

A Study on Teen Slang Translation of Asian Languages

Tiffany Carissa*
Pasin Harisadee[‡]

JiaHui Wang*^b
Sivakorn Seinglek[‡]

Suravich Thammapatee*[‡]
Shiho Hoshi Nobesawa*

Keisuke Magara*
Shiho Hoshi Nobesawa*

1. Introduction

Machine translation facilitates communication between young people and enables them to access information across language barriers. It also allows young people to share ideas and perspectives with a wider audience, regardless of language.

This paper focuses on the translation of teen slangs in Asian languages.

Sakti et al. collected 8 Asian languages and English data related to spoken languages[1]. They made a system to translate common spoken travel conversations from a given source language into multiple target languages in order to facilitate multiparty travel conversations.

Atarashi et al. conducted a study aimed at obtaining appropriate equivalent term in another language for Japanese teen slangs[2]. He et al. proposed a translation method for teen slangs to bilingual phrases based on a method focusing on the similarity and naturalness between the translated sentences[3, 4].

Our aim is to analyse the characteristics of Asian teen slangs, and to clarify the challenges of translating teen slangs for Asian languages.

2. Asian Teen Slang Corpora

In order to grasp the characteristics of Asian teen slangs, several teen slangs and sentences containing them were acquired for 4 languages in this study: Japanese, Chinese, Indonesian and Thai.

2.1. Japanese Teen Slang Corpus

According to Atarashi et al., 10 teen slangs (Table 1) were selected on the basis of their frequency of occurrence in the past five years in in Google Trends¹. For each slang

Table 1: Japanese Teen Slang Set

エモ	リセマラ	ブーメラン	不可避
emotional	reset marathon	backfire	inevitable
ズッ友	リアコ	マウント	すこ
friend forever	in love	one-up someone	like
推し	限界オタク		
fav	deranged nerd		

in Table 1 we obtained 10 sentences from Twitter. We use the set of these 100 Japanese sentences as a Japanese Teen Slang Corpus.

2.2. Chinese Teen Slang Corpus

As for the Chinese language, a total of 10 common teen slangs used in daily life (Table 2) were obtained based on The China Project², which deals with articles on Chinese teen slangs. For each of the slangs in Table 2, ten

Table 2: Chinese Teen Slang Set

一种很新的	刺客	退!退!退!	嘴替
very new	wallet killer	retreat	mouthpiece
大冤种	栓Q	电子榨菜	pua
big loser	thank you	mealtime fun	brainwash
团长		小镇做题家	
pandemic group-buy leaders		small town test taker	

sentences each containing these words were obtained from the Weibo³. We use the set of these 100 Chinese sentences as a Chinese Teen Slang Corpus.

* Tokyo City University

^b Dalian Jiaotong University

[‡] Mahidol University

[‡] Sirindhorn International Institute of Technology, Thammasat University

¹Google Trends, <https://trends.google.co.jp/trends>.

²The China Project, <https://thechinaproject.com>.

³Weibo, <https://weibo.com>.

2.3. Indonesian Teen Slang Corpus

For the Indonesian language, we have selected a total of 10 Indonesian teen slangs (Table 3) which are listed in the Indonesian Dictionary⁴. The 10 teen slangs listed on the news website detik⁵ were selected. For each slang in

Table 3: Indonesian Teen Slang Set

alay	ambyar	bokap	bokek	julid
tacky	shattered	father	broke	jealous
nyokap	kepo	kicep	lebay	mager
mother	busybody	silence due to fear	overacting	lazy

Table 3 we obtained 10 sentences from Twitter. We use the set of these 100 Indonesian sentences as an Indonesian Teen Slang Corpus.

2.4. Thai Teen Slang Corpus

For the Thai language, a total of 10 teen slangs (Table 4) were obtained based on Young Happy⁶ article, which explains about Thai teen slangs. For each slang in Table

Table 4: Thai Teen Slang Set

เกินต้าน (kern tarn)	ว่าซน (wasan)	ตย (tui)	ป้ายยา (pai yar)
irresistible	just like that	dead	persuade
แซบ (sap)	ดีด (due)	ซิ่ง (kzing)	ตัวตึง (tua tung)
delicious looking	nice	brag	the best person
อึ้งย้งวะ (iyhangwa)	ใครไหวไปก่อน (kai wai pai gorn)		
what the hell	fed up with		

4 we obtained 10 sentences from Twitter. We use the set of these 100 Thai sentences as a Thai Teen Slang Corpus.

3. Translations of Asian Teen Slangs

3.1. Evaluation

The translation results were divided into words and sentences, each of which was judged on a 4-point scale to determine the final evaluation (Table 5). The part marked

Table 5: Slang Translation Rating Scale

translation	slang sentence		
	meaning	syntax	impresssion
successful	accurate	natural	same
estimable	estimable	understandable	similar
inestimable	inestimable	not understandable	different
failed	slang translation failed		

“slang meaning” indicates whether the meaning is correctly translated in the teen slangs under test itself (Table 5). For sentences, the appropriateness of the syntax is evaluated, as well as whether the impressions given by the sentences are consistent. If a sentence fails in any one of the 3 items, the translation result will not show the correct meaning, and in this study, each sentence is rated as the lowest value in each of these 3 items. The cases where no translation was carried out, such as when the teen slang part of the sentence is shown as it is in Roman characters, are tagged as “failed.” In our research we take “successful” and “estimable” in Table 5 as success, and “inestimable” and “failed” are referred as failure.

3.2. Japanese-English Slang Translation

A translation study from Japanese to English was carried out and candidate bilingual phrases were extracted. We used Google Translate and DeepL Translator and obtained 200 candidate sentences.

The translation results are evaluated in terms of slang meaning, sentence structure and sentence impression, according to Table 5 (Figure 1). There are various Japanese slangs that are acronyms of common words or even borrowed words, such as “エモ (emo)” which was an acronym

⁴Kamus Besar Bahasa Indonesia, <https://kbbi.kemdikbud.go.id/Beranda>.

⁵detik, <https://detik.com>.

⁶Young Happy, <https://younghappy.com/>.

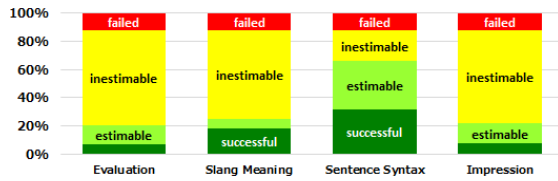


Figure 1: Slang Translation Result (Japanese to English) of an English word “emotional.” These acronyms are often resulted in inestimable output, such as “ズッ友 (zutto: friend forever)” which only shown as “friend (友: tomo),” for the machine translators were unable to recognize the shortened “ずっと (zutto: forever)” inside the slang.

3.3.Chinese-English Slang Translation

A Chinese to English translation test was carried out and candidate bilingual phrases were extracted. We used Google Translate and DeepL Translator and obtained 200 candidate sentences.

The translation results are evaluated in terms of phrase intention, sentence structure and sentence impression (Figure 2). For Chinese teen slangs, a lot of them are

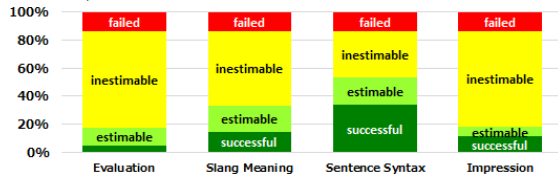


Figure 2: Slang Translation Result (Chinese to English) created as a whole phrase which holds a whole different meaning as compared to when the characters are separated. For example, “电子 (electronic)” and “榨菜 (pickled mustard tuber)” are combined into one phrase which means “video to watch during mealtime” as in Table 2. This is because Chinese words are not separated by space which makes it hard for machine translators to recognize words combined into phrase as a slang. “栓 Q” stands for English word “thank you,” just showing its pronunciation.

3.4.Indonesian-English Slang Translation

A translation test from Indonesian to English was carried out and candidate bilingual phrases were extracted. We used Google Translate and DeepL Translator and obtained 200 candidate sentences.

The translation results are evaluated in terms of phrase intention, sentence structure and sentence impression (Figure 3). For Indonesian teen slangs, a lot of the slangs

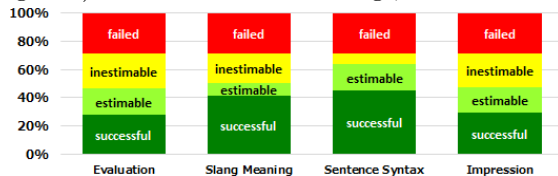


Figure 3: Slang Translation Result (Indonesian to English) are borrowed words from other languages. This is because there are hundreds of language spoken in Indonesia. These data had relatively good result. On the other hand, acronym and newly created word had a bad result on the evaluation. As example, “kicep” is a newly created word and most of the translation was incorrect. What should have been “silence due to fear” ended up translated as “fast.” This is probably because the machine translation often mistranslate it into “sakicep” which is a word from Sunda language which means “fast.”

3.5.Thai-English Slang Translation

A translation study from Thai to English was carried out and candidate bilingual phrases were extracted. We used Google Translate and Bing Translator and obtained 200 candidate sentences.

The translation results are evaluated in terms of phrase intention, sentence structure and sentence impression (Figure 4). For Thai teen slangs, some of them are based on already existing words. But when used as a

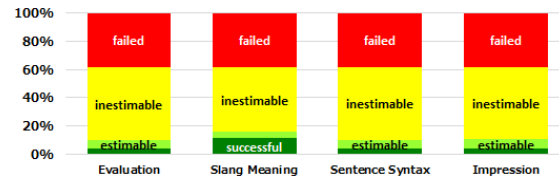


Figure 4: Slang Translation Result (Thai to English) slang word, there is semantic change towards said existing word. For example, “ต้วตึง (tua tung)” for slang means “the best person,” however for normal Thai language the word means “tensioner.” Most people use it more for slang rather than the common word these days.

4.Discussion

The 4 languages evaluated each has their own unique rules and characteristics. Table 6 shows our classification of slang types regarding the influence in translation. Bor-

Table 6: Slang Type Classification for Translation

borrowed word	phrase	neologism
semantic change	acronym	different form

rowed words (such as “栓 Q”), phrases (“电子榨菜”) and neologisms (“推し”) indicate new slang phrases of the language. Semantic changes (“マウント”), acronyms (“ズッ友”) and different forms (“すこ”) indicate phrases created based on existing words. For borrowed words and semantic changes, existing words are used as slangs. Phrases and acronyms are the connections or the omissions of existing words. Neologisms and different forms can be regarded as new words, though they also have original words.

Figure 5 shows the slang types of each language. For

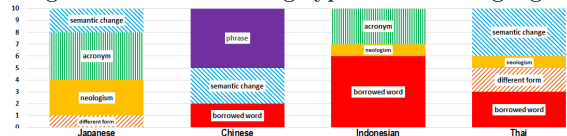


Figure 5: Teen Slang Types in Asian Languages

Japanese slangs, there is a problem with recognizing the acronym of slangs which causes half of the mistranslation. Chinese teen slangs are often written as a phrase which causes problem in word segmentation and causes mistranslation into literal meaning of the characters. For Indonesian slangs, acronyms and newly created words tend to be mistranslated. In case of Thai language, as there are existing words with the same letters, it is still hard to differentiate the meaning of the word in a sentence.

By searching the characteristics of these 4 Asian languages, it was found that in order to increase the accuracy of teen slang translation it is best to focus on the main problem for each individual languages even further.

5.Conclusion

The problem of teen slangs in Asian languages is that they still cannot be translated correctly. In order to improve the translation of Asian teen slangs, we translated Asian teen slangs into several translators and analyzed their characteristics and classify them. The classification allowed us to examine the problems that need to be solved in the translation of Asian teen slangs. We believe that a machine capable of automatically classifying teen slangs is needed to improve the quality of machine translation.

References

- [1] S.Sakti, M.Paul, A.Finch, S.Sakai, T.T.Vu, N.Kimura, C.Hori, E.Sumita, S.Nakamura, J.Park, C.Wutiwiwathchai, B.Xu, H.Riza, K.Arora, C.M.Luong, H.Li, “A-STAR: Toward translating Asian spoken languages,” Computer Speech & Language, Vol.27, pp.509–527, 2013.
- [2] T.Atarashi, K.X.Tan, J.H.Wang, T.Carissa, Y.S.Li, L.Y.Wang, S.H.Nobesawa, “Japanese Teen Slang Translation from Multiple Candidates,” National Convention of IPSJ, pp.815–816, 2023. Written in Japanese.
- [3] W.Y.He, A.A.Khadanovich, K.Fujita, S.Asanuma, R.Tamura, S.Machida, K.Yamada, S.H.Nobesawa, “A Study on Paraphrasing for Translation of Japanese Young People’s Phrases,” IPSJ SIG Notes, vol.2018-NL-236, no.7, pp.1–6, 2018. Written in Japanese.
- [4] W.Y.He, K.Fujita, S.Asanuma, M.Q.Wang, X.J.Guo, K.Nakajima, S.Machida, S.H.Nobesawa, “Analysis of Wakamono Kotoba for Japanese-Chinese Translation,” National Convention of IPSJ, pp.323–324, 2018. Written in Japanese.