WHEN GOOGLE SEARCH ENGINE TALKS TO: CORPUS-BASED TRANSLATION

Mochamad Nuruz Zaman1), Luthfi Muhyiddin2) Universitas Darussalam Gontor

nuruzzaman@unida.gontor.ac.id, luthfimuhyiddin@unida.gontor.ac.id

Abstract

Language development within linguistics rapidly grows and affects the most, particularly in translation studies. The influence has mapped the interest of a lot researcher in Indonesia, due to research method development based on corpus linguistics facilitates the expansion of translation studies scope and address new perspectives to translation raters. Corpus-based translation of this study comes to employ lexicon from statistical tools. The data collection is health protocolary text from several domestic and overseas media where record a large and structured set of texts. The applied method is qualitative research integrated with a corpus-based as descriptive approach and the analysis tool is Google Search Engine patterned by content analysis. Therefore, this study examines translation accuracy and acceptability of English terms as source language toward their Indonesian terms as target language that available online media and the quantity appearance about the COVID-19. This elaboration assessment well creates quality among them to rate how extent the lexicon reaches numerical form.

Keywords – corpus-based translation, COVID-19 lexicon, Google Search Engine, quality of accuracy and acceptability

Introduction

Language structures and cultural patterns in translation involve aspects of complexity that are crucial for equivalence. This is Catford's opinion, "The central problem of translation practice is that of finding TL translation equivalents" (Zaman, 2018: 16). Catford's opinion refers to the current reality, that the main problem in translation is to find the equivalent as accurately as possible according to the message contained in the source language text and the resulting equivalent must meet the rules of the target language.

The most important thing in finding an equivalent is to produce the same or almost the same reaction to the reader of translated text as the reader reacts to the original text. This obstacle must be anticipated if the concept of translation is related to special terms in certain fields, such as the translation of terms related to the COVID-19 Pandemic.

Given these obstacles, an assessment is needed that emphasizes the components of the source language and target language so that it can be seen how the appropriate transfer of messages is assessed. Besides that, because the equivalent of words in translation is very closely related to the instrument, it must be adapted to the analysis tool of translation results that are commensurate and natural. The method of analysis in this research is corpus-based translation.

According to Setiawan (2016: 6), the corpus is a collection of written texts that have either been completed or are still in the writing process produced by certain authors. In its development, the study of a

corpus, which was later known as corpus linguistics, was more interested in studying a language. This is because corpus linguistics is considered to make it easier to research the variety and use of language empirically (Biber and Reppen, 2015: 1). In line with the development of translation practices that are supported by technology, the development of an electronic corpus supporting corpus studies is now making it easier for corpus-based research, including translation studies, in especially in Indonesia. Several types of corpora become the basis of translation research, including parallel corpus, comparable corpus, and multilingual corpus.

For this reason, this research will use a multilingual corpus as a corpus-based translation research tool. The data analysis technique is using Google Search-Engine. This research will also use English language corpus databases such as the British National Corpus (BNC) and Bank of English (BoE). The focus of the study is to assess the accuracy and acceptability of the translation of terms related to the COVID-19 pandemic from English to Indonesian.

The use of technology for education that is applied in the form of software is patterned for the elaboration of a social case that is currently hot, namely the COVID-19 pandemic (as research data) to formulate the information in applied linguistic concepts. This concept is part of the social aspect humanities of the that simultaneously integrates with educational and learning technology that focuses on assessing the quality of the accuracy and acceptability of the translation of terms related to the COVID-19 pandemic in terms of the corpus translation perspective.

The specific target based on the research description is to assess the quality of translation accuracy and acceptability of translation of the lexicon of terms related to the COVID-19 pandemic. While the urgency of this study theoretically has an

impact on scientific contributions in the form of innovations in understanding lexical meanings about the accuracy and acceptability of translations as well as participating in providing expansion of data analysis with the method of corpus analysis tools in the form of Google Search-Engine, while practically this research deserves to be used as new knowledge because it is appropriate with current issues. So that the novelty of reading literacy will be guided according actual and up-to-date to knowledge.

Sujaini, H. (2018) explains that the experimental strategy for improving the quality of the corpus. The test method Bilingual Evaluation Understudy (BLEU) results in an Indonesian-Malay translation 6.97% and Indonesianaccuracy of Javanese 5.55%. This is where it comes from, the assessment of accuracy does not appear as an indicator of a good strategy. So, researchers have a gap to develop it through an assessment of the quality of accuracy assessment. Indrayana, D, et al. (2016) explain that if the data test used Bilingual Evaluation Understudy (BLEU) and linguists by adding the PoS feature there was an increase in the BLEU value of 0.6% in automated testing and 21.67% in testing by linguists. From this it emerged, the data test did not involve linguistic patterns at the lexicon level, so that acceptability still needed to be tested for the next phase. So, researchers have the opportunity to study further at the lexicon level through the acceptability aspect. Abd. Rahman, (2020), explains that the term COVID-19 pandemic is accepted in the context of translation. In this study, the level of acceptance of the translation was not determined and did not discuss accuracy, and was not a corpus-based research.

This update will be sharpened in the discussion flow through a problem formulation that will develop in the form of research questions, namely (1) What is the

quality of accuracy of the translation of terms related to the COVID-19 pandemic demonstrated by Google Search Engine; and (2) How is the quality of acceptability of the translation of terms related to the COVID-19 pandemic demonstrated by Google Search Engine?

Methodology

This study applies a qualitative method with a descriptive approach, namely problemsolving procedures by describing the state of the research object based on the facts that appear as they are (Raco, 2010: 80). With the exposure of social phenomena that are currently hot related to COVID-19. Researchers will investigate in the context of the field of applied linguistics the translation of the product of this phenomenon.

Data sources are sources of data acquisition that are qualitatively capable of being a reference and foothold in the distribution of information and evidence of phenomena at the focus of research (Santosa, 2017: 51). The data sources for this study consist of documents and informants. Documents in the form of words, phrases, or clauses in the Google Search-Engine. Informants who were selected for their contributions to aspects of linguistic context and aspects of translation assessment were Linguistics Scholar and Translation Expert.

Research data is divided into two parts, namely primary and secondary data. Primary data refers to data collected by researchers from the research location directly. Meanwhile, secondary data refers to data collected by other researchers to support their studies (Santosa 2017: 52). The data in this study are focused and fixed on the terms of the COVID-19 field with characteristic patterns of foreign language acquisition, synonyms, abbreviations, acronyms, and common

The data collection technique that examines documents and archives is content analysis, because it explores the content of written or printed information in the mass media (Ardi, 2010: 59). The data collection of this study was carried out by looking for combinations of words and entering the desired words among the searched words. A lexicon search can be performed by entering a specific symbol next to an inflected shape in the search box to enter each inflected form in a concordance in the Google Search-Engine.

The data analysis of this study is by the review of Siyoto and Sodik (2015: 109), namely a series of activities of reviewing, grouping, systematizing, interpreting, and verifying data so that a phenomenon has social, academic, and scientific value. Activities in research data analysis are: grouping data based on the frequency or alphabetical lexicon of COVID-19 terms on Google Search-Engine, tabulating data, presenting data, calculating translation quality. and performing assessment calculations for lexical by comparative validity tests, and drawing conclusions.

The research framework begins with the classification of the English term COVID-19 based on the lexicon category. Next, the equivalent of these terms will be found in Indonesian through Google Search Engine. Furthermore, the researcher will pattern the term with the characteristic pattern of foreign language acquisition, synonyms, abbreviations, and acronyms. Appearance in overriding search engine is displayed for accurate with Google overriding media detection and acceptability with Google overriding number detection. In the final stage, an assessment of the quality of accuracy and acceptability is carried out.

1. Quality of accuracy of the translation of terms related to the COVID-19

pandemic demonstrated by Google Search Engine

Table 1. Data of accuracy of the translation of terms related to the COVID-19 pandemic demonstrated by Google Search Engine.

N	English	Its	Lexicon	Translation	Media	Quality of Accuracy		
0	Term	Indonesian Equivalent	Category	Technique	Detectio - n	(3)	(2)	(1)
		-	D 11 1			(3)		(1)
1	New	Adaptasi	English	Addition &	www.dis		\checkmark	
	normal	kebiasaan	lexicon	established	kes.balip			
		baru	F 1' 1	equivalent	rov.go.id			
		Kenormalan	English	Pure	www.jou			
		baru	lexicon	borrowing &	rnal.ugr.			
				established	ac.id			
	0 1	T · 1	F 1' 1	equivalent	1			
2	Social	Jaga jarak	English	Established	www.kra			
	distancing		lexicon	equivalent	kataume			
		D 1	D 1: 1	A 1	dia.com		1	
		Pembatasan	English	Adaptation	ww.alod			
2		sosial	lexicon	T . 1 1 1	oc.com	1		
3	Physical	Jaga jarak	English	Established	www.dis			
	distancing	fisik	lexicon	equivalent	hub.kuka			
					rkab.go.i			
		D 1 /	F 1' 1	A 1 4 4	<u>d</u>			
		Pembatasan	English	Adaptation	www.alo			
4	T 1	fisik	lexicon	NT / 1' 1	doc.com			
4	Isolation	Isolasi	Synonym	Naturalized	www.hal	\checkmark		
			lexicon	borrowing	odoc.co			
		IZ	C	A .1	m		ν	
		Karantina	Synonym	Adaptation	www.jeo		N	
			lexicon		.kompas.			
5	Disinfecta	Disinfektan	Cumonum	Naturalized	com			
3		Disimertan	Synonym lexicon	borrowing	<u>www.tok</u>	N		
	nt		IEXICOII	Donowing	<u>opedia.c</u> om			
		Pembunuh	Synonym	Description	www.tok			
		bakteri	lexicon	Description	opedia.c			v
		Dartell	юлеон		om			
6	WFH	Bekerja dari	Abbreviation	Established	www.gli			
0	W111	rumah	lexicon	equivalent	nts.com	,		
		Bekerja	Abbreviation	Adaptation	www.id.			
		jarak jauh	lexicon	Muplation	wikipedi		,	
		Juluk Juuli	Технеон		a.org			
7	SFH	Belajar dari	Abbreviation	Established	www.jou			
,	5111	rumah	lexicon	equivalent	rnal.unes	•		
			10	equivalent	a.ac.id			
		Belajar jarak	Abbreviation	Adaptation	www.jou			
		jauh	lexicon		rnal.unes			
		5			a.ac.id			
8	Webinar	Seminar	Acronym	Pure	www.ev			
-		online	lexicon	borrowing &	etx.io			
				adaptation				
		Seminar	Acronym	Pure	www.ev			
		dalam web	lexicon	borrowing &	etx.io		,	
				adaptation	,			
0) (1	A	Generalization	www.tok		V	
9	Masker	Masker	ACIOIIVIII	UCHCIAIIZAHUH			V	
9	Masker N-95	Masker Respirator	Acronym lexicon	Generalization	opedia.c		v	

hidung dan	Acronym Description lexicon	<u>www.tok</u> opedia.c om	V
mulut			

Data 1 demonstrates "new normal", equivalent to "adaptasi kebiasaan baru" is classified by English lexicon. It produces adaptation and established equivalent translation technique effect to the less accurate assessment by appearing overriding media detection by Google Search Engine within www.diskes.baliprov.go.id.

Simultaneously, "new normal", equivalent to "kenormalan baru" is classified by English lexicon. It produces pure and established equivalent borrowing translation technique effect to the less accurate assessment bv appearing overriding media detection by Google Search within Engine www.journal.ugr.ac.id.

Data 2 illustrates "social distancing", equivalent to "jaga jarak" is classified by English lexicon. It produces established equivalent translation technique effect to the accurate assessment by appearing overriding media detection by Google Search Engine within www.krakataumedia.com. Simultaneously, "social distancing", equivalent to "pembatasan sosial" is classified bv English lexicon. It produces adaptation translation technique effect to the accurate assessment by appearing overriding media detection by Google Search Engine within ww.alodoc.com.

Data 3 displays "physical distancing", equivalent to "jaga jarak fisik" is classified by English lexicon. It produces established equivalent translation technique effect to the accurate assessment by appearing overriding media detection by Google Search Engine within www.dishub.kukarkab.go.id.

Simultaneously, "physical distancing", equivalent to "pembatasan fisik" is

classified by English lexicon. It produces adaptation translation technique effect to the less accurate assessment by appearing overriding media detection by Google Search Engine within www.alodoc.com.

Data 4 exposes "isolation", equivalent to "isolasi" is classified by synonym lexicon. naturalized produces borrowing translation technique effect to the accurate assessment by appearing overriding media detection by Google Search Engine within www.halodoc.com. Simultaneously, "isolation", equivalent to "karantina" is classified by synonym lexicon. It produces adaptation translation technique effect to the less accurate assessment by appearing overriding media detection by Google Search Engine within www.jeo.kompas.com.

Data 5 organizes "disinfectant", equivalent to "disinfektan" is classified by synonym lexicon. It produces naturalized borrowing translation technique effect to the accurate assessment by appearing overriding media detection by Google Search Engine within www.tokopedia.com. Simultaneously, "disinfectant", equivalent to "pembunuh bakteri" is classified by synonym lexicon. It produces adaptation translation technique effect to the less accurate assessment by appearing overriding media detection by Search Engine Google within www.tokopedia.com.

Data 6 formulates "WFH", equivalent to "bekerja dari rumah" is classified by abbreviation lexicon. It produces established equivalent translation technique effect to the accurate assessment by appearing overriding media detection by Google Search Engine within www.glints.com. Simultaneously, "WFH", equivalent to "bekerja jarak jauh" is classified by abbreviation lexicon. It produces adaptation translation technique effect to the less accurate assessment by appearing overriding media detection by Google Search Engine within www.id.wikipedia.org.

Data 7 formulates "SFH", equivalent to "belajar dari rumah" is classified by abbreviation lexicon. It produces established equivalent translation technique effect to the accurate assessment by appearing overriding media detection by Google Search Engine within www.journal.unesa.ac.id. Simultaneously, "SFH", equivalent to "belajar jarak jauh" is classified by abbreviation lexicon. It produces adaptation translation technique effect to the less accurate assessment by appearing overriding media detection by Google Search Engine within www.journal.unesa.ac.id.

Data 8 expresses "webinar", equivalent to "seminar online" is classified by acronym lexicon. It produces pure borrowing and adaptation translation technique effect to the less accurate assessment by appearing overriding media detection by Google Search Engine within www.evetx.io. Simultaneously, "webinar", equivalent to "seminar dalam web" is classified by acronym lexicon. It produces pure borrowing and adaptation translation technique effect to the less accurate assessment by appearing overriding media detection by Google Search Engine within www.evetx.io.

Data 9 designs "Masker N-95", equivalent "masker respirator" classified by to lexicon. It produces acronym generalization translation technique effect to the less accurate assessment by appearing overriding media detection by Google Search Engine within www.tokopedia.com. Simultaneously, "Masker N-95", equivalent to "masker kain penutup hidung dan mulut" is classified by acronym lexicon. It produces description effect to translation technique the inaccurate assessment bv appearing overriding media detection by Google Search Engine within www.tokopedia.com.

For evaluating how the accuracy data is. Here is the tabulation of accuracy assessment for ensuring the data is structurally monitored:

Accuracy Level	Score	Qualitive Parameter	Tabel 2. Scale and Information for
Accurate	3	Source language message idea is advanced transferred	Accuracy Instrument (Adapted from Nababan. M.R., et.al: 2012)
		into targeted language; undetected distortion of meaning	Data of accuracy is indicated in to three levels. Accurate level consists of eight data,
Less Accurate	2	Most of source language message idea is intermediate transferred into targeted language; found intervention of idea affected to whole meaning.	less accurate level consists of three data, and inaccurate level consists of seven data. Here is the means calculation:
Inaccurate	1	Target language message idea is meaningless or powerless within source language message idea	

(accurate data number x accurate score) + (less accurate data number x less accurate score) + (inaccurate data number x inaccurate score) accuracy data number

> = means $\frac{(8 \times 3) + (3 \times 2) + (7 \times 1)}{18}$ = <u>2,05</u>

Means calculation of accuracy data is 2.05 by indicating "less accurate"

2. Quality of acceptability of the translation of terms related to the COVID-19 pandemic demonstrated by Google Search Engine

Table 3. Data of acceptability of the translation of terms related to the COVID-19 pandemic demonstrated by Google Search Engine

N	English	Its	Lexicon	Translation	Google	Quality of Acceptability		
0	Term	Indonesian Equivalent	Category	Technique	Number Detecti on	(3)	(2)	(1)
1	New normal	Adaptasi kebiasaan baru	English lexicon	Addition & established equivalent	661,000			
		Kenormal	English	Pure	280,000		\checkmark	

		an baru	lexicon	borrowing & established equivalent				
2	Social distancing	Jaga jarak	English lexicon	Established equivalent	411,000			
	unstantoning	Pembatasa n sosial	English lexicon	Adaptation	308,000			
3	Physical distancing	Jaga jarak fisik	English lexicon	Established equivalent	119,000			
	C	Pembatasa n fisik	English lexicon	Adaptation	102,000			
4	Isolation	Isolasi	Synonym lexicon	Naturalized borrowing	543,000			
		Karantina	Synonym lexicon	Adaptation	7,650,0 00			
5	Disinfecta nt	Disinfekta n	Synonym lexicon	Naturalized borrowing	751,000			
		Pembunuh bakteri	Synonym lexicon	Description	888,000			
6	WFH	Bekerja dari rumah	Abbreviati on lexicon	Established equivalent	1,670,0 00			
		Bekerja jarak jauh	Abbreviati on lexicon	Adaptation	126,000			
7	SFH	Belajar dari rumah	Abbreviati on lexicon	Established equivalent	41,700			
		Belajar jarak jauh	Abbreviati on lexicon	Adaptation	29,800			
8	Webinar	Seminar online	Acronym lexicon	Pure borrowing & adaptation	66,700, 000		V	
		Seminar dalam web	Acronym lexicon	Pure borrowing & adaptation	1,140,0 00			
9	Masker N-95	Masker Respirator	Acronym lexicon	Generalizati on	9,910,0 00			
		Masker kain penutup hidung dan mulut	Acronym lexicon	Description	76,800	V		

Data 1 demonstrates "new normal", equivalent to "adaptasi kebiasaan baru" is classified by English lexicon. It produces adaptation and established equivalent translation technique effect to the

acceptable assessment by appearing overriding number detection by Google Search Engine on 661,000. Simultaneously, "new normal", equivalent to "kenormalan baru" is classified by English lexicon. It produces pure borrowing and established equivalent translation technique effect to the less acceptable assessment by appearing overriding number detection by Google Search Engine on 280,000.

Data 2 illustrates "social distancing", equivalent to "jaga jarak" is classified by English lexicon. It produces established equivalent translation technique effect to the acceptable assessment by appearing overriding number detection by Google 411,000. Engine within on Search Simultaneously, "social distancing", equivalent to "pembatasan sosial" is classified by English lexicon. It produces adaptation translation technique effect to the unacceptable assessment by appearing overriding number detection by Google Search Engine on 308,000.

Data 3 displays "physical distancing", equivalent to "jaga jarak fisik" is classified by English lexicon. It produces established equivalent translation technique effect to the acceptable assessment by appearing overriding number detection by Google Search Engine on 119,000. Simultaneously, distancing", equivalent "physical to "pembatasan fisik" is classified by English lexicon. It produces adaptation translation technique effect to the less acceptable assessment by appearing overriding number detection by Google Search Engine on 102.000.

Data 4 exposes "isolation", equivalent to "isolasi" is classified by synonym lexicon. produces naturalized borrowing It translation technique effect to the unacceptable assessment by appearing overriding number detection by Google Search Engine on 543,000. Simultaneously, "isolation", equivalent to "karantina" is classified by synonym lexicon. It produces adaptation translation technique effect to the acceptable assessment by appearing overriding number detection by Google Search Engine on 7,650,000.

Data 5 organizes "disinfectant", equivalent to "disinfektan" is classified by synonym lexicon. It produces naturalized borrowing technique effect to translation the unacceptable assessment by appearing overriding number detection by Google Search Engine on 751,000. Simultaneously, "disinfectant", equivalent to "pembunuh bakteri" is classified by synonym lexicon. It produces adaptation translation technique effect to the acceptable assessment by appearing overriding number detection by Google Search Engine on 888,000.

Data 6 formulates "WFH", equivalent to "bekerja dari rumah" is classified by abbreviation lexicon. It produces established equivalent translation technique effect to the acceptable assessment by appearing overriding number detection by Google Search Engine on 1,670,000. Simultaneously, "WFH", equivalent to "bekerja jarak jauh" is classified by abbreviation lexicon. produces It adaptation translation technique effect to the unacceptable assessment by appearing overriding number detection by Google Search Engine on 126,000.

Data 7 formulates "SFH", equivalent to "belajar dari rumah" is classified by abbreviation lexicon. It produces established equivalent translation technique effect to the acceptable assessment by appearing overriding number detection by Engine Google Search on 41,700. "SFH", Simultaneously, equivalent to "belajar jarak jauh" is classified by abbreviation lexicon. It produces adaptation translation technique effect to the unacceptable assessment by appearing overriding number detection by Google Search Engine on 29,800.

Data 8 expresses "webinar", equivalent to "seminar online" is classified by acronym lexicon. It produces pure borrowing and adaptation translation technique effect to the less acceptable assessment by appearing overriding number detection by Google Search Engine on 66,700,000. Simultaneously, "webinar", equivalent to "seminar dalam web" is classified by acronym lexicon. It produces pure borrowing and adaptation translation technique effect to the less acceptable assessment by appearing overriding number detection by Google Search Engine on 1,140,000.

Data 9 designs "Masker N-95", equivalent to "masker respirator" classified by lexicon. acronym It produces generalization translation technique effect to the unacceptable assessment by appearing overriding number detection by Google Search Engine on 9,910,000. Simultaneously, "Masker N-95", equivalent to "masker kain penutup hidung dan mulut" is classified by acronym lexicon. It produces description translation effect to the acceptable technique assessment appearing overriding by number detection by Google Search Engine on 76,800.

For evaluating how the acceptability data is. Here is the tabulation of acceptability assessment for ensuring the data is structurally monitored Tabel 2. Scale and Information for Acceptability Instrument (Adapted from Nababan. M.R., et.al: 2012

Acceptability Level	Score	Qualitive Parameter			
Acceptable	3	Targetlanguagemessageideaproper:thetermsarenaturaltothereaderqualifiedwithIndonesianprinciple			
Less Acceptable	2	Mostly, target language message idea is proper, few of them is indicated error grammatical.			
Unacceptable	1	Targetlanguagemessageideadoesn'tsoundproper;technicaltermsrigidto the reader andunqualifiedwithIndonesianprinciple			

(acceptable data number x acceptable score) + (less acceptable data number x less acceptable score) + (unacceptable data number x acceptable score) acceptable data number

= means

$$\frac{(12 \times 3) + (4 \times 2) + (2 \times 1)}{18}$$
$$= 2.56$$

Means calculation of acceptability data is 2.56 by indicating "less acceptable"

Conclusion

Quality of accuracy of the translation of terms related to the COVID-19 pandemic demonstrated by Google Search Engine has an outcome less accurate. Meanwhile, quality of acceptability of the translation of terms related to the COVID-19 pandemic demonstrated by Google Search Engine has an outcome less accurate has an outcome less acceptable. Their quality is influenced by translation techniques, media detection, and number detection.

Acknowledgement:

The researchers highly and gratefully acknowledge the funding support for this research is from Institute of Research and Community Services (LPPM), University of Darussalam Gontor.

References:

- Ardi, Havid. (2010). Analisis Teknik Penerjemahan dan Kualitas Terjemahan Buku "Asal-Usul Elite Minangkabau Modern: Respons terhadap Kolonial Belanda Abad Ke XIX/XX". Tesis yang Dipublikasikan. Surakarta: UNS.
- Biber, D dan Reppen R. (2015). The Cambridge Handbook of English Corpus Linguistics. Cambridge: Cambridge University Press.
- Indrayana, D, dkk. (2016). Meningkatkan Akurasi pada Mesin Penerjemah Bahasa Indonesia ke Bahasa Melayu Pontianak dengan Part of Speech. Jurnal Edukasi dan Penelitian Informatika (JUSTIN) Vol. 3, No. 1, 2016.

- Nababan, M.R., Nuraeni, A., & Sumardiono. (2012). Model Penilaian Kualitas Terjemahan. Jurnal Kajian Linguistik dan Sastra, vol. 24, No. 1, 39-57.
- Raco, R.J. (2010). Metode Penelitian Kualitatif: Jenis, Karakteristik dan Keunggulannya. Jakarta: PT. Gramedia Widiasarana Indonesia.
- Rahman, Abd. (2020). Keberterimaan Istilah-Istilah di Masa Pandemi COVID-19. Jurnal Bidar, Volume 10, Nomor 2, Desember 2020 (68—82).
- Santosa, Riyadi. (2017). Metode Penelitian Kualitatif Kebahasaan. Surakarta: UNS Press.
- Setiawan, T. (2017). Linguistik Korpus dalam Pengajaran Bahasa. Makalah, disajikan dalam seminar nasional Perspektif Baru Penelitian Linguistik Terapan, tanggal 6 Juni 2017 di Program Pascasarjana, Universitas Negeri Yogyakarta.
- Siyoto, S. dan Sodik, A. (2015). Dasar Metodologi Penelitian. Yogyakarta: Literasi Media Publishing.
- Sujaini, H. (2018). Peningkatan Akurasi Penerjemah Bahasa Daerah dengan Optimasi Korpus Paralel. JNTETI, Vol. 7, No. 1, Februari 2018.
- Zaman, M.N., Nababan, M.R., dan Djatmika. (2018). Translation Study of Greetings and Verbs in Accommodating Honorific Expressions of Okky Madasari Novels. Jurnal Pendidikan: Teori, Penelitian, dan Pengembangan 3 (4), 528-537.