JEEF
Journal of English Education Forum
https://jeef.unram.ac.id/index.php/jeef/index

The Battle of the Bots: Comparing Google Translate and Kamus-Sunda's Renditions of *Lutung Kasarung* Story

Dian Budiarti
English Literature Department, UIN Sunan Gunung Djati, Bandung, Indonesia

Received : November 7, 2024
Revised : March 1, 2025
Accepted : March 17, 2025
Published: March 30, 2025

Corresponding Author
Dian Budiarti
dianbudiarti@uinsgd.ac.id

DOI: 10.29303/jeef.v5i1.798
© 2025 The Authors. This open access article is distributed under a (CC-BY License)

Abstract: A number of studies comparing the performances of two different machine translators have been widely carried out, but the study of it involving Sundanese language is still scant and needs to be developed. This present study serves to compare the Indonesian-to-Sundanese translations for a short story performed by Google Translate (GT) and Kamus-Sunda (KS). Data were obtained from their Sundanese translations of five-paragraph Indonesian folklore entitled *Lutung Kasarung*. Furthermore, a number of theoretical frameworks regarding machine translations were used to help analyze the data qualitatively. In general, the findings demonstrate that both GT and KS share a number of similarities and differences in the light of five aspects; (1) improper characters' names, (2) improper target language, (3) the use of speech level, (4) shifts in meaning, and (5) improper sentence structure. Drawing a brief conclusion, this study found that KS makes less errors rather than GT does, so it can be stated that KS performs better than GT in translating the story.

Keywords: *Google Translate; Kamus-Sunda; machine translator*

INTRODUCTION

In recent years, the development of machine translation (MT) has drawn significant attention, particularly in facilitating communication between different languages, including regional ones like Sundanese (Ramadhan et al., 2022). Despite the rapid progress of MT for widely spoken languages such as English and Mandarin, its effectiveness in translating Indonesian to Sundanese remains an open question, as previous studies have shown inconsistencies in translation accuracy for low-resource languages (Susanto et al., 2023). Indonesian and Sundanese, though both belonging to the Austronesian language family, have distinct grammatical structures, morphology, and honorific speech levels (*undak-usuk basa*). Additionally, word choice and sentence structure in Sundanese are strongly influenced by politeness levels and social relationships between speakers and listeners (Primandhika & Munawar, 2019). These linguistic and cultural differences pose challenges for MT systems, which often struggle to capture nuanced meanings accurately (Nada et al., 2024). Therefore, assessing the performance of MT in translating Indonesian to Sundanese is crucial, especially as digital tools are increasingly relied upon for language learning and the preservation of regional languages.

Regarding this, several previous studies have investigated the translation of Sundanese using MT. Komariah (2018) examined Google Translate's (GT) translation of the short story *Si Kabayan Nguseup* 'Kabayan Went Fishing' from Sundanese to English, while Budiarti (2018) evaluated GT's translation from Indonesian to Sundanese using the short story *Manusia Bertubuh Kerdil* 'The Dwarfs'. Both studies found that GT still faced difficulties in accurately translating sentence structures and cultural meanings, leading to the conclusion that this MT system requires further refinement.

In this case, however, it is interesting to note that comparing a short story translation performed by GT and any other machine translators has received almost no attention in the research literature, so this disciplinary knowledge gap triggered this research project. Thus, this present study endeavors to fill the lacuna by exploring this issue deeper to new data through comparing GT and KS translations for the story of *Lutung Kasarung* by addressing at least one research question: how does Google Translate's performance compared to Kamus-Sunda's in translating the story of *Lutung Kasarung*? By way of background, *Lutung Kasarung* is a well-known legend from the Land of Pasundan, one of Sundanese regions in West Java. This story tells of a prince who was punished by the gods and sent to earth in the form of a lutung (a type of monkey). On earth, the prince became lost in the forest and was given the name Lutung Kasarung, which in Sundanese means "the lost lutung". This story was chosen as the subject of analysis because it contains cultural elements and unique Sundanese expressions that may challenge MT systems in preserving meaning accurately.

To clarify, Google Translate (translate.google.co.id) and Kamus-Sunda (kamus-sunda.com/penerjemah.html) are both instant and free machine translators offering Indonesian-Sundanese (or vice versa) translation service. Differently, GT is a multilingual one, while KS is a bilingual one. As they suggest, GT supports the translation from a certain source language to more than a hundred of target languages, while KS only focuses on the translation of Indonesian and Sundanese language pair. Since then, it is hypothesized that KS would show a better performance in this study. Finally, this preliminary research results are expected to contribute theoretically to the development of machine translation studies, particularly for low-resource languages, and practically to help Indonesian-Sundanese machine translator users in selecting the most suitable translation tool for their needs.

RESEARCH METHOD

In order to scientifically answer the research question presented in the previous section, this study employed a descriptive comparative method since the data were examined by primarily collecting textual data and examining it using comparative analysis (Creshwell, 2012). In collecting the data,

a five-paragraph story entitled Lutung Kasarung written in Indonesian language was chosen as the source text considering that it is one of Indonesian folklores from West Java where most Sundanese live. Additionally, it was taken from *dongengceritarakyat.com*, a website that contains a collection of folklores, fairy tales, children's stories, and fables both originated from domestic and foreign countries.

Furthermore, the text was translated into Sundanese by using two different machine translators, Google Translate and Kamus-Sunda. In this study, Indonesian-Sundanese translation was chosen under the consideration that the researcher is a native speaker of those two languages, so it is expected that she is able to give deeper interpretation based on the possessed lingual intuition.

The translated text were then analyzed qualitatively to compare how GT and KS perform in translating the text supported by various theoretical frameworks and the previous studies related to machine translations. Finally, as for presenting the research results more effectively, some codes were used; SL for source language, TL for target language, GT for Google Translate, KS for Kamus-Sunda, V for Verb, C for Complement, A for Adverb, N for Noun, det for Determiner, O for Object, and P1, P2, P3, ... n for Paragraph 1, 2 3, ... n. Additionally, it is noted that Sundanese language has unique vowel characteristics, namely /ɛ/ and /ö/, which are respectively represented by the letters *é* and *eu* (Budiarti, 2024).

RESULT AND DISCUSSION

This section will demonstrate the results and discussion in the light of GT and KS performances in translating the story of *Lutung Kasarung*. Generally, the overall results have revealed that both of them share some similarities and differences in several aspects that will be further elaborated as follows.

Improper Characters' Names

There are four characters mentioned in the whole story, namely *Prabu Tapa Agung*, *Purbasari*, *Purbararang*, and *Lutung Kasarung*. Generally speaking, characters' names or personal names are included into proper nouns – specific referents that serve to distinguish a particular individual from others (Pour, 2009). Therefore, in the process of translating the words, they should be kept as they are (Newmark, 1988). Following such a rule, this study found that the first three characters' names have been properly translated by both GT and KS. However, there are omissions regularly made by GT in translating *Lutung Kasarung* becoming only *Kasarung* for the whole text, while KS keeps translating it becoming *Lutung Kasarung* as it should be. This case can be seen in the following excerpts.

SL: *Akhirnya, ia menarik tangan **Lutung Kasarung**. Tiba-tiba, **Lutung Kasarung** berubah menjadi seorang pangeran tampan.* (P5)
'Finally, she pulled Lutung Kasarung's hand. Suddenly, Lutung Kasarung transformed into a handsome prince.'

GT: *Tungtungna, manéhna ditarik leungeun-Na **Kasarung**. Ujug-ujug, **Kasarung** robah kana pangeran ganteng.* (P5)
'Finally, she was pulled by **Kasarung's** hand. Suddenly, **Kasarung** transformed into a handsome prince.'

KS: *Ahirna, manehna metot panangan **Lutung Kasarung**. Anjog-anjog, **Lutung Kasarung** robih barobah kaayaan saurang pangeran tampan.* (P5)
'Finally, she took **Lutung Kasarung**'s hand. Suddenly, **Lutung Kasarung** transformed into a handsome prince.'

Based on the data, it can be concluded that KS system has a capacity in recognizing Lutung Kasarung as a proper noun, while GT has not. Although further research needs to be conducted, it is assumed that the regular omissions made by GT for the word *Lutung* might refer to the fact that *Lutung* is a kind of animal, whose close meaning to *kera berbulu hitam* 'a black ape' that has been described at the very first time the name appeared as can be seen below.

SL: *Selama di hutan, Purbasari berteman dengan **hewan**, salah satunya dengan seekor **kera berbulu hitam**. Purbasari menamai **kera** itu **Lutung Kasarung**.* (P2)
'During her time in the forest, Purbasari befriended **animals**, one of them being **a black-furred ape**. Purbasari named the **ape Lutung Kasarung.'**

GT: *Bari di leuweung, babaturan Purbasari jeung **sato**, salah sahijina ku **kunyuk buluan hideung**. Purbasari **kunyuk** ngaranna eta **Kasarung**.* (P2)
'While in the forest, Purbasari befriended **animals**, one of them being **an ape with black fur**. Purbasari named the ape **Kasarung**.'

In other words, it is interesting to note that GT system probably intends to prevent such a redundancy, but unfortunately, this case is not the right place.

Improper Target Language

In some cases, both GT and KS do not translate all the words into the TL. This study found that there are 5 English words and 3 Indonesian words (keep being in the SL) in the GT translation, while there are 10 Indonesian words and 2 code-mixed Indonesian-Sundanese words in the KS translation. These simultaneously confirm Anisa (2018), Budiarti (2018), and Komariah's (2018) studies which also found the similar results, except for the one performed by KS, in which it produces, surprisingly, code-mixed words. This case, several of which, can be seen as follows.

SL – English

SL: (1) *Ia **berniat** mencelakakan Purbasari.* (P1)
'She **intended** to harm Purbasari'
(2) *Purbararang mengakui kesalahannya dan **meminta maaf**.* (P5)
'Purbararang admitted her mistake and **apologized**.'

GT: (1) *Anjeunna **intends** ngarugikeun Purbasari.* (P1)
'Purbararang **intends** to harm Purbasari'
(2) *Purbararang ngaku kasalahan sarta **apologized**.* (P5)
'Purbararang admitted her mistake and **apologized**.

As stated previously, the translations from Indonesian to English – the outer language in this study – only occur in GT, not in KS. This might be correlated to the fundamental differences in their translation systems. GT operates as a multilingual translator, supporting translations across a vast range of language pairs, including Indonesian to English. In contrast, KS is designed specifically as a bilingual system, solely focusing on the Indonesian-Sundanese language pair. Given this limitation, it is unsurprising that KS does not

facilitate Indonesian-to-English translation, as its system lacks any built-in English reference database that GT, with its broader linguistic scope, inherently possesses.

SL – Indonesian (SL)

SL: *"Seorang ratu harus mempunyai suami yang **tampan**. Mana calon suamimu?"* (P4)
'A queen must have a **handsome** husband. Where is your future husband?'

GT: *"A ratu kudu boga salaki **ganteng**. Dimana calon salaki?"* (P4)
'A queen must have a **handsome** husband. Where is future husband?'

KS: *"Saurang ratu kedah ngagaduhan salaki anu **tampan**. Manten calon salaki anjeun?"* (P4)
'A queen must have a **handsome** husband. Where is your future husband?'

Differently, Indonesian-Indonesian translations occur in both GT and KS that one of which they interestingly show the same error in translating *tampan* 'handsome' into *ganteng* and *tampan* respectively which share the same meaning.

SL – Code-mixed Indonesian-Sundanese

SL: (1) *Sang Raja terpaksa **mengasingkan** Purbasari ke dalam hutan.* (P2)
'The King was forced **to exile** Purbasari to the forest.'
(2) *Ia berniat **mencelakakan** Purbasari.* (P1)
'She intended **to harm** Purbasari.'

KS: (1) *Sang Raja kapaksa **mengasingkeun** Purbasari ka jero leuweung.* (P2)
(2) *Manehna boga niat **mencelakakeun** Purbasari.* (P1)

Mengasingkeun and *mencelakakeun* consist of partly Indonesian and Sundanese languages. They are respectively started with Indonesian prefixes *meng-* and *men-*, followed by Indonesian base words *-asing-* and *-celaka-*, and ended with Sundanese suffix *-keun.* In other words, the one translated is only its suffixes. This word formation does not fulfill any standard form of any languages, so it makes us need more data to think further why. To put them clearer, the correct ones should be *ngasingkeun* and *nyilakakeun.*

The Use of Speech Level

Sundanese is one of local languages in Indonesia implementing speech levels – *basa kasar* (low level/rough), *basa loma* (refined low level/neutral), and *basa lemes* (high level/polite). According to Suwarno & Yanwar (2019), *basa kasar* is normally used in an emotional situation, *basa loma* is usually used in casual interactions, particularly among peers, close acquaintances, or when addressing younger individuals, while *basa lemes* is usually addressed to older or respected people or in formal situations. Regarding this, GT and KS show a significant difference. In consistent with Budiarti (2018), GT tends to use *basa loma* more. Since then, she further suggests that it cannot be used for all kinds of text, such as religious or wedding ceremony speech texts, which basically need vocabularies in *basa lemes*. Oppositely, this study found that KS tends to use *basa lemes* more. This difference can be clearly seen in the table below.

Table 1. Speech Level Use of GT and KS

	GT	KS
P1	*Prabu Tapa Agung kagungan dua putri, **ngaranna** Purbararang sarta Purbasari.*	*Prabu Tapa Agung ngagaduhan dua jalmi putri, **namina** Purbararang sarta Purbasari.*
	'Prabu Tapa Agung had two daughters, **named** Purbararang and Purbasari.'	
P1	*Hiji **poé**, raja mutuskeun **pikeun** nunjuk Purbasari jadi ratu.*	*Hiji **dinten**, raja megatkeun **kanggo** nunjuk Purbasari barobah kaayaan ratu.*
	'One **day**, the King decided **to** make Purbasari a queen.'	
P2	***Kusabab** sieun infecting **sakabeh** pang-eusi karajaan, ...*	***Margi** rempan menulari **sakumna** nunyicing-an karajaan, ...*
	'**Because of** feeling afraid of infecting **all** royal inhabitants, ... '	
P4	*"A ratu **kudu boga** salaki ganteng. Dimana calon salaki?"*	*"Saurang ratu **kedah ngagaduhan** salaki anu tampan. Manten calon salaki anjeun?"*
	"A queen **should have** a handsome husband. Where is your future husband?"	
P5	Sadar **geus nyieun kasalahan**, Purbara-rang ngaku **kasalahan** sarta apologized.	Sadar **atos migawe kalepatan**, Purbararang ngajirim **kalepatan** na sarta **neda hapunten**.
	'Consciously **made a mistake**, Purbararang admitted **it** and **apologized**.'	

In this study, the research object is a folklore that aims at entertaining readers (Ihueze, 2015), so it basically needs *basa loma* to show such an impression of intimacy. Therefore, it can be stated that the language performed by GT is, to some extent, a way more appropriate than that performed by KS.

Shifts in Meaning

As usually found in most machine or even, sometimes, human translations, the case of shifts in meaning almost always occurs. The following excerpts present some examples of them found in this study.

SL: (1) *Akhirnya, ia **menarik** tangan Lutung Kasarung.* (P5)
'Finally, she **pulled** Lutung Kasarung's hands.'
(2) *Sementara itu di istana, Purbararang **memutuskan** untuk melihat kondisi adiknya.* (P4)
'Meanwhile in the palace, Purbararang **decided** to see her sister's condition.'

GT: (1) *Tungtungna, manéhna **ditarik** leungeun-Na Kasarung.* (P5)
'Finally, she **was pulled** by Lutung Kasarung's hands.'

KS: (2) *Samentara eta di karaton, Purbararang **megatkeun** kanggo ningali kaayaan adi na.* (P4)
'Meanwhile in the palace, Purbararang **broke** (it) to see her sister's condition.'

As can be seen above, both GT and KS seem to indicate their incapacity in translating all the words faithfully as they should be. For instance, GT translates *menarik* 'pulled' into *ditarik* 'was pulled'. Such translation shows that there is a change in terms of the sentence's main verb from the active form into the passive one. Hence, it simultaneously gives the sentence a totally different meaning since there will be a subject and object alteration. Meanwhile, an instance of this case made by KS is when it translates *memutuskan* 'decided' into *megatkeun* 'broke', which also gives a totally different meaning. Based on the context, the proper one should be

mutuskeun 'decided'. To go further, both *megatkeun* and *mutuskeun* can be expressed by the same word *memutuskan* by Indonesian. In other words, the error made by KS can be related to the certain polysemy they both share.

Improper Sentence Structure

In terms of sentence structure, generally, several errors only occur in GT translation, while in KS does not. Some evidences can be seen as follows.

SL: (1) *Setelah kejadian itu, Purbasari* ***kembali menjadi seorang ratu***. (P5)
'After the incident, Purbasari **became a queen again**.'
(2) *Ia pun pergi menemui* ***seorang penyihir***. (P1)
'She then went to meet **an enchanter**.'
(3) *Purbasari* ***menamai kera itu*** *Lutung Kasarung.* (P2)
'Purbasari **named the ape** Lutung Kasarung.'

GT: (1) *Sanggeus kajadian éta, Purbasari* ***deui janten ratu***. (P5)
(2) *Manéhna indit ka tingali* ***dukun a***. (P1)
(3) *Purbasari* ***kunyuk ngaranna eta*** *Kasarung.* (P2)

Starting from the Sentence (1), *kembali menjadi ratu* 'returned to be a queen' [V + C] is translated into *deui janten ratu* '*again became a queen' [A + V + C]. It can be seen that the error relies on the use of the word *deui* 'again' to express *kembali* 'returned'. Basically, those two words refer to the same meaning. However, since there is a changed word function from a verb into an adverb, the correct structure to follow Sundanese rules should be *janten ratu deui* 'became a queen again' [V + C + A]. This case might be due to GT's use of word-for-word translation strategy.

Meanwhile, Sentence (2) shows that *seorang penyihir* 'an enchanter' [det + N] is translated into *dukun a* '*enchanter a' [N + det]. Almost similar, Sentence (3) shows that *menamai kera itu* 'named that ape' [V + O] is translated into *kunyuk ngarana eta* '*ape named that' [O + V + O]. However, it has not been possible to provide any definitive reason to these findings.

Finally, to help drawing a conclusion, the overall findings regarding the comparison of GT and KS translations for the story of *Lutung Kasarung* will be presented in the following table.

Table 2. The Comparison of GT and KS Translations for the Story of *Lutung Kasarung*

Findings		GT	KS
Improper Characters' Names		Occurred	-
Improper Target Language	English	Occurred	-
	Indonesian (SL)	Occurred	Occurred
	Code mixed Indonesian-Sundanese	-	Occurred
The Use of Speech Level		*Basa loma*	*Basa lemes*
Shifts in Meaning		Occurred	Occurred
Improper Sentence Structure		Occurred	-

CONCLUSION

The foregoing results and discussion have demonstrated that both Google Translate (GT) and Kamus-Sunda (KS) share several similarities and differences in various aspects as well as show their individual strengths and weaknesses. Drawing a brief conclusion, this study found that KS produces fewer errors compared to GT, indicating that KS performs better in translating the story of *Lutung Kasarung*. This suggests that KS, as a specialized bilingual translator, may be more reliable for handling Indonesian-Sundanese translations than GT, which is a multilingual system designed for a broader range of language pairs.

However, it is important to note that this study focused solely on a single short story, which limits the generalizability of the findings. Different text genres, such as academic papers, news articles, or conversational texts, may pose different challenges for machine translation systems. Therefore, while KS showed superior performance in this context, we are not yet in a position to determine which machine translator is the most suitable for all types of texts. Further research is highly recommended to address this gap by examining a more diverse range of texts and translation contexts. By continuing research in this area, we can contribute to the development of more accurate and culturally sensitive machine translation systems, ultimately supporting the preservation and accessibility of regional languages like Sundanese in the digital age.

REFERENCES

_____. (2017). *Kumpulan Cerita Rakyat Jawa Barat: Sangkuriang dan Lutung Kasarung*. [Online]. http://dongengceritarakyat.com/kumpulan-cerita-rakyat-jawa-barat-sangkuriang-dan-lutung-kasarung/.

Anisa, N. (2018). *Terjemahan Ungkapan Bahasa Sunda-Bahasa Indonesia dengan Berbantuan Google Translate* (Unpublished research report). Universitas Gadjah Mada.

Budiarti, D. (2018). *Penerjemahan antarbahasa Serumpun Berbasis Mesin: Analisis Terjemahan Google Translate dari Bahasa Indonesia ke Bahasa Sunda.* (Unpublished research report). Universitas Gadjah Mada. Yogyakarta.

Budiarti, D. (2024). Konjungsi 'Jeung' dalam Bahasa Sunda: Kajian atas Perilaku Sintaksis dan Semantisnya. *Pendekar: Jurnal Pendidikan Berkarakter*, *2*(6), 1–16. https://journal.politeknik-pratama.ac.id/index.php/Pendekar/article/view/811.

Creswell, J. W. (2012). *Educational research: planning, conducting, and evaluating quantitative and qualitative research 4th edition.* Boston: Pearson Education. https://doi.org/10.51903/pendekar.v2i06.811.

Ihueze, O. A. (2015). Folklore in literature: A tool for culture preservation and entertainment. *International Journal of Humanities and Social Science Invention*, *4*(4), 57-61. https://www.researchgate.net/publication/375744530_Folklore_in_Literature_A_Tool_for_Culture_Preservation_and_Entertainment.

Komariah, S. (2018). *Pergeseran Makna dan Perubahan Struktur Bahasa Sunda ketika Diterjemahkan ke dalam Bahasa Inggris pada Cerita Si Kabayan Nguseup.* (Unpublished research report). Universitas Gadjah Mada. Yogyakarta.

Nada, A. Q. et al. (2024). Optimizing Indonesian-Sundanese bilingual translation with Adam-Based Neural Machine Translation. *Jurnal Resti: Rekayasa Sistem dan Teknologi Informasi*, *8*(6), 690-700. https://doi.org/10.29207/resti.v8i6.6116.

Newmark, P. (1988). *A textbook of translation.* London: Prentice Hall.

Pour, B. S. (2009). How to translate personal names. *Translation Journal*, *13*(4). https://translationjournal.net/journal/50proper.htm.

Primandhika, R. A., & Munawar, A. (2019). Experiment on a transformer model Indonesian-to-Sundanese neural machine translation with Sundanese speech level evaluation. *Proceedings of the 12th International Conference on Applied Linguistics (CONAPLIN 2019)*, 1-6. https://doi.org/10.2991/assehr.k.210427.069.

Ramadhan, T. I., Ramadhan, N. G., & Supriatman, A. (2022). Implementation of Neural Machine Translation for English-Sundanese Language using Long Short Term Memory (LSTM). *Building of Informatics Technology and Science (BITS)*, *4*(3). https://doi.org/10.47065/bits.v4i3.2614.

Susanto, L., Diandaru, R., Krisnadhi, A., Purwarianti, A., & Wijaya, D. (2023). Replicable benchmarking of Neural Machine Translation (NMT) on low-resource local languages in Indonesia. *Proceedings of the First Workshop in South East Asian Language Processing*, 100–115. https://doi.org/10.18653/v1/2023.sealp-1.8.

Suwarno, B. & Yanwar, A. P. (2019). Speech level in some Indonesian regional languages and implication for conservation. In *Proceeding of the 2nd International Conference on Social Science and Character Educations (ICoSSCE 2019)*, 328-332. https://www.atlantis-press.com/proceedings/icossce-19/125933420.

www.kamus-sunda.com/penerjemah.html

www.translate.google.co.id