

Developing Multilingual Automotive E-Dictionary Based on Corpus Linguistics

*Yuliana Ningsih, Adhiela Noer Syaief, Kurnia Dwi Artika, Herpendi Herpendi

Department of Agroindustry, Automotive Technology, Computer & Network Engineering Technology
Politeknik Negeri Tanah Laut
South Kalimantan, Indonesia
*yuliananingsih@politla.ac.id

Abstract— The study aims to compile a dictionary that makes students, Indonesian workers, especially mechanics, supervisors in the automotive sector who work or study in the automotive industry need an automotive dictionary. The Design and Development (D&D) model used in this study and obtaining data were taken from automotive books, automotive websites for technical language specifically for automotive engines, and additional data from official and non-official workshops for workshop languages commonly used daily. This study used the Antconc application for automotive text corpus processing. This dictionary has three languages: English (automotive)-Indonesian-local (workshop language), which were arranged to start from the first alphabet that is pronounced. The product's result was a multilingual automotive e-dictionary which several experts validate categorized as a "valid and excellent" product. The multilingual automotive e-dictionary group in the experimental group was more effective; 0.6 than the control group; 0.3. The Cronbach Alpha value from six variables: content, format, accuracy, ease of use, timeliness and user satisfaction was greater than 0.60 (the existing provisions). It showed that the variable was "reliable" as a measure of reliability test.

Keywords—*automotive; E-dictionary; multilingual; corpus; linguistics*

I. INTRODUCTION

Vocabulary has a significant role in everyday life, work, and even in language learning. The attention to vocabulary learning has increased, and the need for an approach to vocabulary learning to

encourage vocabulary enhancement and emphasize comprehension. The dictionary has much information, has excellent benefits, and is an excellent forum for knowledge, especially language knowledge. The dictionary is also a means to enrich vocabulary for everyone, making it easier to express ideas through written and spoken language.

In either print or electronic form, a reference source provides words that are usually alphabetically sorted and information about the words' conditions and pronunciations, functions, etymologies, meanings, syntactic and idiomatic usage, and other relevant information. [1].

The automotive dictionary is indispensable for automotive students, automotive engineers, and even mechanical repair shops. A large number of terminology dictionaries are now available in print and online [2]. Also stated, an increasing number and variety of books or dictionaries are available today, both online and in print. However, there is currently no unique tool (glossary, reference material on the technical translation) that can convey the knowledge correctly. The automotive dictionary is indispensable for automotive students, automotive engineers, and even mechanical repair shops. A large number of terminology dictionaries are now available in print and online.

Students and regular mechanics prefer to use the workshop's language, which is used daily, without knowing the automotive engineering language. Moreover, the availability of a dictionary of terminology for the workshop's language, which is usually used in the workshop, is still limited. In response, the researcher took the initiative to develop a multilingual automotive dictionary called the workshop Language. This dictionary aims to help automotive students in significant schools and

colleges, workshop mechanics, and other learners easily express automotive language and workshop language. Also stated, an increasing number and variety of books or dictionaries are available today, both online and in print. However, there is currently no unique tool (glossary, reference material on the technical translation) that can convey the knowledge correctly. The automotive dictionary is indispensable for automotive students, automotive engineers, and even mechanical repair shops. A large number of terminology dictionaries are now available in print and online.

The development of this multilingual automotive dictionary based on corpus linguistics. Corpus linguistics is a statistically retrieved language database for investigation, description, application, and analysis relevant to linguistics. [3].

In this corpus, data collected in a written form, either in the form of ordinary or digital data, contains various linguistic information, starting at the word level and progressing through structure, meaning, and discourse, which may be used for research purposes. [4].

The corpus produced by digitization can be used in various linguistic fields, such as morphology, syntax, semantics, lexicology, or other linguistic fields, such as translation, literature, etc. [4].

Based on this explanation, corpus linguistics can be used to study and carry out various language studies. In this study, the researcher attempted to describe the optimization of the corpus linguistic approach in developing a multilingual e-automotive dictionary[5][6]. Fadila stated that the research as a benchmark used a mobile dictionary to find the meaning of new words in the classroom and outside the classroom [7]. Also stated that the research results are used as a source of research to be carried out that students' vocabulary knowledge was affected by the application of an electronic dictionary, and students with a high reading interest were more adapted to using an electronic dictionary than students with such a low reading interest. Alhatmi also stated that the study that used technology-based (digital) dictionaries than when using a traditional dictionary and using of dictionary mainly to comprehend new terms while reading [8]. Moreover, the research results were used as a source of research to be carried out that the descriptive method reveals the web-based material processing mechanism factually into the Arabic-Indonesian medical dictionary corpus format and compiles a word list [9].

Based on the background mentioned earlier, this study aims to develop a multilingual automotive dictionary for engineering students; it is also possible that this dictionary can be used by the repair shop,

either official or non-official. The dictionary provides words in three languages, namely English, Indonesian, and the language used in the workshop daily. The dictionary was developed for specialized English learners from machinery, be it college or school students. This dictionary is also used to make it easier for machine teachers or lecturers so that students can easily recognize the proper Engineering Language in machines.

This research aims to make it easier for students in the automotive sector, Indonesian workers, especially mechanics, supervisors, and those working or studying in the automotive sector who need an automotive dictionary.

II. CONTEXTUAL FOUNDATION

2.1 Dictionary

Students' vocabulary mastery can be improved through dictionary as media [8].

Students find it easier to improve vocabulary mastery by using a dictionary that provides a source of words, meanings and pronunciation [6]. They also stated that A dictionary is a well-known learning tool that provides a significant role in the context of learning a new language. It has been shown to help learn vocabulary and developing language proficiency.

Dictionary is used to find meaning, looking for translation, synonym, and paraphrasing [10][11]. Currently, the dictionary the development of electronic dictionaries has added new resources in learning vocabulary because the use of dictionaries is beneficial in learning vocabulary [12].

Electronic dictionaries are very helpful for students in learning new vocabulary and are an engaging, fun tool to use [5].

There has been rapid development in the last two decades in learning new vocabulary using and utilizing dictionary and electronic and e-dictionary [13][14][15][16].

The existing literature on vocabulary learning strategies illustrates that language learners can learn new vocabulary independently using a dictionary [12].

2.2 Corpus Linguistic

Compared to most other linguists, the person who creates the dictionary has a stronger and more immediate need for a corpus and has thus been at the forefront of corpus development for a long time. The user corpus tool has significantly improved since it was first used in dictionary projects [17].

A corpus is a set of documents and samples of words from various origins, as well as documents of various text formats, that are used to create a lexicon. In the case of a lexicographic that examines examples of a word, the type of text from which a particular example comes, i.e., from which document, and characteristics of that document, such as the date of publication, author, mode (pronounced or written), domains, and so on, from the text, may be of interest. Every form must be provided with information in order to function properly. This metadata is typically found in the 'header,' which defines the characteristics of the document in such a way that it can be interpreted by a corpus the instrument that can interpret it.[18].

III. METHOD

The application method is Design and Development (D&D) models and based on the English text-based corpus processing application, especially the automotive engineering language, will be used in this study [19]. Stated that the systematic stage is the preparation phase, data and material collection, data and material processing or data processing, material evaluation or data and material assessment, and dictionary printing preparation [20]. Moreover, a descriptive method in which the factual web-based processing technique:

1. The author uses the AntConc tool to construct a word list derived from articles and automotive news found on the internet, then download it into Microsoft Word and save it in the format (*.doc).
2. This procedure re-formats the text downloaded into the form (*.txt) by converting it from the UTF-8 model to the save-as process in Microsoft Word.
3. Manually re-checking the corpus materials to ensure that only those related to the automotive sector are included.
4. Transfer it to Microsoft Excel; a parallel dictionary was constructed to verify the meaning of words in the text.
5. Then, using the N-gram format provided by the AntConc application, add lemma/entries in the form of basic vocabulary terms that have not yet been included in the file after performing data cleaning with particular

emphasis to basic vocabulary, derivative vocabulary, and affix words.

6. To revise data and convert it to a * doc format file by the template that has been generated.
7. It is necessary to search for an equivalent word in the draft dictionary in *.doc format.

Then, the dictionary was assessed for its validity by three experts. The first expert is an expert in English and mechanic who uses workshop language. The next expert is a lexicographer, and the Gregory formula used to analyze the results of content validity and quality judgment was investigated using guidance provided [21].

The data on the effectiveness of the multilingual automotive e-dictionary was obtained from the difference between the post-test and pre-test scores. The assessment was obtained from the answers to the questions given with 25 questions with measuring instruments in this study using a questionnaire or a google form link.

Furthermore, the results of the test scores were analyzed by looking for a Gain score. The Gain score is also known as an increase or difference in pre-test and post-test scores. Gain score data analysis can show the achievement of increasing user abilities by considering their initial abilities, and the calculation results of the N-Gain can show the effectiveness of the developed dictionary. The N-Gain formula according to Meltzer.

IV. RESULTS AND DISCUSSION

This research results in developing an automotive dictionary as a medium for Automotive and English for Specific Purposes. Compiling this dictionary makes it easier for students and Indonesian workers, especially mechanics, supervisors in the automotive sector who work or study in the automotive industry who need an automotive dictionary. Expert in lexicography, a material expert, and two consumers of the product, evaluated the product's validity after it was assembled (teachers and students). When developing the dictionary, the experts involved aspects such as (1) entry type, (2) definition, (3) simplicity, (4) readability, and (5) appropriateness for the student's skill level.

It was discovered that there were 570 vocabularies related to the automotive sector as a result of the first stage's data collection and analysis using corpus linguistics techniques. It is possible to collect this information from a data analysis performed with the sketch engine application's wordlist study feature. Following the second stage of data collection and processing, which included editing and cleaning the data as well as adding vocabulary, it was discovered that 426 vocabularies were comprised

of essential words as well as derived words in the form of verbs, nouns, adjectives, and other terms related to the automotive sector. Four hundred twenty-six vocabularies that will later be included in the draft dictionary in the form of *.doc and then look for the equivalent by the guidebook for developing this dictionary are described below.

In developing a multilingual automotive dictionary, the first draft was designed around two themes: motorcycles and vehicles, with terms in English, Indonesian, and Workshop Languages. Additionally, it includes a picture of each phonetic alphabet. The following step is to improve the draft and develop it into a finished product. Seven sections comprise the dictionary: the dictionary cover, prologue, acknowledgments, abbreviations, the dictionary body, including the definitions and glossary, biographical sketch, and synopsis. There are a total of 67 pages in the dictionary. After that, the dictionary was evaluated by three experts based on its content validity. The first expert is a specialist in Lexicography, as well as in automotive and English (linguist). The products of this dictionary were pocket book (printed), E-book, and play store application.

The findings of the content validity test were examined with the use of the Gregorian equation. Content validity testing has revealed that the dictionary is relevant to the five requirements for an excellent multilingual dictionary, as determined by the test results. In several sections of the lexicon, however, there are still some adjustments to be made. The first half of the sentence contains the term equivalent. Fact that there are multiple equivalent words in each of the three languages, the usage of capital letters impacts the creation of the dictionary. This dictionary will make it easier for everyone in the automotive sector to acquire technical languages primarily derived from English and have been translated into two languages, namely Indonesian and workshop languages, through repetition. Because many technical phrases are utilized in everyday life, or what we refer to as "workshop language," this dictionary has a distinct local sense in the language of workshops worldwide, particularly in South Kalimantan. As a result, there must be some level of coordination between them.

After revising some parts of the dictionary, three experts judged it for the validity of the product stated that the references validating dictionaries of experts in various fields (linguists, lexicographers, information scientists) [22]. This product validated by linguist, lexicographer, and automotive expert. The result of the validity result is shown in Table 1.

TABLE I. VALIDITY RESULT

Aspects	Average (%)	Criteria
Entry type	75.15	Valid enough
Definition	85.20	Very valid

Simplicity	85.10	Very valid
Readability	85.15	Very valid
Appropriate for users	86.05	Very valid

Based on Table 1, it was known that all validated aspects are categorized as very valid for definition, simplicity, legibility, suitable for users, and type of entry as valid enough. There were several minor revisions before use in the small group test.

Five users judged the overall quality of the dictionary as well. As in Table 2, the result of the quality judgment showed.

TABLE II. THE RESULT OF QUALITY JUDGMENT

Users	Score	Category
User 1	54.5	Excellent
User 2	53	Excellent
User 3	54	Excellent
User 4	53	Excellent
User 5	53.5	Excellent

The results explained that the research results are used as a source of research to be carried out that dictionary of multilingual thematic pictures are categorized as highly valid (excellent) [23].

The effectiveness test was conducted on 20 users for the control group or comparison group with conventional automotive terminology and 20 users for the experimental group by applying multilingual automotive e-dictionary. The results were obtained by giving pre-test and post-test to the control class and experimental class. The control class used an ordinary dictionary at the study, while the experimental class used a multilanguage automotive e-dictionary assisted by two dictionaries; an e-book dictionary and the multilingual automotive dictionary in the play store application. The operational data of scoring with the N-Gain calculation from the results of the pre-test and post-test calculations were in the following table:

TABLE III. N-GAIN SCORE

Group	Pre-test	Post-test	N-gain	Criteria
Control Group	35	55	0.3	Low
Experiment Group	15	70	0.6	Medium

The data calculation above showed that the multilingual automotive e-dictionary group in the experimental group was more effective than the control group.

TABLE IV. RELIABILITY TEST

Variable	Cronbach Alpha	Criteria
Content	0.626	Reliable
Accuracy	0.612	Reliable
Format	0.726	Reliable
Ease of Use	0.882	Reliable

Timeliness	0.732	Reliable
User Satisfaction	0.825	Reliable

The table above showed the Cronbach Alpha value from six variables: content, accuracy, format, ease of use, timeliness, and user satisfaction was greater than 0.60 (the existing provisions). It showed that the variable was "reliable" as a measure of reliability test.

V. CONCLUSION

This study developed the product that was a multilingual automotive e-dictionary for the students, Indonesian workers, especially mechanics, supervisors in the automotive sector who work or study in the automotive industry who need an automotive dictionary. The dictionary was evaluated by three experts and five users which the result showed that the dictionary was categorized as a "valid and excellent" product. The multilingual automotive e-dictionary group in the experimental group was more effective; 0.6 than the control group; 0.3. The Cronbach Alpha value from six variables: content, accuracy, format, ease of use, timeliness and user satisfaction was greater than 0.60 (the existing provisions). It showed that the variable was "reliable" as a measure of reliability test. It is suggested for the students and teacher to use the dictionary in the teaching-learning process. Moreover, it is suggested for other researchers to continue this study by conducting field implementation and developing the dictionary into a better product.

ACKNOWLEDGMENTS

This dictionary was made possible by the generous support of DIPA Research Grant POLITALA (Politeknik Negeri Tanah Laut). Some knowledge input, editing, review, and validity of this dictionary could not have been accomplished without the collaboration of many personnel that aid as follows: Director Politeknik Negeri Tanah Laut. Department of Automotive Technology Politeknik Negeri Tanah Laut, Balai Bahasa Banjarbaru South Kalimantan, State Vocational High School 2 Pelaihari, State Vocational High School 1 Takisung, Yamaha Pelaihari, Wira Toyota Banjarmasin, Mitra Suzuki Banjarmasin, Harindo Workshop Pelaihari, and Adi Workshop Pelaihari.

REFERENCES

[1] "Dictionary by Merriam-Webster: America's most-trusted online dictionary." <https://www.merriam-webster.com/>

- (accessed Sep. 17, 2021).
- [2] R. Billero and M. C. N. Martínez, "Nuove risorse per la ricerca del lessico del patrimonio culturale corpora multilingue LBC," *undefined*, 2017.
- [3] N. S. Dash, "Corpus linguistics: an introduction," p. 175, 2008, Accessed: Sep. 17, 2021. [Online]. Available: https://books.google.com/books/about/Corpus_Linguistics.html?hl=id&id=ikOZWjkxzEC.
- [4] N. HIZBULLAH, F. FAZLURRAHMAN, and F. FAUZIAH, "LINGUISTIK KORPUS DALAM KAJIAN DAN PEMBELAJARAN BAHASA ARAB DI INDONESIA," *Pros. Konf. Nas. Bhs. Arab*, vol. 1, no. 2, 2016, Accessed: Sep. 17, 2021. [Online]. Available: <http://prosiding.arab-um.com/index.php/konasbara/article/view/84>.
- [5] K. A. Barham, "The Use of Electronic Dictionary in the Language Classroom: The Views of Language Learners.," *Online Submiss.*, Mar. 2017.
- [6] W. T. M. M. F. A. E. Abbasi, "Learners' Perceptions of Monolingual Dictionaries in Learning English as a Foreign Language.," *Int. J. Educ. Lit. Stud.*, vol. 7, no. 3, pp. 10–18, Jul. 2019, Accessed: Sep. 17, 2021. [Online]. Available: <http://dx.doi.org/10.7575/aiac.ijels.v.7n.3p.10>.
- [7] S. F. Fadila, "The Effect of Electronic Dictionary and Reading Interest on Students' Vocabulary Knowledge (A Quasi-experimental Study at the Eighth Grade of SMP Al Falah, Mampang Prapatan)," Jan. 2019, Accessed: Sep. 17, 2021. [Online]. Available: <https://repository.uinjkt.ac.id/dspace/handle/123456789/44022>.
- [8] S. Alhatmi, "A Survey Study of the Dictionary Use Substrategies of English Majors in Saudi Arabia: Dictionary Related Aspects," *English Lang. Teach.*, vol. 12, no. 3, p. 139, Feb. 2019, doi: 10.5539/ELT.V12N3P139.
- [9] S. F. Azzahra, "PENYUSUNAN KAMUS KEDOKTERAN ARAB-INDONESIA DENGAN PENDEKATAN LINGUISTIK KORPUS," *TSAQOFIYA Jur. Pendidik. Bhs. Arab IAIN Ponorogo*, vol. 2, no. 2, pp. 60–66, Sep. 2020, doi: 10.21154/TSAQOFIYA.V2I2.37.
- [10] B. Nematollahi, F. Behjat, and A. Kargar, "A Meta-Analysis of Vocabulary Learning Strategies of EFL Learners," *English Lang. Teach.*, vol. 10, no. 5, p. p1, Apr. 2017, doi: 10.5539/ELT.V10N5P1.
- [11] K. J. Tsai, "Corpora and dictionaries as learning aids: inductive versus deductive approaches to constructing vocabulary knowledge," *Comput. Assist. Lang. Learn.*, vol. 32, no. 8, pp. 805–826, Nov. 2019, doi: 10.1080/09588221.2018.1527366.
- [12] A. Asgari and G. Mustapha, "The Type of Vocabulary Learning Strategies Used by ESL Students in University Putra Malaysia.," *undefined*, 2011.
- [13] S. KNIGHT, "Dictionary Use While Reading: The Effects On Comprehension and Vocabulary Acquisition For Students Of Different Verbal Abilities," *Mod. Lang. J.*, vol. 78, no. 3, pp. 285–299, Sep. 1994, doi: 10.1111/J.1540-4781.1994.TB02043.X.
- [14] B. Laufer and J. Hulstijn, "Incidental vocabulary acquisition in a second language: the construct of task-induced involvement," *Appl. Linguist.*, vol. 22, no. 1, pp. 1–26, Mar. 2001, doi: 10.1093/APPLIN/22.1.1.
- [15] L. Li and H. Xu, "Using an Online Dictionary for Identifying the Meanings of Verb Phrases by Chinese EFL Learners," *Lexikos*, vol. 25, no. 0, pp. 191–209, Nov. 2015, doi: 10.4314/lex.v25i0.
- [16] J. Stirling, "The Portable Electronic Dictionary: Faithful friend or faceless foe?," *undefined*, 2005.
- [17] A. Kilgarriff and I. Kosem, "Corpus tools for lexicographers," *Electron. Lexicogr.*, Jan. 2013, doi: 10.1093/ACPROF/OSO/9780199654864.003.0003.
- [18] A. Stefanowitsch, "Corpus linguistics: a guide to the methodology," 2015, Accessed: Sep. 17, 2021. [Online].

- Available:
https://books.google.com/books/about/Corpus_linguistics.html?hl=id&id=3ZHeDwAAQBAJ.
- [19] R. C. Richey and J. D. Klein, "Design and Development Research," *Handb. Res. Educ. Commun. Technol. Fourth Ed.*, pp. 141–150, Jan. 2014, doi: 10.1007/978-1-4614-3185-5_12.
- [20] S. J. Schierholz, "Methods in Lexicography and Dictionary Research," *Lexikos*, vol. 25, no. 0, pp. 323–352, Nov. 2015, doi: 10.4314/lex.v25i0.
- [21] J. Pallant, "SPSS Survival Manual : A Step by Step Guide to Data Analysis Using IBM SPSS," Jul. 2020, doi: 10.4324/9781003117452.
- [22] J. Ranitovic, A. Kavagic, V. Nikolic, L. Velimirovic, and M. Stankovic, "A theoretical approach in designing Bilingualized electronic dictionary and its importance in ESP teaching," *Mediterr. J. Soc. Sci.*, vol. 5, no. 13 SPEC. ISSUE, pp. 239–243, 2014, doi: 10.5901/MJSS.2014.V5N13P239.
- [23] N. M. Suniyasih, N. M. Ratminingsih, and I. G. Budasi, "Development of Multilingual Thematic Picture Dictionary: A Support for Literacy," *J. Pendidik. dan Pengajaran*, vol. 53, no. 3, p. 236, Sep. 2020, doi: 10.23887/JPP.V53I3.27508.