>lolo</i><br>'grandfather'      | <i>mo</i> 'you,'<br>'your'      | <i>mura</i> 'cheap'                 |

---

Like the examination of the “beginner set” textbooks, the vocabulary lists of the five “intermediate set” textbooks were also subjected to comparison, and the count of “common” words was determined. Results indicate a total of 5,767 unique words across the five “intermediate set” textbooks, with 674 identified as “common.” This represents only 11.7% of the overall word count. When assessing the “common” words relative to the total vocabulary of each textbook, as illustrated in Table 7, the “common” words constitute approximately 19% to 41% of the respective total word counts. BOOK #4, with the lowest total word count, only encompasses 19% of the “common intermediate words.” This implies that “intermediate” learners share approximately 28% of their knowledge concerning “intermediate” vocabulary, suggesting that the remaining

vocabulary could already be beyond the “intermediate” level, or a big portion contains beginner words from BOOK #6, #7 and #8 which BOOK #4 and #5 did not include any more in their glossaries. Words like *aso* ‘dog’ and *pusa* ‘cat’ or *banyo* ‘toilet’ and *sala* ‘living room’ can all be found in the “common beginner” wordlist especially in BOOK #1, #2, #6, #7 and #8, but are not found in the “common intermediate” wordlist under BOOK #4 and #5. Like with the case of “common beginner” words, it can be assumed that the words and expressions included in these textbooks were chosen with the needs of the target students in mind during the development process which may not always align with the vocabulary needs of learners using other textbooks.

Table 7 shows the total number of unique words found in all five “intermediate set” textbooks and as well as the percentage of “common intermediate” words in each textbook.

**Table 7. Total Number of Words and Percentage of “Common” and “Intermediate” Words in each “Intermediate” Textbook**

| TEXTBOOK  | TOTAL NUMBER OF WORDS |
|---|-----------------------|
| <b>BOOK #4:</b> <i>Intermediate Tagalog</i><br>(Ramos & Goulet, 1982)             | 149 of 795 (19%)      |
| <b>BOOK #5:</b> <i>Intermediate Tagalog</i><br>(Barrios, 2015)                    | 438 of 1,888 (23%)    |
| <b>BOOK #6:</b> フィリピン語<br>( <i>Firipin-go</i> ‘Filipino’) (Oue & Yoshizawa, 2017) | 560 of 1,991 (28%)    |

## Evaluating Consistency Across Beginner and Intermediate Filipino Textbooks for a Model of Word Level Identification

| TEXTBOOK  | TOTAL NUMBER OF WORDS |
|---|-----------------------|
| <b>BOOK #7:</b> 大学のフィリピン語<br><i>(Daigaku no Firpino-go</i><br>‘Komprehensibong Tekstbuk ng<br>Filipino/University Filipino’ ‘A<br>Comprehensive Filipino Textbook’)<br>(Yamashita et al., 2018) | 615 of 2,321 (27%)    |
| <b>BOOK #8:</b> <i>Elementary Tagalog</i><br>(Domigpe & Domingo, 2014)  | 556 of 1,369 (41%)    |
| <b>TOTAL NUMBER OF WORDS COMPARED</b>   | 5,767                 |
| <b>COMMON WORDS (appeared in 3 out of 5 textbooks)</b>  | 674 (11.7%)           |

From all five textbooks, only 38 words appeared in all “intermediate set” textbooks. It was found that all these words are also found in the list of “common beginner” words. All 38 words are shown in Table 8.

**Table 8. Words which Appeared in All Five “Intermediate Set” Textbooks**

|                                   |   |                           |   |
|-----------------------------------|---|---------------------------|---|
| <i>ang</i> ‘subject marker’       | <i>buwan</i> ‘moon,’ ‘month’            | <i>lolo</i> ‘grandfather’ | <i>pamangkin</i> ‘nephew, niece’        |
| <i>araw</i> ‘sun,’ ‘day’          | <i>na</i> ‘two,’ ‘already’              | <i>at</i> ‘and’           | <i>panahon</i> ‘time/period,’ ‘weather’ |
| <i>ate</i> ‘older sister’         | <i>hindi</i> ‘no,’ ‘not’                | <i>luma</i> ‘old’         | <i>panganay</i> ‘eldest’                |
| <i>ayaw</i> ‘pseudo-verb dislike’ | <i>ho</i> ‘politeness marker’ ‘like po’ | <i>luya</i> ‘ginger’      | <i>pinsan</i> ‘cousin’                  |

|  |  |                                       |                                   |
|--|--|---------------------------------------|-----------------------------------|
| <i>bagoong</i><br>'fermented fish or<br>shrimp'    | <i>kita</i> personal<br>pronoun <i>ko + ka</i>       | <i>gulay</i> 'vegetable'              | <i>puti</i> 'white'               |
| <i>bagyo</i> 'typhoon'                             | <i>mabuti</i> 'good'                                 | <i>mahaba</i> 'long'                  | <i>mo</i> 'you,' 'your'           |
| <i>ng</i> 'object focus<br>marker,' 'linker'       | <i>loob</i> 'inside'                                 | <i>mataas</i> 'tall,'<br>'elevated'   | <i>tag-ulan</i> 'rainy<br>season' |
| <i>baon</i> 'money'<br>'food brought<br>from home' | <i>kulay</i> 'color'                                 | <i>matamis</i> 'sweet'                | <i>taon</i> 'year'                |
| <i>bawang</i> 'garlic'                             | <i>sa</i> 'place marker,'<br>'future time<br>marker' | <i>may</i> 'existential<br>verb have' |                                   |
| <i>bunso</i> 'youngest<br>sibling'                 | <i>lola</i><br>'grandmother'                         | <i>ako</i> 'I'                        |                                   |

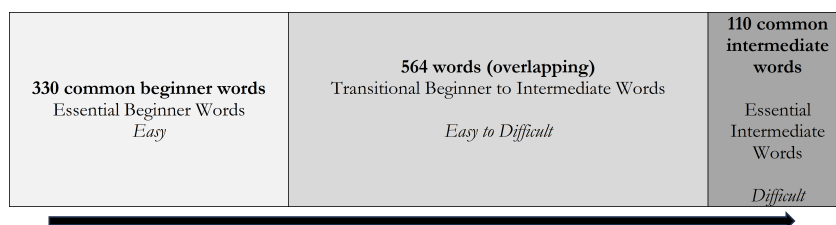
---

Next, the number of “common beginner” words (894 words) was compared with the “common intermediate” words (674 words). A total of 564 words were found to overlap or are common. This suggests that the “intermediate” vocabulary is not entirely distinct from the “beginner” vocabulary, instead, the “intermediate” vocabulary can be assumed to build on the “beginner” vocabulary with the 564 overlapping words as what was also pointed by the EVP or English Vocabulary Profile about the nature and relationship of A Level vocabulary and B Level vocabulary (English Profile, n.d.). Figure 2 presents a diagram, a continuum, distinguishing the “common beginner” vocabulary from the “common intermediate” vocabulary based on the analysis of the 564 overlapping words. The overlapping words are analyzed as transitional beginner to intermediate words with easy to difficult difficulty level. On the other hand, non-overlapping common beginner words are the essential beginner words that must be learned at the beginner level



## Evaluating Consistency Across Beginner and Intermediate Filipino Textbooks for a Model of Word Level Identification

and the non-overlapping common intermediate words are the essential intermediate words that are crucial to be learned during the intermediate level.



**Figure 2. Simple Description of Beginner and Intermediate Vocabulary**

It was not yet done in this study; however, a clearer description and distinction of the beginner and intermediate words can be done by analyzing the nature of the words found in each column of Figure 2. This is suggested as the next step of the study.

Finally, to further distinguish “common beginner” words from “common intermediate” words, the words were categorized by parts of speech and by themes. Here, the textbooks were consulted on how they tagged the words that can be categorized under two or more parts of speech and then these tags were counted. For example, if a beginner word is tagged as *noun* in all six textbooks, then it was counted as a *noun*. However, if a beginner word is tagged as *noun* in three textbooks and is tagged as *verb* in the other three textbooks then it was counted as a *noun*. Table 9 shows the categorization of the “common beginner” words.

Table 9. Categorization of “Common Beginner” Words

| NOUNS<br>(476)   |  | ADJECTIVES<br>(101)   | VERBS (161)  | OTHERS<br>(156)  |
|--|--|---|--|--|
| Plants; Food<br>(Fruits,<br>Vegetables,<br>Condiments,<br>other<br>Ingredients<br>for cooking;<br>Filipino food<br>names<br>(proper<br>names);<br>Drinks;<br>Utensils and<br>Tablewares;<br>Family<br>members;<br>Occupation;<br>People terms;<br>Parts of the<br>body; Illness;<br>Medicine;<br>Clothes;<br>Accessories | Days of the<br>week;<br>Months;<br>Numbers;<br>Shapes;<br>Colors;<br>School<br>materials;<br>Things inside<br>the house;<br>Parts of the<br>house; Places<br>in the<br>community;<br>Transportation;<br>Directions;<br>Names of<br>places<br>(Proper<br>names);<br>Landscapes;<br>Weather;<br>Disaster;<br>Location<br>/position | Temperature;<br>Size (people,<br>things);<br>Weight;<br>Height;<br>Speed;<br>Distance;<br>Condition of<br>things;<br>Attitude<br>/personality;<br>Feelings<br>(basic);<br>Feelings<br>( <i>nakaka-</i> );<br>Taste;<br>Comparison<br>( <i>kasing-</i> ) | Everyday<br>activities<br>( <i>mag-</i> , <i>ma-</i> ,<br><i>-um-</i> );<br>Cooking ( <i>-in</i> ,<br><i>i-</i> ); <i>Magpa-</i> ;<br>Direction<br>( <i>-um-</i> ) | Markers ( <i>ang</i> ,<br><i>mga</i> , <i>ng</i> , <i>sa</i> );<br>Linkers;<br>Clitics; <i>Po</i> ,<br><i>Ho</i> ; Time<br>(indefinite);<br>Time affixes<br>( <i>ala</i> , <i>alas</i> );<br>Pronouns;<br>Conjunctions<br>( <i>pero</i> , <i>dahil</i> ,<br><i>kasi</i> );<br>Question<br>words ( <i>+ba</i> );<br>Expressions<br>/greetings;<br>Pseudo-verbs;<br><i>Oo</i> , <i>Hindi</i> ;<br>Existential<br>verbs |

Table 10 presents the categorization of “common intermediate” words. As with the findings, there are more “common beginner” words (894)

than “common intermediate” words (674), so Table 9 “common beginner” words have more words under the columns noun, verbs and others. So, it can be said that distinguishing between “common beginner” and “common intermediate” words based solely on their part of speech is challenging. A closer look at the themes within each part of speech in both tables also reveals minimal differences between the two wordlists. The key distinction is that the “common intermediate” wordlist (Table 10) contains more themes within the verb category. For instance, Table 9 shows that the “common beginner” wordlist includes verbs with the affix *um* related to direction, such as *kumanan* ‘turn right,’ but lacks verbs with the affix *pa*, like *papunta* ‘going’ which can be found in the “common intermediate” wordlist. Another example would be about the verbs related to cooking. Both Table 9 and Table 10 include verbs related to cooking, such as *tadtarin* ‘to chop or to mince.’ However, only the “common intermediate” wordlist features *dikdikin* ‘to crush or to pulverize,’ suggesting that learners encounter the concept of chopping earlier than pulverizing. Additionally, there are fewer nouns in the “common intermediate” wordlist, but it does include abstract nouns related to beliefs, which are absent from the “common beginner” list. Examples of these are *malas* ‘bad luck’ and *suwerte* ‘good luck.’

Table 10. Categorization of “Common Intermediate” Words

| NOUNS<br>(343)  |  | ADJECTIVES<br>(105)  | VERBS<br>(107)   | OTHERS<br>(119)  |
|---|--|--|--|--|
| Animals;<br>Plants; Food<br>(Fruits,<br>Vegetables,<br>Condiments,<br>other<br>Ingredients<br>for cooking);<br>Filipino food<br>names<br>(proper<br>names);<br>Drinks;<br>Utensils and<br>Tablewares;<br>Family<br>members;<br>Occupation;<br>People terms;<br>Parts of the<br>body; Illness;<br>Medicine<br>Clothes;<br>Accessories; | Days of the<br>week;<br>Months;<br>Numbers;<br>Shapes;<br>Colors;<br>School<br>materials;<br>Things inside<br>the house;<br>Parts of the<br>house; Places<br>in the<br>community;<br>Transportation;<br>Directions;<br>Names of<br>places<br>(Proper<br>names);<br>Landscapes;<br>Weather;<br>Disasters;<br>Media; | Temperature;<br>Size (people,<br>things);<br>Weight;<br>Height;<br>Speed;<br>Distance;<br>Condition of<br>things;<br>Attitude<br>/personality;<br>Personality<br>(-in);<br>Feelings<br>(basic);<br>Feelings<br>( <i>nakaka-</i> );<br>Taste;<br>Texture<br>(touch,<br>food/eating);<br>Comparison<br>( <i>mas</i> );<br><i>Napaka-</i> | Everyday<br>activities<br>( <i>mag-</i> , <i>ma-</i> ,<br><i>-um</i> );<br>Cooking ( <i>-in</i> ,<br><i>i-</i> , <i>-an</i> );<br>Destroy ( <i>-in</i> );<br>Illness ( <i>ma-</i> );<br>Direction<br>( <i>-um-</i> , <i>pa-</i> );<br>Focus ( <i>-in</i> ,<br><i>-an</i> ) | Markers ( <i>ang</i> ,<br><i>mga</i> , <i>ng</i> , <i>sa</i> );<br>Linkers;<br>Clitics; <i>Po</i> ,<br><i>Ho</i> ; Time<br>(indefinite);<br>Time affixes<br>( <i>ala</i> , <i>alas</i> );<br>Pronouns;<br>Conjunctions<br>( <i>pero</i> , <i>dahil</i> ,<br><i>kasi</i> );<br>Question<br>words ( <i>+ba</i> );<br>Expressions<br>/greetings;<br>Pseudo-verbs;<br><i>Oo</i> , <i>Hindi</i> ;<br>Existential<br>verbs |

# Evaluating Consistency Across Beginner and Intermediate Filipino Textbooks for a Model of Word Level Identification

| NOUNS<br>(343)  | ADJECTIVES<br>(105) | VERBS<br>(107) | OTHERS<br>(119) |
|---|---------------------|----------------|-----------------|
| Special Occasions;<br>Beliefs;<br>Cultural activities | Location /position  |                |                 |

## 4 Conclusion and Recommendation

Despite the growing interest in learning the Philippine language and culture, there is still a notable lack of learning materials, particularly when compared to other languages. Additionally, there is no available proficiency assessment tool for learners who want or need it. This study sought to evaluate the consistency across beginner and intermediate Filipino textbooks to contribute in creating a model of Filipino word level identification by confirming the language referenced in Filipino textbooks used globally and differentiating vocabulary between “beginner” and “intermediate” levels. By comparing eight textbooks, it was revealed that despite variations in terminology, the majority of textbooks target the national language of the Philippines, Filipino. The findings suggest that learners across various countries are being exposed to the same language, though textbooks employ different names for it.

For the second objective, the glossaries or vocabulary lists of the eight textbooks were analyzed. From the six beginner-level labeled textbooks, 4,889 unique words were extracted, of which 894 words were common, representing 18.3% of the total number of words. The “common” words accounted for between 32% and 71% of the total vocabulary in each

textbook, suggesting that “beginner” learners share only about 51% of their knowledge concerning “beginner” vocabulary and the remaining vocabulary they are learning already exceeds their level. Similarly, the five intermediate-level labeled textbooks contained 5,767 unique words, with 674 common words identified. These “common intermediate” words represented between 19% and 41% of the total vocabulary in each textbook. This implies that “intermediate” learners share approximately 28% of their knowledge concerning “intermediate” vocabulary with the remaining words likely extending beyond the intermediate level or beginner words which were not mentioned anymore in two of the exclusively intermediate textbooks. These findings indicate that each textbook incorporates specific vocabulary tailored to its target learners, which may not always align with the vocabulary needs of learners using other textbooks.

The comparison between the categorization of “common beginner” and “common intermediate” words by parts of speech and themes further revealed that the intermediate textbooks cover a broader range of themes, particularly related to nouns and verbs. Notably, the nouns in the intermediate vocabulary list include abstract nouns not found in the beginner vocabulary.

Although the number of “common” beginner and intermediate words found is low, the vocabulary comparison showed significant overlap between beginner and intermediate levels, with intermediate vocabulary incorporating a substantial portion of beginner words. This overlap suggests a continuum in language acquisition, where intermediate learners continue building on foundational vocabulary while expanding into more complex thematic areas.

From this study, the author was able to create “common beginner” and “common intermediate” wordlists, which can serve as foundational resources for future developments. The ultimate goal of the study is to use these findings to develop a user-friendly word-level checker tool. Such a tool could become a valuable reference for creating teaching materials and proficiency assessments in Filipino language education.

The compilation of data for Data 2 which were analyzed, and wordlists can be viewed in this folder: <https://shorturl.at/m9cwO>

Future textbook development should aim for clearer distinctions between vocabulary levels to enhance learner progression. Additionally, standardizing terminology across textbooks would ensure consistency in language instruction globally. It is recommended that educators and material developers create resources that address the distinct needs of both beginner and intermediate learners, emphasizing the gradual introduction of more abstract and complex concepts at higher proficiency levels. Furthermore, the creation of a word-level checker tool or Filipino vocabulary profile could help educators and learners assess proficiency more accurately and tailor instruction to individual learning paths.

## **5 References**

- ACTFL. (2022). *Assigning CEFR ratings to ACTFL assessments*. [https://www.actfl.org/uploads/files/general/Assigning\\_CEFR\\_Ratings\\_To\\_ACTFL\\_Assessments.pdf](https://www.actfl.org/uploads/files/general/Assigning_CEFR_Ratings_To_ACTFL_Assessments.pdf)
- ACTFL. (2023). *ACTFL proficiency guidelines 2012: English*. <https://www.actfl.org/educator-resources/actfl-proficiency-guidelines/english>

- Almario, Virgilio S., & Kilates, Marne L. (2014). *Madalas itanong hinggil sa Wikang Pambansa* [Frequently asked questions on the National Language]. Komisyon sa Wikang Filipino.
- Barrios, Joi. (2014). *Tagalog for beginners: An introduction to Filipino, the National Language of the Philippines (online audio included)*. Tuttle Publishing.
- Barrios, Joi. (2015). *Intermediate Tagalog: Learn to speak fluent Tagalog (Filipino), the National Language of the Philippines* [Downloadable material included]. Tuttle Publishing.
- Byrnes, Heidi. (1990). *Foreign language program articulation from high school to the university*. ERIC Digest.
- Constantino, Pamela. (2012). Wikang Filipino bilang konsepto. In *Salindaw: Varayti at baryasyon ng Filipino*. Sentro ng Wikang Filipino.
- Council of Europe. (2023). The framework: The CEFR levels. Common European Framework of Reference for Languages (CEFR). <https://www.coe.int/en/web/common-european-framework-reference-languages/level-descriptions>
- Domigpe, Jiedson R., & Domingo, Nenita Pambid. (2014). *Let's speak Tagalog! Tara, mag-Tagalog tayo! Come on, let's speak Tagalog! (online audio download included)*. Tuttle Publishing.
- Ellis, Rod. (1997). *Second language acquisition*. Oxford University Press.
- English Profile. (n.d.). English vocabulary profile [Retrieved September 30, 2024]. <https://www.englishprofile.org/wordlists>
- Interagency Language Roundtable. (n.d.). An overview of the history of the ILR language proficiency skill level descriptions and scale



## Evaluating Consistency Across Beginner and Intermediate Filipino Textbooks for a Model of Word Level Identification

- by Dr. Martha Herzog [Retrieved September 30, 2024]. <https://www.govtilr.org/Skills/IRL%20Scale%20History.htm>
- Japan Foundation. (2010). JF standard for Japanese-language education. [https://ny.jpj.go.jp/karashi/wp-content/uploads/jfs2010\\_all\\_en\\_small.pdf](https://ny.jpj.go.jp/karashi/wp-content/uploads/jfs2010_all_en_small.pdf)
- Krashen, Stephen. (1982). *Principles and practice in second language acquisition*. Pergamon Press.
- Laranjo, Ronel O. (2020). Mapping Philippine studies in North East Asia: A SWOT analysis of Southeast Asian Studies Programs from China, Japan, and Korea. *SUVANNABHUMI*, 12(1), 111–130.
- Laranjo, Ronel O., & Palma Gil, Florinda Amparo A. (2023). Preliminaryong paglalapat ng CEFR at TLLFTF sa Filipino bilang ikalawang wika tungo sa pagbuo ng istandardisadong kurikulum at pagsusulit [Preliminary Application of CEFR and TLLFTF to Filipino as a Second Language Towards the Development of a Standardized Curriculum and Test]. *Kawing*, 7, 18–42.
- McFarland, Curtis D. (1989). *A frequency count of Filipino*. Linguistic Society of the Philippines.
- Mizumoto, Atsushi, Pinchbeck, Geoffrey G., & McLean, Stuart. (2021). Comparisons of word lists on new word level checker [Comparisons of word lists on new word level checker]. *Vocabulary Learning and Instruction*, 10(2), 30–41.
- Oue, Masanao, & Yoshizawa, Jenny. (2017). *Osaka University School of Foreign Studies world languages Series 6: Filipino*. Osaka University Press.

- Peregrino, Jovy, Kim, Dongyeob, Enriquez, Ma. Althea, & Laranjo, Ronel. (2016). *Panimulang pag-aaral ng wikang Filipino. (Basic Filipino Language) interactive book application (app) format*. Busan University of Foreign Studies.
- Ramos, Teresita V. (1985). *Conversational Tagalog: A functional-situational approach*. University of Hawaii Press.
- Ramos, Teresita V., & Goulet, Rosalina Morales. (1982). *Intermediate Tagalog: Developing cultural awareness through language*. University of Hawaii Press.
- Ramos, Teresita V., & Mabanglo, Ruth. (2012). *Southeast Asian language teaching: The language learning framework for teachers of Filipino, special issue*.
- Saville-Troike, Muriel, & Barto, Karen. (2016). *Introducing second language acquisition*.
- Schmidt, Helmut. (1994). Probabilistic part-of-speech tagging using decision trees. *Proceedings of International Conference on New Methods in Language Processing*.
- Takahata, Sachi. (2022). Filipino: A nationwide migrant language and culture. In J.C. Maher (Ed.), *Language communities in Japan*. Oxford University Press.
- The Japan Foundation & Japan Educational Exchanges and Services. (2023, July 1). CEFR level to be added for reference to JLPT score reports. <https://www.jlpt.jp/e/cefrlevel/index.html>
- Tono, Yukio. (2019). Coming full circle - from CEFR to CEFR-J and back. *CEFR Journal Research and Practice*.

- Tupas, Ruanni. (2014). The politics of ‘p’ and ‘f’: A linguistic history of nation-building in the Philippines. *Journal of Multilingual and Multicultural Development*, 36 (6), 587–597.
- Yamashita, Michiko, Casel, Leith, & Takano, Kunio. (2018). 大学のフィリピン語 (*daigaku no firipino-go*) [Komprehensibong Tekstbuk ng Filipino (A Comprehensive Textbook of Filipino)]. Tokyo University of Foreign Studies University.
- Yap, Fe Aldave. (2012). *Global flipino crossing borders*. De La Salle University Publishing House.

## **6 Appendices**

### **6.1 CEFR Ratings Assigned to ACTFL Assessments (ACTFL, 2022)**

| <b>Rating on ACTFL<br/>Assessment</b> | <b>Corresponding<br/>CEFR Rating<br/>Receptive Skills –<br/>Reading and<br/>Listening</b> | <b>Corresponding<br/>CEFR Rating<br/>Productive Skills –<br/>Speaking and<br/>Writing</b> |
|---------------------------------------|---|---|
| Distinguished                         | C2  |   |
| Superior                              | C1.1  | C2  |
| Advanced High                         | C1.1  | C1  |
| Advanced Mid                          | B2  | B2.2  |

| <b>Rating on ACTFL<br/>Assessment</b> | <b>Corresponding<br/>CEFR Rating<br/>Receptive Skills –<br/>Reading and<br/>Listening</b> | <b>Corresponding<br/>CEFR Rating<br/>Productive Skills –<br/>Speaking and<br/>Writing</b> |
|---------------------------------------|---|---|
| Advanced Low                          | B1.2  | B2.1  |
| Intermediate High                     | B1.1  | B1.2  |
| Intermediate Mid                      | A2  | B1.1  |
| Intermediate Low                      | A1.2  | A2  |
| Novice High                           | A1.1  | A1  |
| Novice Mid                            | o   | o   |
| Novice Low                            | o   | o   |
| o                                     | o   | o   |

## 6.2 Comparison of the Expected Learning Objectives of BOOK #2 and BOOK #3

| Summary of Expected Attainments from BOOK #2: Conversational Tagalog (Ramos, 1985, pp. xiv–xv)  | English Translation of a Part of the Introduction of BOOK #3: Panimulang Pag-aaral ng Wikang Filipino/Introductory Study of the Filipino Language (Peregrino et al., 2016)  |
|---|---|
| <p><b>Functional Skills:</b><br/>Students will be able to handle simple conversations in common social contexts, including:</p> <ol style="list-style-type: none"> <li>1. Greetings and leave-takings.</li> <li>2. Introducing themselves and others.</li> <li>3. Expressing likes and dislikes.</li> <li>4. Telling time and giving dates using Spanish and Tagalog.</li> <li>5. Expressing simple discomforts.</li> <li>6. Narrating daily activities and future plans.</li> <li>7. Asking for and giving directions.</li> <li>8. Ordering food and meals.</li> </ol> <p><b>Linguistic Skills:</b><br/><i>Structure:</i> Students will gain proficiency in basic grammatical structures, including:</p> | <p>This book aims to help Koreans learn the basic knowledge of the Filipino language and use it in the context of both Korean and Filipino culture. The book presents activities that will teach students to speak, write, and understand the Filipino language for meaningful communication. The activities designed in this book will help students:</p> <ol style="list-style-type: none"> <li>1. Understand the background and nature of the Filipino language;</li> <li>2. Pronounce Filipino words correctly while speaking and reading texts;</li> <li>3. Use simple and basic sentence structures of the Filipino language in conversation and writing;</li> <li>4. Understand and correctly use basic words or vocabulary of the Filipino language; and</li> </ol> |

| Summary of Expected Attainments from BOOK #2: Conversational Tagalog (Ramos, 1985, pp. xiv–xv)  | English Translation of a Part of the Introduction of BOOK #3: Panimulang Pag-aaral ng Wikang Filipino/Introductory Study of the Filipino Language (Peregrino et al., 2016)  |
|---|---|
| <ol style="list-style-type: none"><li>1. Basic statements and question patterns (affirmative and negative).</li><li>2. Use of question words.</li><li>3. Basic actor and object focus verbs.</li><li>4. Pseudo-verbs and existentials.</li><li>5. Pronouns, demonstratives, and markers for phrases.</li><li>6. Simple modifications and adverbial phrases.</li><li>7. Sentence connectors and linkers.</li></ol> | <ol style="list-style-type: none"><li>5. Express simple thoughts and feelings in daily activities, situations, and interactions using the Filipino language.</li></ol>  |
| <p><b>Vocabulary Skills:</b></p> <p>Students will acquire common nouns for objects, people, shapes, colors, and food, along with basic actor and object focus verbs.</p>  | <p>The book will cover and include the following linguistic activities and exercises:</p> <ol style="list-style-type: none"><li>1. Background and nature of the Filipino language</li><li>2. Sounds and alphabet of the Filipino language</li><li>3. Common expressions for greetings and introductions</li><li>4. Basic descriptions of people, things, and places</li><li>5. Basic knowledge of recognizing numbers and correct counting, telling time, and basic knowledge of purchasing and using money</li></ol> |
| <p><b>Pronunciation and Comprehension Skills:</b></p> <p>They will approximate key sounds like glottal stops, nasal sounds, and initial stops. They will also understand and answer simple questions about daily routines and family topics.</p>  | <ol style="list-style-type: none"><li>6. Primary pronouns</li><li>7. Basic family relationships</li><li>8. Expressing likes and dislikes</li><li>9. Basic knowledge of location and simple verbs</li></ol>  |

## Evaluating Consistency Across Beginner and Intermediate Filipino Textbooks for a Model of Word Level Identification

|  |   |
|--|---|
| <b>Summary of Expected Attainments from BOOK #2: Conversational Tagalog (Ramos, 1985, pp. xiv–xv)</b>      | <b>English Translation of a Part of the Introduction of BOOK #3: Panimulang Pag-aaral ng Wikang Filipino/Introductory Study of the Filipino Language (Peregrino et al., 2016)</b> |
| <hr/>  |   |
| This section focuses on practical conversational skills, emphasizing both linguistic and cultural fluency. |   |

### Acknowledgements

The author would like to thank Teru Watanabe for his invaluable help in cleaning the data and for creating a simple program in R studio used in comparing the vocabulary lists gathered in this study. Moreover, this research has been completed using the Grants-in-Aid for Scientific Research (Fundamental Research (A)) Project Number 22H00076 (Research Representative: Yukio Tono).