

## Classification of Serial Verb Constructions in Indonesian based on the corpora survey

Kanta HASEGAWA

Kanda University of International Studies

---

Serial Verb Construction (SVC) is a monoclausal construction consisting of multiple independent verbs with no element linking them and no predicate-argument relation between the verbs according to Haspelmath (2016: 292).

There are several studies on SVCs in Indonesian. However, none of them have a unified view, and their classification varies from study to study. The results of those studies are summarized in a total of 15 types of semantic relations: aspect, auxiliaries, causatives, cause-effect, coordinated actions, manner, motion & direction, overt connectors, purpose, resultatives, simultaneous actions, source, state, supplementary explanation, and synonym & antonym.

Considering the above, here I mainly use MALINDO Conc. (Nomoto et al. 2018), a reclassified corpus based on Leipzig Corpora Collection for Malay and Indonesian, to investigate the usages of SVCs and collect the examples that have not been mentioned previously. MALINDO Conc. is highly searchable and suitable for finding SVCs because it can search sentences by specifying prefixes and the collocation as well.

In this way, I finally assert that there are eight categories of SVCs: cause-effect, destination, manner, motion & direction, purpose, sequential action, simultaneous action, and source.

Cause-effect relation represents a structure in which *karena* 'because' can be inserted between the verbs. Therefore, a verb representing effect often precedes a verb representing cause, but sometimes a verb representing cause precedes.

Destination relation refers to a structure in which *sampai* 'until' can be inserted between the verbs. This relation is very few.

Manner relation represents the manner of the head verb. Often *dengan* 'with' can be inserted in this structure.

Motion & direction relation includes peculiar structures found only in verbs representing motion or direction. No element can be inserted between the verbs. The second verb of this relation tends to be prepositional.

Purpose relation represents a structure in which *untuk* 'for' can be inserted between the verbs.

Sequential action relation refers to the two verbs making a line in chronological order. In this relation, *dan* can be inserted between the verbs.

Simultaneous action relation refers to both verbs' actions carried out simultaneously. In this relation, *sambil* 'while' can be inserted between the verbs.

Source relation represents that one action carried out before another. In other words, it shows the opposite flow of sequential action relation. In this relation, *dari* 'from' can be inserted. This relation is very few.

Other classifications should not be included in SVCs, following Haspelmath's definition. However, depending on the analysis, there are also relations that should not be included in SVCs introduced here. Those issues will be discussed.

### References

- Haspelmath, M. 2016. The Serial Verb Construction: Comparative Concept and Cross-linguistic Generalizations. *Language and Linguistics* 17(3). 291–319. <https://doi.org/10.1177/2397002215626895>.
- Nomoto, Hiroki, Shiro Akasegawa & Asako Shiohara. 2018. Reclassification of the Leipzig Corpora Collection for Malay and Indonesian. *NUSA: Linguistic studies of languages in and around Indonesia* 65. 47–66. <https://doi.org/10.15026/92899>.

# The effects of word class, semantic variability and lexical characteristics on Malay-English bi-directional translation

Soon Tat Lee<sup>1</sup>, Walter J. B. van Heuven<sup>2</sup>, Jessica M. Price<sup>1</sup> and Christine Xiang Ru Leong<sup>1</sup>

<sup>1</sup>*University of Nottingham, Malaysia;*

<sup>2</sup>*University of Nottingham, Nottingham, United Kingdom*

Translation equivalents, or words from two different languages that share similar meaning, are widely used in bilingual research (e.g., Jouravlev & Jared, 2020). However, selection of appropriate translation equivalents can be challenging because many words have more than one translation from one language to the other. For example, the English word *thick* can be translated in Malay as *pekat* (meaning for a liquid substance) or *tebal* (for a solid substance). This paper presents the first Malay-English translation norms used to investigate the impact of word class, semantic variability, and Malay and English lexical characteristics on translation ambiguity. Distinct groups of proficient Malay-English bilinguals performed translation tasks in two experiments. In Experiment 1, English translations for 1,004 Malay words were first collected via forward translation (FT;  $N = 30$ ), and in Experiment 2, Malay translations were gathered for the 845 English words obtained from the FT task using backward translation (BT;  $N = 30$ ). Translation equivalents for source words with different word classes, number of senses, word frequency and length were presented in the translation norms. The data revealed a high prevalence of translation ambiguity in both translation directions. Specifically, verbs, adjectives, and class ambiguous words were more translation ambiguous than nouns. Furthermore, within-language semantic variability (number of senses carried by a word) and word length were positively correlated with translation ambiguity, whereas word frequency only positively correlated with translation ambiguity in FT. Word length and word frequency of the source words and their translations also positively correlated with each other. The overall results will be compared with translation studies involving other language pairs (e.g., Spanish-English: Prior et al., 2007; Chinese-English: Wen & van Heuven, 2017) and the implications will be discussed. The translation norms will be a useful resource for researchers conducting bilingual language studies involving Malay and English.

*Keywords:* translation equivalents, translation ambiguity, bilingualism

## References

- Jouravlev, O., & Jared, D. (2020). Native language processing is influenced by L2-to-L1 translation ambiguity. *Language, Cognition and Neuroscience*, 35(3), 310–329. <https://doi.org/10.1080/23273798.2019.1652764>
- Prior, A., MacWhinney, B., & Kroll, J. F. (2007). Translation norms for English and Spanish: The role of lexical variables, word class, and L2 proficiency in negotiating translation ambiguity. *Behavior Research Methods*, 39(4), 1029–1038. <https://doi.org/10.3758/BF03193001>
- Wen, Y., & van Heuven, W. J. B. (2017). Chinese translation norms for 1429 English words. *Behavior Research Methods*, 49. <https://doi.org/10.3758/s13428-016-0761-x>

## Epistemic and affective meanings of discourse marker *nih* in colloquial Jakartan Indonesian

One of the discourse markers in colloquial Jakartan Indonesia that is understudied is *nih*. Sneddon (2006) mentioned that *nih* functions to give a stress. There is no detailed explanation on how the speakers use *nih* to build discourse. The present study followed a suggestion by Djenar (2018) to scrutinize discourse markers based on modality meaning. Here, I focused on the epistemic and affective meanings of *nih*. The present study is a discourse analysis that focuses on the sequences of turn taking and the information flow. There are twenty two utterances of *nih* in the present study. The data is taken from the videos of food reviews on YouTube. This genre is getting more popular now. The interactions are between two reviewers or the reviewers and the seller. In the YouTube video, the reviewers are aware that they have audience (video watchers). Thus, they do not only talk to each other but also to the audience.

The speakers build the description by asserting the epistemic and affective meaning first. Usually, the utterance with *nih* is used in initial part of information given. It functions to attract hearers' attentions. As the hearers give more attention, the speakers give more detailed information about the object or actions. Mostly, the speakers use *nih* when they describe the object such as food as in *Ni benar-benar asam manisnya dapat banget nih*. Furthermore, it occurs when the speakers mention actions. The actions could take place at the moment or they are about to do by the speakers as in *Nih gue mau nyobain nih, yang ini*. The speakers try to engage the watchers in the activities of trying the food.

In terms of epistemic meaning, *nih* indicates shared information and feeling. The speakers position themselves as the one who have knowledge and share it to the watchers. Thus, the watchers who are not there receive knowledge from the speakers. *Nih* also delivers the affective meaning in the utterances. By using *nih*, the speakers can pinpoint their attitude towards the object mostly the food more clearly to the watchers who do not taste it. When the speakers shared their feeling, mostly they give assessment about the food.

*Nih* tends to collocate with demonstratives *nih*, *ni*, and *ini*. With the use of demonstratives, the speakers point the objects. It makes the interlocutors and the viewers pay attention to the same objects. Thus, the speakers want to make sure that they are in the same page. The collocations indicate how the speakers guide the hearers to notice the objects or the actions.

## References

- Djenar, D. N., Ewing, M., & Manns, H. 2018. *Style and intersubjectivity in youth interaction* (Issue February). Walter de Gruyter Inc. <https://doi.org/10.1515/9781614516439>
- Sneddon, J. N. 2006. Colloquial Jakartan Indonesian. *The Australian National University*. <https://doi.org/10.1017/CBO9781107415324.004>

## **Baru and baharu: Which is new?**

It is generally believed that of the two forms meaning ‘new’ in Malay, namely *baru* and *baharu*, *baru* is new and *baharu* is old, with the former developed from the latter. Another belief that has gained currency in Malaysia recently as a rationale for prescriptivists to replace certain instances of *baru* by *baharu* is that the two forms used to be distinguished based on their syntactic categories: *baru* was an adjective meaning ‘new’ whilst *baharu* was an adverb meaning ‘only just’. This paper argues that these two popular beliefs are both problematic based on the Classical and Early Modern Malay texts available in the Malay Concordance Project (MCP; Proudfoot 1991). MCP contains 5.7 million words from more than 150 sources written or compiled between 1300s and 1950s.

The MCP data reveals that both *baru* and *baharu* have co-existed at least since 1370s, although one can trace back to Proto Malayo-Polynesian \**baqeru* (Blust & Trussel 2010). The table below shows the frequencies of various forms based on the roots *baru* and *baharu* during different periods. The forms occurring in place names (i.e. *Johor Bah(a)ru*, *Kota B(a)haru*, *kampung Baharu*) are not included. The attested forms (after normalization) are listed in (1). The presence of forms that are extremely rare in Modern Malay such as (*mem(di)barui* and *baharu-baharu*) suggests that the two forms were distinguished less clearly before than now, contrary to the second belief above. Moreover, free variations in a single text are observed in many texts.

Period	<i>baru</i>	<i>baharu</i>	% of <i>baru</i>
1370s–1590s	27	57	32.1
1600s–1640s	14	109	11.4
1650s–1690s	7	15	31.8
1700s–1740s	4	22	15.4
1750s–1790s	54	93	36.7
1800s–1840s	76	198	27.7
1850s–1890s	107	221	32.6
1900s–1950s	58	95	37.9

- (1) a. *baru, barulah, barula, barunya, perbaruan, barukan, membarukan, membarui, dibarui, baru-baru, baru-baruan, sebaru-barunya, membaru-barukan, dibaru-barui*  
b. *baharu, baharulah, baharukah, baharuan, baharunya, pembaharuan, berbaharu, beharulah, baharukan, membaharukan, membaharukannya, dibaharukan, dibaharukannya, baharui, baharuihi, membaharui, dibaharui, dibaharuilah, dibaharuinya, pembaharui, memperbaharui, baharu-baharu, baharu-baharuan, baharu-baharukan*

Regarding the second belief, *baru* and *baharu* were both used as an adjective and an adverb. In fact, for *baharu*, the adverbial use was far more common than the adjectival use in the past.

- (2) Telah setahun lamanya hamba berkawal ini, **baharulah** Tuanku melihat patik hadir pada malam ini. ‘I’ve stood guard for a year, and it is not until tonight that Your Majesty saw my presence.’ (Hikayat Bayan Budiman 18:26, 1371)

## **References**

- Blust, Robert & Stephen Trussel. 2010. *Austronesian Comparative Dictionary*, web edition.  
<https://www.trussel2.com/acd/>  
Proudfoot, I. 1991. Concordances and Classical Malay. *Bijdragen tot de Taal-, Land- en Volkenkunde* 147: 74–95. doi: 10.1163/22134379-90003200

## Philippine cognate of the Malay applicative suffix *-kan*

The present study aims to identify a Philippine cognate of the Malay applicative suffix *-kan*. Focusing on the instrumental function of the suffix, it mainly examines two Philippine languages, namely Ivatan (a Batanic language that is spoken in the northern region of the Philippines) and Cebuano (a Bisayan language in the central area of the Philippines). In Ivatan, instruments are marked by the accessory (that is, comitative and instrumental) preposition *no* for common nouns. The preposition contains the case-marking segment *n-*, which is also attested in the Ivatan pronominal prefix *ni-* (< \*n-i-), as in *ni-akən* ‘1.SG.ACCESSORY’. Furthermore, our data show that this case-marking pronominal prefix *ni-* reflects the earlier form \*kani-, which consists of the case-marking prefix \*kan- and the prefix \*i- to denote that the following root is a personal pronoun. By contrast, the Cebuano oblique prefix *kan-*, as in *kan-áku?* ‘1.SG.OBL’, while also reflecting the form \*kani-, has more functions, such as comitative, accusative, locative, dative, and benefactive functions. In addition, the presence of both the prefixal and the suffixal uses of the reflex of \*kan in the Pamona language in central Sulawesi indicates that the nominal prefix \*kan- altered to become the verbal suffix at some point in the prehistory of Malay. Based on these data, this study concludes that one of the sources of Malay *-kan* is the Proto-Malayo-Polynesian accessory prefix \*kan-, which later came to function as an oblique marker to encode various cases in some Philippine languages.

## Is Malay part of the Mainland Southeast Asian Linguistic Area? A phonological perspective

Malay is often considered to be a lect spoken at the southern end of the Mainland Southeast Asian (MSEA) linguistic area (Enfield 2021; Matisoff 2019; Nomoto and Ling Soh 2019). Sidwell and Jenny (2021), however, excludes Malay from MSEA, for it “retains much of its inherited [Austronesian] typology” (p. 3) rather than having converged into the MSEA features.

Indeed, Malay does not seem to bear most of the typical phonological features of MSEA. It is not tonal; its vowel inventory is not complex; it allows most of its phonemes at the coda position; it does not have the creaky voice register; and it is not predominantly monosyllabic or sesquisyllabic. Thus, at least at the phonological level, it seems likely that Malay cannot be considered part of the MSEA area.

Using Phonotacticon, a cross-linguistic database in progress containing basic phonotactical information, I test this hypothesis by calculating the phonological “distance” between Standard Malay and other lects of the Eurasian macroarea (as defined by Hammarström and Donohue 2014). Table 1 shows that, among the 20 lects that are phonologically the most similar to Standard Malay, only Bulu Stieng is spoken in the typically defined area of MSEA (other than Baba Malay, which is also Malayic). On the other hand, among the 20 lects that are the least similar to Malay phonologically, 14 are spoken in the MSEA. The data clearly suggests that Malay phonology is not similar to other MSEA phonologies at any meaningful level, although morphosyntactic and semantic aspects may tell a different story.

## References

- Enfield, Nick J. (2021). *The Languages of Mainland Southeast Asia*. Cambridge Language Surveys. Cambridge: Cambridge University Press. DOI: 10.1017/9781108605618.
- Hammarström, Harald and Mark Donohue (2014). “Some principles on the use of macro-areas in typological comparison”. In: *Language Dynamics and Change* 4.1, pp. 167–187.
- Matisoff, James A. (2019). “Preface”. In: *The Mainland Southeast Asia Linguistic Area*. Ed. by Alice Vittrant and Justin Watkins. Berlin; Boston: De Gruyter Mouton, pp. V–XVI. DOI: 10.1515/9783110401981-202.
- Nomoto, Hiroki and Hooi Ling Soh (2019). “Malay”. In: *The Mainland Southeast Asia Linguistic Area*. Ed. by Alice Vittrant and Justin Watkins. Berlin: Mouton, pp. 465–522.
- Sidwell, Paul and Mathias Jenny (2021). “Introduction”. In: *The Languages and Linguistics of Mainland Southeast Asia: A comprehensive guide*. Ed. by Paul Sidwell and Mathias Jenny. Berlin; Boston: De Gruyter Mouton, pp. 1–20. DOI: 10.1515/9783110558142-001.

Lect vs. Lect	Distance
Baba Malay vs. Standard Malay	0.085480188
Atong (India) vs. Standard Malay	0.095067466
Hokkaido Ainu vs. Standard Malay	0.167061519
Chaldean Neo-Aramaic vs. Standard Malay	0.168546185
Bonan vs. Standard Malay	0.199138977
Bantawa vs. Standard Malay	0.202939158
Bulo Stieng vs. Standard Malay	0.203508065
Evenki vs. Standard Malay	0.206556515
Korean vs. Standard Malay	0.214251562
Jarawa (India) vs. Standard Malay	0.215325429
Bunan vs. Standard Malay	0.216197864
Jejueo vs. Standard Malay	0.220863482
Bih vs. Standard Malay	0.22132463
Chintang vs. Standard Malay	0.231281428
Dandami Maria vs. Standard Malay	0.232917562
Asturian-Leonese-Cantabrian vs. Standard Malay	0.233773806
Kachchi vs. Standard Malay	0.242538466
Akajeru vs. Standard Malay	0.267086895
Dutch vs. Standard Malay	0.267844915
Daman-Diu Portuguese vs. Standard Malay	0.26814023
...	...
Cao Miao vs. Standard Malay	0.508648085
Standard Malay vs. Yue Chinese	0.509814795
Cosao vs. Standard Malay	0.511978446
Eastern Panjabi vs. Standard Malay	0.513931913
Japanese vs. Standard Malay	0.526323943
Daai Chin vs. Standard Malay	0.52833456
E vs. Standard Malay	0.53042274
Burmese vs. Standard Malay	0.533191225
Hakka Chinese vs. Standard Malay	0.546852544
A'ou vs. Standard Malay	0.575587103
Jiongnai Bunu vs. Standard Malay	0.583000116
Bugan vs. Standard Malay	0.596951413
Hui Chinese vs. Standard Malay	0.611087521
Angami Naga vs. Standard Malay	0.626449066
Jinyu Chinese vs. Standard Malay	0.639892157
Biyo vs. Standard Malay	0.659756729
Daohua vs. Standard Malay	0.680680035
Standard Malay vs. Western Xiangxi Miao	0.754079458
Ersu vs. Standard Malay	0.764277111
Enu vs. Standard Malay	0.791573537

Table 1: Eurasian lects phonologically similar to and dissimilar from Malay

## Abstract

# Popular culture and local lingua franca: The Emerging of Hip-hop and Rap Communities of Eastern Indonesia

by

Nazarudin ([nazarudin.hum@ui.ac.id](mailto:nazarudin.hum@ui.ac.id)) and Silva Tenrisara  
([silva.tenrisara@gmail.com](mailto:silva.tenrisara@gmail.com))

Popular culture embodies the beliefs, ideas, perspectives, attitudes, and images of various cultures. For the past decade, the Internet and social media have been a significant influence on pop culture. In Indonesia, for example, the widely spread of social media has opened new access for the youth in eastern Indonesia, such as youth community from Papua, Maluku, and NTT, to popular culture. One of it is the emerging eastern Indonesia rap and hiphop music. The urbanization in big cities in eastern Indonesia, such as Kupang, Ambon, and Jayapura encourages popular cultures among its inhabitants. They created local hiphop and rap communities and use their own local Malay mixed with Indonesian and English to deliver their musics. These emerging music communities had influences even people who live in western part of Indonesia, through social media platform such as TikTok, YouTube, and even Instagram.

The aim of this research is to investigate how these emerging eastern Indonesia youth hiphop and rap communities show and shape their identity through this popular culture. Moreover, we also investigate the code-mixing and code-switching in the local Malay that they use in their lyrics. The results of this ongoing research shows that these youth had more flexibility and creativity in code-switching and code-mixing. It means that the degree of multilingualism among these younger speakers are high. Another finding that also important to highlight is that the appearance of middle-identity among this youth communities and this identity embedded through local Malay, such as Papuan Malay, Ambonese Malay, and Kupang Malay. The term “anak Papua” not only belongs to the indigenous tribes originally from Papua, but it also applies to the descendants of the Makassarese or Buginese, or other people who already live there for a long period of time.

**Keywords:** local Malay, Papuan Malay, youth language, popular culture, multilingualism



The Ternate Malay Position in North Maluku:  
Language Attitude Studies of Hiri Island People

Filia, Nazarudin, Kushartanti

**ABSTRACT**

This paper aims to investigate the language attitude of Ternate Malay speakers on Hiri Island, North Maluku, Indonesia. In this research, I will also show the emerging phenomena of language use in the social media among the Ternate speakers. The data was taken during the fieldwork in 2021 based on questionnaire on language attitudes and observations. Respondents are people we met on the street, at the traditional market, beach, and port (86 respondents). The result shows that there is a polyglossic situation in Ternate. Indonesian language use as national language and also nationally acknowledge as the language of education. However, instead of the formal Indonesian, the Ternate speakers often use the Ternate Malay than Indonesian. Nevertheless, the growth in the use of Ternate Malay could be an increasing threat to the long-term viability of local languages in the region. Meanwhile, in the social media, the Ternate speakers tend to use the Ternate Malay side to side with the Ternate language.

**Keywords:** language attitude, Ternate Malay, sociolinguistic, North Maluku

## **The Myth of Malay Decline in the Malay World**

*Najib Noorashid*

*Universiti Brunei Darussalam, Brunei Darussalam*

The interminable ‘Malay vs. English’ debate is ubiquitous in Brunei Darussalam, Malaysia, Indonesia and Singapore – the Malay World. Although Malay is the national language, and a cultural and identity marker for the Malays, the hegemony of English has prompted concerns on the decline of the Malay language among scholars and the general public. By incorporating multi-modal approaches and sociolinguistic perspectives of language use and attitudes, this study uses empirical data to compare the sociolinguistic situations in the four Malay countries. This study found that there are misconceptions concerning the decline of Malay due to the misunderstanding of language contact issues with English, and the legacy left from the Western colonial period. This study further discusses that the belief of English has a negative effect on Malay is another language myth in the Malay World. This is predominantly due to the propensity of promoting a zero-sum game between the use of English and Malay, consequently raising the alarming trend of recycling colonial knowledge about Malay and the Malay World as postulated by Collins (2002). The sociolinguistic situation is further exacerbated by the politicization of sociolinguistic issues, while other themes including purism, religion and security also play major roles in influencing Malay attitudes towards bilingualism and Malay vitality. This study therefore attempts to address the possibility of ‘Malay endangerment’ and its declining vitality resulting from English hegemony as another exaggerated claim raised by certain people with vested interests.

**Keywords:** Malay, English, Brunei, Malay World, language contact, language myth, sociolinguistics.

## *Can the Indonesian census inform us about vitality? Analyzing Malayic speaker data from West Kalimantan*

Karl Anderbeck, Nat Wong, and Christina Natasha

### Abstract

People across the world are shifting from smaller, local languages to larger languages (Krauss 1992; Quakenbush & Simons 2015). Indonesia's 700+ languages, many of them with small speaker populations, are not exempt from this trend (Anderbeck 2015; Abtahian, Cohn & Pepinsky 2016a).

Given Indonesia's large and geographically dispersed population, high linguistic diversity, and low numbers of researchers, statistical methods utilizing data sources like the national census are an attractive way to get a broad understanding of language dynamics in Indonesia (Abtahian, Cohn & Pepinsky 2016b; Abdullah, Yunita & Cleopatra 2014).

This paper details the development of a pragmatic model for gauging present and future intergenerational transmission of heritage languages using data from the 2010 Indonesian census. The model includes calculations of speaker population, comparison of old and current speaker populations, comparison of ethnic and speaker populations, and comparison of old and young ethnics. This model is then applied among the (mostly Malayic) languages of West Kalimantan and compared to field observations of language transmission.

### References

- Abdullah, Suparman Ibrahim, Yunita & Maria Cleopatra. 2014. Distribution of daily use local language in Indonesia. Paper presented at the Second International Conference on Education and Language (2nd ICEL), Bandar Lampung University.
- Abtahian, Maya, Abigail Cohn & Thomas Pepinsky. 2016a. Modeling social factors in language shift. *International Journal of the Sociology of Language* 242. 139–170.
- Abtahian, Maya, Abigail Cohn & Thomas Pepinsky. 2016b. Methods for modeling social factors in language shift. *University of Pennsylvania Working Papers in Linguistics* 22(2 Selected Papers from New Ways of Analyzing Variation (NWAV 44)). 1–10.
- Anderbeck, Karl. 2015. Portraits of language vitality in the languages of Indonesia. In I Wayan Arka, Ni Luh Nyoman Seri Malini & Ida Ayu Made Puspani (eds.), *Language documentation and cultural practices in the Austronesian world* (A-PL 019: Papers from 12-ICAL), vol. 4, 19–47. Canberra: Pacific Linguistics.
- Krauss, Michael. 1992. The world's languages in crisis. *Language* 68(1). 4–10.
- Quakenbush, J. Stephen & Gary Simons. 2015. Looking at Austronesian language vitality and endangerment through EGIDS and the sustainable use model. In I Wayan Arka, Ni Luh Nyoman Seri Malini & Ida Ayu Made Puspani (eds.), *Language documentation and cultural practices in the Austronesian world* (A-PL 019: Papers from 12-ICAL), vol. 4, 1–17. Canberra: Pacific Linguistics.



## Direct speech in Indonesian: Functional change of non-verb frames

Asako Shiohara

This paper deals with a stylistic change from classical Malay (CM) to modern Indonesian (MI) related to the strategies for framing direct speech (DS, hereafter) observed in narrative.

In CM, as represented by Hikayat, a type of prose typically delivering legends or histories, DS is most frequently framed by a non-verb constituent followed by the phrase denoting the speaker. In excerpt (1), cited from Hikayat Raja Pasai, the noun *kata* ‘words’ frames DS followed by the speaker NP *orang dalam jong itu* ‘the person in the ship’ in the first and the third clause, and the verb stem *sahut* frames DS in the second clause followed by the speaker NP *orang mengali itu* ‘person who dug that’.

- (1) *Maka kata orang dalam jong itu, “Apa ada khabar dalam negeri ini?” Maka sahut orang mengali itu, “Adapun khabar dalam negeri ini badak makan anaknya.” Maka kata orang dalam jong itu, “Tiada kami tahu apa ertinya katamu itu.”* (Mead 1914:45)

‘So, the person in the ship **said**, “Is there any news in this country?” Then the man who dug **answered**, “As for the news in this country, rhinos eat their young.” Then the person in the ship **said**, “We don’t know what you mean by that word.”’ (119: 18, in Jones 1987: 63).

Besides the noun *kata* and the verbal root *sahut* found in excerpt (1), some other forms such as *ujar* and *jawab* may perform similar functions, which can be opposed to verb frames with verbs such as *berkata*, *menyahut*, *berujar*, and *berjawab*. We shall label such non-verb frames “NVF” henceforth.

Cumming (1983: 67) mentions the stylistic change from CM to MI; MI employs NVF, but often after DS. Excerpt (2) exemplifies this pattern, which is cited from an article in the newspaper KOMPAS in 2018.

- (2) *“Mereka bersyukur karena sudah bisa lolos, bisa ikut pemilu. Intinya itu,” kata Jokowi.*

“They are grateful because they have been able to pass, they can participate in the election. The point is that,” said Jokowi. (https://nasional.kompas.com/read/2018/03/02/14102441/kata-presiden-soal-pertemuan-90-menit-dengan-psi-di-istana?page=all.)

After addressing the question of when and how the change occurred based on literature of the late 1800s and early 1900s—a period of several years before and after 1900 seems to be the turning point—I will examine the factors that caused the change. I posit that, aside from the most direct factor of the influence of foreign novels, the change from oral to written discourse might have affected the style. In Hikayat, which had developed as oral literature (Sweeney 1987), the restriction of linearity requires the frame to precede the DS; otherwise, interpretational confusion may arise. Contrariwise, in modern novels, which have developed as written text, visual information such as quotation marks helps the readers to identify the marked section as DS.

### References

- Cumming, Susanna. 1991. Functional change: The case of Malay constituent order. The Hague: Mouton de Gruyter.  
Mead, J. P. 1914. A Romanized Version of the Hikayat Raja-Raja Pasai. Journal of the Straits Branch of the Royal Asiatic Society 66. URL: <https://www.jstor.org/stable/41561000>  
Sweeney, Amin. 1987. *A full hearing : orality and literacy in the Malay world*. Berkeley : University of California Press.

## Indonesian classifiers are required by syntax, not semantics

In this paper, I investigate the syntactic behavior of classifiers in Indonesian. The main theoretical point made in my work is that the syntactic distribution of Indonesian classifiers can (and should) be explained mainly by their syntactic properties, and not by their semantics. In the existing literature on classifier languages (see Little et al. 2020 for an overview), the syntactic facts are accounted for by a (formal) semantic analysis; I show that, for Indonesian, a syntactic analysis is more motivated than a semantic one.

I claim that, unlike in other classifier languages (cf. Bale & Coon 2014), Indonesian classifiers do not have any number-related semantics (contra Little & Winarto 2018, Nomoto 2013). From a semantic point of view, they only introduce a presupposition that a given nominal referent belongs to a certain nominal class (e.g. humans or animals). For this reason, using a wrong classifier renders the whole expression illicit, as shown in the pair of examples in (1), (2); in (2), the classifier for human referents occurs in a NP denoting an animal (a cat), which is impossible.

(1) dua (ekor) kucing  
two CLF.ANIMAL cat  
'two cats'

(2) \*dua orang kucing  
two CLF.HUMAN cat  
int. meaning: two cats

The facts that are to be accounted for by a syntactic theory of Indonesian classifiers are the following. First, as shown in (3), classifiers cannot occur without numerals (or other quantifiers) – in such a case, they are interpreted as “homonymous” lexical nouns.

(3) a. ekor kucing  
CLF/tail cat  
1. <sup>ok</sup> 'a cat's tail'  
2. \* 'a cat / one cat'  
(own data)

b. buah mannga  
CLF/fruit mango  
1. <sup>ok</sup> 'mango fruit' lit. 'fruit of mango'  
2. \* 'a mango / mango'  
(Little & Winarto 2018: sect 2.1)

Secondly, classifiers and nouns are not required to be linearly adjacent, as shown in (4a), (4b):

(4) a. **Postnominal numerals**  
saya me-lihat kucing dua (ekor)  
I ACT-see cat two CLF  
'I see two cats.'

b. **'floated' numerals**  
udang purba itu tinggal dua (ekor)  
shrimp ancient this remain two CLF  
'Only 2 (of those) ancient shrimps remain.'

Thirdly, classifiers are optional in any syntactic context, as shown in (1-4).

In my analysis, I claim that Indonesian classifiers form a phrase with a numeral rather than the noun (as has been claimed in Little & Winarto 2018, Nomoto 2013). I suggest that classifiers carry an unvalued selectional feature, uNum, which requires them, once they enter the syntactic workspace, to combine directly with the numeral (i.e. adjuncts select for the phrases they merge with; cf. Zeijlstra 2020). Numerals carry an unvalued selectional feature uN, which percolates up to the NumP; for this reason, NumP is required to combine with N, being a N' level adjunct.

## References

- Bale, A., & Coon, J. (2014). Classifiers are for numerals, not for nouns: Consequences for the mass/count distinction. *Linguistic Inquiry*, 45(4). Little, C. R., & Winarto, E. (2018). Classifiers and the definite article in Indonesian. In Baird, M. & J. Pesetsky (eds), 49th meeting of the North East Linguistics Society (NELS). GLSA. Little, C. R., Moroney, M., & Royer, J. (2020, January). Classifying classifiers: Two kinds of numeral classifiers across languages. In 94th Annual Meeting of the Linguistic Society of America. New Orleans, LA. Nomoto, H. (2013). Number in classifier languages. University of Minnesota. Zeijlstra, H. (2020). Labeling, selection, and feature checking. *Agree to Agree*, 31.

## Pronoun substitution in Javanese, Indonesian, and Malay

Yoshimi Miyake (Akita University), Hiroki Nomoto (Tokyo University of Foreign Studies)

Sri Budi Lestari (Ritsumeikan Asia Pacific University)

The phenomenon of pronoun substitutes as well as pronoun ellipses in Asian languages have been studied from sociolinguistic points of views and grammatical feature point of views (Ewing 2014, Lestari 2020, Utsumi 2020). This study will discuss pronoun substitution in three languages, i.e., Javanese, Indonesian, and Malaysian. Sixteen different categories of nouns have been investigated to see whether each of the nouns can be used to substitute pronouns. The sixteen categories are: age differences, kinship and family terms, fictional kinship terms, occupations related to education, occupations in service industry, or occupation related to religion such as priest, personal relations, social status, royal status, as well as 6 grammatical categories, i.e. quantifiers, demonstrators, personal pronouns, proper names, modifier + noun, and others. Having examined whether each of the terms can be used as substitutes for 1) 1<sup>st</sup> person pronoun, 2) 2<sup>nd</sup> person pronoun, 3) address term, and 4) title, our current studies show that terms for family or kinship are the most frequently observed nouns which can substitute personal pronouns.

Discussing Malay, Nomoto reports that family and kinship terms include the most nouns which can be used both as address terms and second person pronoun substitutes, and nouns which can be used not only as address terms & second person pronouns, but also as the first person pronouns (Nomoto 2022).

Although Indonesian and Javanese show a similar trend, that is, family and kinship terms are most frequently used as pronoun substitutes for the second person, address terms, as well as title, they show a distinctive difference from Malaysian family and kinship terms in fictional uses, that is, family or kinship terms are used for non-family or kin members. Malay terms for mother and father, i.e. *bapa*, *ayah*, *ibu emak*, and a term for child, that is, *anak/nak* cannot be fictionally used neither as address terms, pronoun substitutes or titles—those terms are used only for ‘actual’ family members. On the other hand, an Indonesian term for father *bapak* (not a Malay origin *ayah*) and *ibu* ‘mother’ can be fictionally used as pronoun substitutes and titles, or rather, those terms are the most frequently used pronoun substitutes as well as address terms for adults.

Terms for older siblings, i.e. *abang*, *kakak/kak* in both Malay and Indonesian can be fictionally used. In Javanese, *mas* ‘older brother’ and *mbak* ‘older sister’ are the most commonly used pronoun substitutes for young men and women.

We will discuss why a strong correlation is observed between address terms and second person pronouns. We will also discuss what kinds of factors block certain nouns from substituting pronouns such as shortened forms, newly coined terms, politeness, etc.

## A comparison of speech act participant referring expressions in Malay and Japanese

**Four functions.** Malay utilizes non-pronominal expressions such as kinship terms (e.g., *ibu* ‘mother’) and professions (e.g., *cikgu* ‘teacher’) when referring to interlocutors in contexts where personal pronouns (e.g., *I, you*) are expected in languages such as English. Such non-pronominal expressions are referred to as **pronoun substitutes** (McGinn 1991). A pronoun substitute can refer to the speaker, as in (1), or the addressee, as in (2).

- (1) *Ibu tak sedar yang kamu dah balik.* ‘**Mother (= I)** didn’t realize you were back.’
- (2) *Esok pagi saya akan hantar tugasan itu ke bilik prof.* ‘Tomorrow morning I will turn in the assignment to **professor’s (= your)** room.’

Expressions of pronoun substitutes are often used as a **title** as in *Cikgu Abu* ‘Mr. Abu’ and/or as an **address term** as in *Selamat pagi, cikgu!* ‘Good morning, Sir/Miss!’ In this paper, we compare the distributions of pronoun substitutes in Malay and Japanese and point out similarities and differences between the two languages.

**Data collection method.** The data was collected using a common questionnaire for investigating pronoun substitutes and address terms (Okano et al. 2022). It contains 147 concepts that can potentially function as pronoun substitutes and address terms cross-linguistically. The questionnaire classifies them into 16 semantic categories such as Age, Family, Position, Anaphor, Quantifier, etc. First, we chose one or more than one expression for each concept. Then, we examined whether a given expression has either of the four functions (i.e., 1st and 2nd person pronoun substitutes, address term, title).

**Results.** A total of 171 and 236 expression tokens were chosen for Malay and Japanese, respectively. The function distribution was expressed by a four-digit number as illustrated in the table below, where 1 means a given function is available and 0 means it is not.

Category	ID	Concept	Expression	1st	2nd	Address term	Title	Distribution
Family	8	father	ayah	1	1	1	0	[1110]
Fictional_family	35	father	pak	0	0	0	1	[0001]

**Findings.** [1] Prominent distributions. The most prominent distributional patterns in Malay included 2nd person and address terms ([0110], [1110], [0111]) and few expressions allow all four functions ([1111]). By contrast, [1111] was the most frequent pattern in Japanese. This difference can be ascribed to different natures of kinship terms in the two languages. [2] Family and Fictional Family. In Malay, Fictional family items behaved independently from their Family counterparts. Thus, *mak* ‘mum’ showed [1110] for Family but [0001] for Fictional family. By contrast, most expressions in these categories behave identically in Japanese. [3] Unattested distributions. Patterns [1001], [1101] and [1011] were attested in neither language.

**References.** McGinn, R. 1991. Pronouns, politeness, and hierarchy in Malay. In R. Blust (ed.) *Currents in Pacific linguistics: Festschrift in honor of George W. Grace*, 197–221. Canberra: Australian National University. • Okano, K., H. Nomoto, S. Wittayapanyanon, Thuzar Hlaing and A. Kasuga. 2022. An investigation of pronoun substitutes and address terms in three Asian languages based on a common questionnaire. *Proceedings of the Twenty-Sixth Annual Meeting of the Association for Natural Language Processing*, 69–73.



# ISMIL 25 (6-8 August 2022), ILCAA/Tokyo University of Foreign Studies

Timothy C Brickell: Tokyo University of Foreign Studies / The University of  
Melbourne

## A comparison of sound symbolism in Indonesian and two languages of North Sulawesi

Lexical items denoting sound symbolism are often labelled synonymously as ‘onomatopoeia’, ‘ideophones’, ‘iconicity’, ‘mimetics’, or ‘expressives’ (Lee 2017:181). A broad definition is that all refer to words which display a more direct association between form and meaning, and which describe a range of properties and sounds in the external world, for example those made by natural elements, animals, humans, tools, or machines. This paper seeks to address, at least in part, the paucity of description of this topic in Austronesian (AN) languages by comparing sound symbolism in varieties of Indonesian with what can be observed for two Minahasan languages indigenous to North Sulawesi: Tondano and Tombulu.

Sound symbolic words represent concepts which can be considered universal to human experience. Descriptions and definitions in extensively researched languages are therefore not uncommon, with previous literature seeking to examine these words from a cross-linguistic perspective, for instance Voeltz & Kilian-Hatz (2001) and Hinton, Nichols, & Ohala (1994). However, descriptions of sound symbolism in the AN languages are notably lacking, as outlined in Lee’s (2017) article on this aspect of Seediq (Atayalic) grammar. Grammatical descriptions and dictionaries of AN languages (including Indonesian) often omit or give limited space to sound symbolism. Publications which at least mention this phenomenon in AN languages are those for: Kambera (Klamer 1998, 1999), Balinese (Clynes 1995, 1998, 2000), and Malay (Ade-laar 1983). More detailed descriptions are rare; in addition to Seediq (Lee 2017), exceptions are found for Indonesian (Mandjusri 1987), Tagalog (Rau 1994), Ilocano (Rubino 2001), and Numbami & Jabem (Bradshaw 2006).

For this study, the approach will be to examine Indonesian lexical items from a variety of sources, from the more formal *bahasa baku* in dictionaries to colloquial varieties in blogs and/or social media. Data for Tondano and Tombulu come from archived recordings, a dictionary, and discussions with native speakers. Lexical items will be categorised as: Natural sounds (the four elements, animals, and humans) or Artifacts (musical instruments, vehicles, and mechanical and electronic equipment). Sound symbolism in the two categories of languages will be examined with regard to any shared characteristics of phonology and morphology. Despite hundreds of years of close contact between Minahasan languages and other Malay varieties, linguistic transfer has been minimal (Brickell 2020). Essentially no morphosyntax is shared, although a small amount of lexical transfer has occurred. The primary questions to be answered are to what extent do Indonesian and the Minahasan languages share sound symbolic words or their morphophonological features, and are these the result of linguistic transfer.

# Recent Contact Features in Northwest Papuan Malay

David Gil

The Malay/Indonesian language bears testament to thousands of years of language contact, with contact features dating back to the original expansion of Austronesian languages into Nusantara, other features associated with the early diversification of the Malayic sub-branch, and yet others reflecting ongoing contact between contemporary varieties of Malay/Indonesian and the surrounding languages.

This paper presents a case study of the latter state of affairs, focusing on Papuan Malay, specifically the northwest subdialect spoken in the Cenderawasih Bay and Bird's Head regions. The paper presents five grammatical features that distinguish between Northwest Papuan Malay and other closely related Eastern varieties of Malay, and which may accordingly be attributed to recent or even ongoing contact between Northwest Papuan Malay and neighboring languages, either Austronesian (eg. Biak, Wamesa) or non-Austronesian (e.g. Meyah, Hatam).

## 1. *Distribution of nasal consonants*

Whereas most eastern Malay varieties replace final [n] with [ŋ], Northwest Papuan Malay tends to replace [ŋ] with [n], resulting in a complex phrasally-conditioned pattern of variation between the two consonants.

## 2. *Dual number in pronouns*

In Northwest Papuan Malay, the periphrastic expressions *tong dua*, *kam dua* and *dong dua* are in the process of undergoing lexicalization, resulting in a set of dual pronouns, *tendua*, *kamdua*, *dendua*.

## 3. *Bilateral instrumental verbal marking*

Bilateral instrumental verbal marking indexes the presence of an instrumental argument that, in addition, bears a different thematic role with respect to a second (overt or understood) verb. In Northwest Papuan Malay, such marking occurs periphrastically, with *pake* as the first term in a serial verb construction.

## 4. *Generalized argument indexation*

In Northwest Papuan Malay, nouns are commonly marked for person and number by means of an enclitic pronoun.

## 5. *Null content interrogatives*

In Northwest Papuan Malay, content questions are commonly formed without any overt interrogative word, the position occupied by the expected interrogative remaining empty.

Each of these five features of Northwest Papuan Malay is rare or absent from other varieties of eastern Malay, but common in other neighboring languages; moreover, these features range from uncommon to exceedingly rare worldwide. Thus, it is argued that each of these five features reflects the outcome of recent or ongoing contact between Northwest Papuan Malay and other languages of the region.

## **Aspiration in Baling Malay Dialect**

Mohammad Khairulanwar Abd Ghani & Mohd Tarmizi Hasrah

[ujanasastera.k3@gmail.com](mailto:ujanasastera.k3@gmail.com); [h.tarmizi@uum.edu.my](mailto:h.tarmizi@uum.edu.my)

Pusat Pengajian Bahasa, Tamadun dan Falsafah  
Universiti Utara Malaysia

### **Abstract**

In the classification of Aslian languages, phonemic aspiration of stops is considered to be characteristic only of members of the southern branch of the Aslian group (Geoffrey Benjamin, 2012). In this paper, we try to provide all available information on the presence of aspirated voiceless stops in Baling Malay dialect. In addition to the presentation of relevant lexical data, we also discuss the phonemic status, and the origin of aspiration. This research employs a qualitative study and data was scrutinized using Descriptive Linguistic method. Area of data gathering includes Malay villages along the Baling district in Kedah. Cross-linguistic comparison shows that the rise of phonemic aspiration in Baling Malay dialect is an areal phenomenon, finding close parallels in other languages in the northern Malay Peninsula, as a result of increasing contact with Thai and Patani Malay dialect. In other words, circumstances such as boundaries, politics, and mass migration of the Thai community to Kedah between the 16<sup>th</sup> and 17<sup>th</sup> centuries, have indirectly resulted to phenomenon what we call as linguistic contact (Asmah, 1985).

**Keywords:** Aspiration, Baling Malay, linguistic typology, Patani Malay, phonemes

## Evidence of contact with Malay/Indonesian in the Enggano language

This paper presents evidence of contact with Malay/Indonesian in the Enggano language, spoken by c. 1,500 speakers on Enggano Island, Sumatra. Through analysis of the lexicon, phonology and morphosyntax, we demonstrate that increased contact with Indonesian in the period after Independence has led to a greater degree of lexical borrowings as well as triggering language change. This can be shown by comparing a corpus collected by Hans Kahler in the 1930s with contemporary materials collected as part of ongoing documentation since 2018.

In the Kahler materials, there is already evidence of contact with Malay in the form of lexical borrowings, e.g. *dupia* ‘money’ < rupiah, *baw* ‘onion’ < bawang, *bidi’i* ‘room’ < bilik (Kahler 1987). These undergo assimilation: final codas are lost, nasalisation spreads from a nasal consonant before deletion, non-phonemes are replaced with the nearest equivalent. However, these loans are not regularly used in the text collection (e.g. Kahler 1955). In comparison, lexical borrowings are common in contemporary Enggano: 243 out of 1065 words in the lexical database are loans from Malay/Indonesian, and only 15 of these were attested in the Kahler dictionary. Moreover, 102 of 279 clauses in a small text corpus contained one or more loans, suggesting they are relatively discourse-frequent.

Many borrowings are thought to come from Bengkulu Malay or Minangkabau, since they reflect the change \*-a > -o that distinguishes these varieties (Nothofer 1992: 23), e.g. *apo* ‘what’, *tuo* ‘old’. Like older loans, new borrowings often show assimilation, e.g. through simplification of consonant clusters (e.g. *biku* ‘week’ < minggu, *kuti* ‘key’ < kunci), loss of velar nasals (e.g. *ik* ‘corn’ < jagung, *tk* ‘bear’ < tanggung), replacement of non-phonemes (e.g. *tekora* ‘school’ < sekolah, *kuru* ‘teacher’ < guru), word-level nasality (e.g. *nanau* ‘lake’ < danau) and use with Enggano morphology (e.g. *ka’kari* ‘work’ (ki-a’-kari ‘FOC-VBLZ-work’ < kerja). However, there are also instances of Indonesian forms that appear in the texts in unassimilated form, including discourse markers like *jadi* ‘so’, and larger phrases like *dak tau nya* ‘before we knew it’. These may represent code-switching with Indonesian rather than borrowing and are perhaps reflective of a change from Enggano-dominance to Indonesian-dominance among the speech community. This in turn suggests a change in contact situation from a typical borrowing context to shift-induced interference (see Haspelmath 2008).

As well as an increase in lexical borrowing, contact with Malay/Indonesian has also resulted in language change. In the Kahler materials, loans beginning with [s] and [t] are systematically borrowed with initial [k], e.g. *e-kapii* ‘cow’ < sapi, *e-kumu* ‘well’ < sumur, reflecting the fact that [t] only exists as an allophone of /d/, whilst PMP \*t and \*s merged as Enggano *k* (Edwards 2015). As seen above, many contemporary borrowings replace [s] with [t], suggesting that *t* may now be viewed as a phoneme of Enggano. On a morpho-syntactic level, contemporary Enggano appears to have undergone several changes. For example, main clauses in the Kahler texts have predicates in *bu*-form, typically with verb-initial order. In contemporary Enggano texts, over half of *bu*-verbs occur in subject-initial. This may reflect contact with predominantly SVO Indonesian.

Consequently, comparing the Kahler corpus with the contemporary Enggano corpus provides ample evidence of increased contact with Malay/Indonesian through an increase in lexical borrowing/code-switching, as well as contact-induced change. The aim of the paper is to illustrate these changes and reflect on the implications for our understanding of language contact in minority/endangered language contexts, where the relative status of the donor and recipient language has changed.

## References:

- Edwards, Owen. 2015. The Position of Enggano within Austronesian. *Oceanic Linguistics* 54(1): 54–109.
- Haspelmath, Martin. 2008. Loanword typology: Steps toward a systematic cross-linguistic study of lexical borrowability. In T. Stolz, D. Bakker and R. Salas Palomo (eds.) *Aspects of Language Contact: New Theoretical, Methodological and Empirical Findings with Special Focus on Romancisation Processes*, 43-62. Berlin, New York: De Gruyter Mouton.
- Kähler, Hans. 1955. Ein Text von der Insel Enggano. *Afrika und Asien* 39 (1955): 89--94
- Kähler, Hans. 1987. Enggano–deutsches Wörterbuch. Berlin: Dietrich Reimer Verlag.
- Nothofer, B. 1992. Lehnwörter im Enggano. In F. Schultze & K. Tauchmann (eds.) *Kölner Beiträge aus Malaiologie und Ethnologie zu Ehren von Professor Dr. Irene Hilgers-Hesse*, 21-31. Bonn: Hols.