

# Spelling Errors in Bahasa of Nahdlatul Ulama Surabaya University Students

Nadya Afdholy<sup>1</sup>, Terra Bellatrix Aden Nashahta<sup>2\*</sup>

<sup>1</sup>Universitas Negeri Surabaya

<sup>2\*</sup>Universitas Nahdlatul Ulama Surabaya

<sup>1</sup>nadyaafdholi@yahoo.com, <sup>2\*</sup>terrabelatrix888@gmail.com

**Abstract:** The study of errors in Bahasa spelling is still interesting to be researched in multi-lingual countries such as Indonesia, considering that Indonesia has more than 700 languages and many are endangered. This research aims to expose spelling errors in the Bahasa of Nahdlatul Ulama Surabaya University (UNUSA) Students. This study used quantitative methods with a corpus approach and descriptive qualitative with secondary data of Bahasa course assignment documents. The results of this research are an error in spelling that is influenced by various languages both mother tongue and second language in addition to the Bahasa, namely English, Arabic, Javanese, and *Slank*.

**Keywords:** *corpus, error, Bahasa, spelling, UNUSA*

## Introduction

Bahasa as a national language is an example of successful language planning and language standardization for the benefit of national development. The lack of concern for language can result in language extinction. As an archipelago with more than 18,000 islands, Indonesia has a rich and diverse multi-cultural and linguistic heritage. More than 700 languages are spoken by around 300 different ethnic groups. The largest ethnic groups are Javanese with 45% of the total population, Sundanese 14%, Madurese 7.5%, and coastal Malays 7.5% (Czermak, Delanghe, & Wenig, 2003: 2). Specific word, phrases, saying, sign and symbol can be shown by the culture. (Kwary & Miller, 2013: 258). However, this does not matter as long as the Indonesian people hold Indonesian national motto 'Unity in Diversity', which reflects that although the people are diverse, Indonesia remains one.

Bahasa is the only official language of the Republic of Indonesia and since independence has been actively promoted as a link between many people throughout Indonesia (Hajek, 2006: 125). Nababan said by looking at a large number of Indonesian people which are more than 170 million, it can be concluded that Indonesian people speaks many languages (Nababan, 1991: 116). It shows how diverse languages in Indonesia.

There are 726 languages in Indonesia which makes it as the second of the diverse world, after Papua New Guinea which has 823 local languages (Martí et al in Riza, 2008: 93). The language spoken in Indonesia is divided into three classes: Bahasa, local or everyday language, and foreign languages (Nababan, 1985: 3). In 1980, the National Center for Language Development surveyed bilingualism in Indonesia. The results of this survey indicate that there is a significant increase in the use of Bahasa as the first language in the Provincial Capital and small towns in 7 of the 13 provinces covered by the survey (Nababan, 1985: 6). The bilingualism survey is based on the assumption that through the status and function of Bahasa as an official language, and more Indonesians in everyday life feel the need to speak Bahasa and local language. In general, people use Bahasa in more modern and public activities, while the use of local languages in more traditional and regional aspects of their lives.

The increase in the sector of formal second language teaching (from pre-school to university level) has resulted in several second language speakers with varying degrees of skill in public

use and official recognition (Comrie, 2009: 782). Therefore, in a multi-language country like Indonesia, the study of Bahasa spelling errors is very interesting considering that Indonesia has more than 700 languages and many are endangered.

In this study, the area that used as a place of research is the University of Nahdlatul Ulama in Surabaya. The language used in this academic domain is Bahasa. Participants involved were undergraduate students in nurse department by giving the task of obtaining descriptive Bahasa essays. These essays will be analyzed by the researchers.

This research applies corpus linguistics by using AntConc, it will shows language variation that is used by students from these two classes. The word classes studied in this study consist of four-word classes, namely nouns, verbs, adjectives, and adverbs.

Previous research on the corpus has been researched several times in international journals, including Journal entitled *Selecting and Creating a Word List for English Language Teaching* written by Kwary and Jurianto (Kwary & Jurianto, 2017: 60). The study analyzes the claims made by the existing word list and proposes a method for selecting words and making word lists. The results of the study indicate that there are differences in the range of word lists because differences in the corpora and source texts are analyzed. Besides, the study suggests that educators should make a list of words that are personal and comprehensive so that it is beneficial for students in learning English.

Besides, research on corpus found in the vocabulary entitled *A Corpus-Based Study on the Technical Vocabulary of Islamic Religious Studies* written by Simbuka, Hamied, Sundayana, and Kwary (Simbuka, Hamied, Sundayana, & Kwary, 2019: 47). The research mapped out the construction of vocabulary lists in Islamic religious studies developed through a corpus of Islamic religious studies in Islamic-based tertiary institutions. The vocabulary list was extracted from five sub-disciplines, namely the Qur'an, the hadith, Islamic law and jurisprudence, Islamic philosophy and theology, and theology of Islamic mysticism.

Grabowski in his research entitled *Stance Buncles in English-to-Polish Translation: a Corpus-Informed Study* uses the parallel English-Polish Paralela corpus and the National Corpus of Polish to extract and explore the use of frequencies from the Polish equivalent through selected English lexical collections that show epistemic attitudes and attitudes (Grabowski, 2018: 404). This corpus study reveals a large number of Polish equivalents, both single and multiple words, which express an attitude which shows that the majority of Polish equivalents are often used in original Polish texts and potentially help improve the compatibility of translated texts.

Based on observations from previous studies of the linguistic corpus, spelling errors in Bahasa of Nahdlatul Ulama University Surabaya students have never been studied, so the formulation in this study lies in how spelling errors in Bahasa were produced by students at the Nahdlatul Ulama University in Surabaya.

## **Backgrounds of the Study**

Corpus linguistics is used by linguists to observe language patterns in various types of texts (Novita & Kwary, 2018: 53), given that communication, today tends to be more informal, so that the corpus can also be used to see informality (Praminatih, Kwary, & Ardaniah, 2018: 101). Campoy in Kwary and Artha said that research in corpus linguistics had led to the elaboration of higher quality student input and gave researchers and teachers a broader and better perspective on the language used (Kwary & Artha, 2017: 2). Pedagogical-oriented research on specially designed lists containing high-frequency words in a collection of texts aims to benefit the

learning and teaching of second or foreign language vocabulary (Simbuka, Sundayana, Hamied, & Kwary, 2019: 277). Simbuka, Hamied, Sundayana, and Kwary said that the general definition of vocabulary is all words that exist in a particular language or subject (Simbuka, Hamied, Sundayana, & Kwary, 2019: 50). The motivation behind the corpus linguist's efforts in compiling a list of words, whether they will be general vocabulary or special discipline vocabulary is largely to weigh pedagogical lists of the benefits of learning second or foreign language vocabulary (Simbuka, Sundayana, Hamied, & Kwary, 2019: 281).

Corpus is a relatively new approach to applied linguistics that emphasizes its application in teaching in the second language which includes various specific techniques with certain characteristics (Spolsky & Hult, 2008: 539). A corpus is needed to document and explore languages (Kwary, 2019: 102). Language descriptions are developed from systematic observations of language behavior, which emphasizes empirical analysis, to see what is typical and what is unusual (Spolsky & Hult, 2008: 540). All corpus linguistic works use a corpus that looks at a large collection of texts that have special features in digital form (Spolsky & Hult, 2008: 540). This refers to the fact that these texts are produced by language users in real communication situations. That is, they are not based on intuition or anecdotal evidence about how language might be used. The language can be completely spontaneous, especially in the case of language learners, part of a particular task, which in this study uses text in the form of student essays as class assignments. Thus it will be represented what is the variation of language in the analysis of a very specific text.

Corpus linguistics has begun to make valuable contributions to the teaching and learning of second and foreign languages (Spolsky & Hult, 2008: 543). Inter-language studies of students always rely on empirical analysis and thus share characteristics with corpus linguistic works. However, the corpus approach to analysis between languages has several different strengths and limitations. Its strengths include the ability of studies to enter more data and more participants and to examine the interaction of more variables than is generally feasible with previous approaches. However, at the same time, corpus-based work is limited in the number of contexts provided for analysis (Spolsky & Hult, 2008: 543). By examining patterns in language use, especially on a broader scale than before, corpus linguistics has begun to make valuable contributions to the field of teaching and learning of second and foreign languages.

## **Research Method**

This research uses a quantitative method with a corpus and descriptive qualitative approach. The corpus relies on computer-assisted techniques or software. The most commonly used software is concordance which will display all occurrences of words in the surrounding context (Spolsky & Hult, 2008: 541). A concordance is produced by appropriate software (Kwary, Howard Jackson, 2009: 3). According to Kwary, corpus users will be able to compare the results obtained from this corpus and concordancer with data found in studies or other textbooks. (Kwary, 2018: 95). In this study, researchers were assisted by the application of AntConc to see concordance, calculate word frequencies, analyze collates, and also see statistical measures of the strength of word associations.

Recognizing patterns in how language is used certainly requires making quantitative judgments; just saying something is "typical" or "general" means that it happens more often than other choices. However, the count alone explains a little about the use of language. Even the most complex quantitative analyzes must be linked to functional interpretations of language patterns (Spolsky & Hult, 2008: 543).

The data from the study came from descriptive essays in Bahasa with the theme of self-data written by 120 undergraduate Nursing undergraduate students from the Nahdlatul Ulama University in Surabaya, 700-1000 words long. Researchers convert student essays from Microsoft Word in doc format to txt format. Furthermore, the researchers used AntConc to analyze the data by uploading the student essay into AntConc and using the wordlist column to identify spelling errors that corresponded to the word classes of Alwi, et al (Alwi, Soenjono, Lapoliwa, & Anton, 2010) such as nouns, verbs, adjectives, and adverbs which were then interpreted as results.

**Results and Discussion**

In Bahasa, the vocabulary used by the students of the University of Nahdlatul Ulama Surabaya is diverse. This was written in a self-description essay as Bahasa assignment. Based on the results of the descriptive essay writing assignments in Bahasa, it was found that the students used Bahasa in the formal context, namely in the Bahasa lecture, but there were still many spelling errors in the students of the Nahdlatul Ulama University in Surabaya, namely spelling errors found in the noun word class, verbs, adverbials, adjectives, as described below.

*Spelling Errors in Nomina Word Classes*

Spelling mistakes in Bahasa in the noun word class of Nahdlatul Ulama University students can be found in the writing error of the person's name, university name, month name, city name, island name, and country name. This can be seen in the following table;

**Table 1.1** Spelling Errors in the Noun Word Class

Name of People	Name of Universities	Name of Months	Name of Cities/Regionsf	Name of Countries	Name of Islands
cindy	brawijaya	juni	banten	malaysia	madura
hasyim	airlangga	oktober	tanggulangin	indonesia	kalimantan
lestari	untag	nopember	jakarta	singapore	bawean
khadijah	nadhlatul		sidoarjo	korea	sulawesi
intan	ulama		jombang	australia	ternate
iskandar	unusa		tanjung sari		
nabilla	hangtuah		banyuwangi		
achmad			probolinggo		
lidy			mojokerto		
nabila			ponorogo		
novia			balikpapan		
puspita			banjarmasin		
regita			banyumas		
sabila			purwosari		
			nganjuk		

The error spelling the name in the Bahasa essay assignment can be seen in the software namely AntConc. In AntConc, students write their middle and last name prefixes in the majority using lowercase letters. Then, for writing the name of the University, students write the prefix of the university name using lowercase letters and the abbreviation of the name of the university uses the whole lowercase letter. For example, Universitas Brawijaya is written as a University

of Brawijaya, UNTAG is written as Untag. Furthermore, in writing the name of the month, the name of the city/region, country and island also be written using a lowercase prefix.

Researchers also found several spelling errors in the noun word classes that were influenced by foreign languages such as English, Arabic, Javanese and *slank* languages. This can be seen in table 1.2;

**Table 1.2** Effect of Foreign, Local, and *Slank* Languages

English	Arabic	Javanese	<i>Slank</i>
klue	kyai	angkot	cowok
badminton	fiqih	tebel	udah
hobby	hadits	laok	ga papa
standart	ustadz	mbak	
thypus		hamper	
maroon		sampek	
stay		nopember	
exkstrakurikuler		negri	
email		malem	
fashion		pedenya	
favorite		pendiem	
food		guyon	
emoticon			
tehnik			
channel			
culture			
entrepreneurship			
focus			
guide			
important			
incubator			
hangout			

After knowing the writing errors in Bahasa writing, the writer also found that ‘saya’ word which is a standard language in Bahasa appeared 4732 times, while the word ‘aku’, which is a non-standard language in Bahasa appeared 631 times. The word ‘hobby’ appears 9 times while the word ‘hobi’ which is standard Bahasa appears 38 times. Furthermore, for ‘nopember’ and ‘November’ the words appear 9 times. The authors found non-standard writing for the word ‘apotek’ written to be an ‘apotik’.

### *Spelling Errors in Verbs Word Classes*

Spelling Errors in Bahasa verb word classes for Nahdlatul Ulama University students can be seen in table 1.3;

**Table 1.3** Effects of Foreign, Local and *Slank* Languages on Bahasa

English	Arabic	Javanese	<i>Slank</i>
browsing	fikir	ngaji	alamin
searching	meridhoi	nglarang	bertemen
diintruksikan	wudhu	ikutin	dibolehin
bullying	difahami	nangisan	inget
block	berfikir	motoran	diinget
labi			keinget
mensupport			
jugje			
make			

After knowing the data from table 1.3 above, we can know that foreign languages, Javanese and *Slank* languages can influence the writing of the spelling used by students in writing a written work.

*Spelling Errors in Adverbial Word Classes*

Spelling errors in the variety of words in the adverbial Bahasa word-class includes; incorrect writing of affixes to the word 'membangga'. The word 'membangga' should have standard writing like 'membanggakan'. Then, the spelling error in the 'memberikan' is written as 'memebrerikan'. Furthermore, the error writing of the 'akhirnya' written by students becomes 'ahirnya, akhinya, and akirnya'. Wrong spelling of Bahasa in the adverb word-class is influenced by students' lack of understanding of affixes.

*Spelling Errors in Adjective Word Classes*

Spelling errors in the adjective word class are seen in writing the word 'gemar' which are written as 'gemari'. Students should write to be 'menggemari'. Furthermore, in writing adjectives are also influenced by *Slank* language, this can be seen in writing the word 'kemalasan' which is written as 'kegabutan'. In addition to *Slank* language, it turns out students are also influenced by Javanese in writing spelling in adjective word classes. This can be seen in the writing of the word 'jengkel' which is written as 'mangkel'. The data shows that students still lack understanding about writing standard Bahasa in adjective word classes.

*Spelling Errors in Preposition*

Mistakes in spelling are not only found in the four classes of nouns, verbs, adverbials, and adjectives. But it is also found in prepositional writing. This can be seen in table 1.4;

**Table 1.4** Error Writing Spelling in Prepositions

Affix <i>di</i> -	Affix <i>ke</i> -
di acc	ke egoisanku
di antar	ke pergian nya
di belikan	ke Pendidikan
di cepatkan	ke sekian kali nya
di mana	ke 4
di anugerahi	ke 3
di bangunin	
di bawah	

---

di bentak  
 di butuhkan  
 di buru  
 di buka  
 di cepatkan  
 di cap  
 di dampingi  
 di didik  
 di dudukkan  
 di extra  
 di fikiranku  
 di gendong  
 di gunakan  
 di haruskan  
 di ikutkan  
 di ingat  
 di jadikan  
 di jaga  
 di jelaskan  
 di lahirkan  
 di lalui

---

Based on table 1.4 we can see that students separate prepositions *di-* and *ke-* to all words that have an influence and those that do not affect. This shows that students lack understanding of affixes and prepositions.

### *Social Reasons of Spelling Error*

After knowing that some of students who are studying in the Nahdlatul Ulama University in Surabaya are not only coming from Java Island but also they come from Kalimantan, Sumatra, Ternate and Madura Island. It shows that education in Java Island is more advanced than another Island in Indonesia. Thus, students who come from Java Island are going to have more understanding to write the standard of Bahasa. Meanwhile this study shows that those students who come from Kalimantan, Sumatra, Ternate, Madura and Java Islands in Nahdlatul Ulama University have the same problem in lack of understanding to write the standard of Bahasa. Due to Nahdlatul Ulama University in Surabaya is an Islamic University it causes Arabic and Javanese languages influence students to have spelling error when they write an essay in Bahasa. The same things happened with *Slank* language, Surabaya is the metropolis city in East Java and the age range of students in this study are between 18-21 thus these cause they use some *slank* language when they communication and influence when they write an essay in Bahasa.

### **Conclusion**

This study proves that spelling errors in Bahasa are influenced by English, Arabic, Javanese language, *Slank*. Besides, there is a lack of understanding of the standard form of writing in Bahasa. It would be better to hold standard and non-standard form writing exercises as well as further understanding of the affixes for students, to preserve the standard Bahasa form at the

tertiary level. As researchers, our wish is to give the understanding about writing properly thus it can maintain an ideal writing for every individual.

## References

- Alwi, H., Soenjono, D., Lapoliwa, H., & Anton, M. M. (2010). *Tata Bahasa Baku Bahasa Indonesia 3rd Edition*. Jakarta: Pusat Bahasa dan Balai Pustaka.
- Comrie, B. (2009). *The World's Major Languages Second Edition*. USA: Routledge.
- Conrad, S. M., & Levelle, K. R. (2008). Corpus Linguistics and Second Language Instruction. In B. Spolsky, & F. M. Hult, *The Handbook of Educational Linguistics* (pp. 539-556). USA: Blackwell Publishing Ltd.
- Czermak, K., Delanghe, P., & Wenig, W. (2003). Preserving Intangible Cultural Heritage in Indonesia: A Pilot Project on Oral Tradition and Language Preservation. *UNESCO Jakarta Office - Indonesia*, 1-8.
- Grabowski, Ł. (2018). Stance Buncles in English-to-Polish Translation: a Corpus-Informed Study. *ТЕОРИЯ И ПРАКТИКА ПЕРЕВОДА THEORY AND PRACTICE OF TRANSLATION: Russian Journal of Linguistics.* , 404-422.
- Hajek. (2006). On the Edge of the Pacific: Indonesia and East Timor. In D. Cunningham, D. E. Ingram, & K. Sumbuk, *Language Diversity in the Pacific: Endangerment and Survival (Multilingual Matters)* (pp. 121-132). Canada: MPG Books Ltd.
- Kwary, D. A. (2009). Howard Jackson, Key Terms in Linguistics. *Continuum*, 1-6.
- Kwary, D. A., & Miller, J. (2013). A Model for an Online Australian English Dictionary Database. *Terminology. International Journal of Theoreticak and Applied Issues in Specialized*, 258-276.
- Kwary, D. A., & Artha, A. F. (2017). The Academic Article Word List for Social Sciences. *Mextesol Journal*, 1-11.
- Kwary, D. A., & Jurianto. (2017). Selecting and Creating a Word List for English Language Teaching. *Teaching English with Technology*, 60-72.
- Kwary, D. A. (2018). A Corpus and a Concordancer of Academic Journal Articles. *Data in Brief*, 94-100.
- Kwary, D. A. (2019). A Corpus Platform of Indonesian Academic Language. *SoftwareX*, 102-106.
- Nababan, P. W. (1985). Bilingualism in Indonesia: Ethnic Language Maintenance and the Spread of the National Language. *Southeast Asian Journal of Social Science*, 1-18.
- Nababan, P. W. (1991). Language in Education: The Case of Indonesia. *International Review of Education*, 115-131.
- Novita, H., & Kwary, D. A. (2018). Comparing the Use of Lexical Bundles in Indonesian-English Translation by Student Translators and professional Translators. *Humanities & Social Sciences Collection*, 53-74.
- Praminatih, G. A., Kwary, D. A., & Ardaniah, V. (2018). Is EFL Students' Academic Writing Becoming More Informal? *Journal of World Languages*, 101-112.
- Riza, H. (2008). Resource Report on Language of Indonesia. *Proceedings of the 6th Workshop on Asian Language Resources (IJCNLP)*, (pp. 93-94).
- Simbuka, S., Hamied, F. A., Sundayana, W., & Kwary, D. A. (2019). A Corpus-Based Study on the Technical Vocabulary of Islamic Religious Studies. *Teflin Journal*, 47-71.

Simbuka, S., Sundayana, W., Hamied, F. A., & Kwary, D. A. (2019). Membandingkan Pendekatan Klasifikasi Kosakata dan Analisis Kata Kunci untuk Menentukan Kosakata Teknis Studi Islam. *Eleventh Conference on Applied Linguistics (CONAPLIN 2018)* (pp. 277-282). Atlantis Press.

Spolsky, B., & Hult, F. M. (2008). *The Handbook of Educational Linguistics*. US: Blackwell Publishing Ltd.