

# Pronoun substitutes in Malay textbooks

Hiroki Nomoto

Tokyo University of Foreign Studies

3/12/2021, CLS's 20th Anniversary e-Symposium, NUS

# Pronoun substitutes

- **Pronoun substitutes**

- Non-pronominal expressions used to refer to the speaker and the addressee.
- Can be substituted by 1st and 2nd person personal pronouns.
- E.g., MOTHER is used where English employs *I/you*.
- Common in East and Southeast Asian languages.

# Examples from TALPCo

- TUFS Asian Language Parallel Corpus (Nomoto et al. 2018, 2019)  
<https://github.com/matbahasa/TALPCo>
- [Spoken to Mr. Lee] 'Which drink would **you** like?' (#3756)
  - Japanese: **イさん**はどちらの飲み物がいいですか。
  - Korean: **이 씨**는 어느 쪽 음료수가 좋습니까?
  - Malay: **Encik Lee** nak minuman yang mana?
  - Indonesian: **Pak Lee** mau minuman yang mana?
  - Thai: **คุณอิ**รับเครื่องดื่มอันไหนดีคะ
  - Vietnamese: **Anh Lee** thích loại đồ uống nào?
  - Burmese: **မစ္ဆူတာလီ** ဘယ်သောက်စရာ သောက်မလဲ။

# Pronoun substitutes ≠ address terms

- **Address terms**

- Identify who the intended addressee is.
- Outside the proposition expressed by the sentence.
- Not limited to East and Southeast Asian languages.

- Example (Malay, TALPCo #2744)

Cikgu, ini ibu saya.

teacher this mother 1SG

'Miss, this is my mother.'

# Ways of referring to speaker & addressee

- Indo-European & Chinese languages Novice
  - A limited number of personal pronouns
  - English: I, my, me, myself, you, your, yourself, yourselves, we, our, us, ourselves (12 items)
  - Mandarin: 我, 我們, 咱們, 你, 您, 你們 (6 items)
- East and Southeast Asian languages Novice - Superior
  - Many personal pronouns
  - Many pronoun substitutes
  - Null pronouns ("omission")
  - How to choose from many available options

# Mistake by a student

- Final exam of Malay Grammar
- Bonus question: Write a message of gratitude to Mrs. Faridah.
- Student: Studied Malay for a year (5 classes a week)
- Answer

Pn. Faridah, terimah kasih selama setahun. Kelas <sup>puan 'Mrs.'</sup> awak sangat menarik, jadi saya boleh belajar bahasa Melayu dengan seronok.

'Mrs. Faridah, thank you for a year. **Your** class was very interesting, so I could enjoy studying Malay.'

2nd person pronoun for people of equal or lower status

# Issues

1. How have pronoun substitutes been taught?  
→ Focus on Malay textbooks
2. How should we teach pronoun substitutes and other speaker- and addressee-referring expressions?

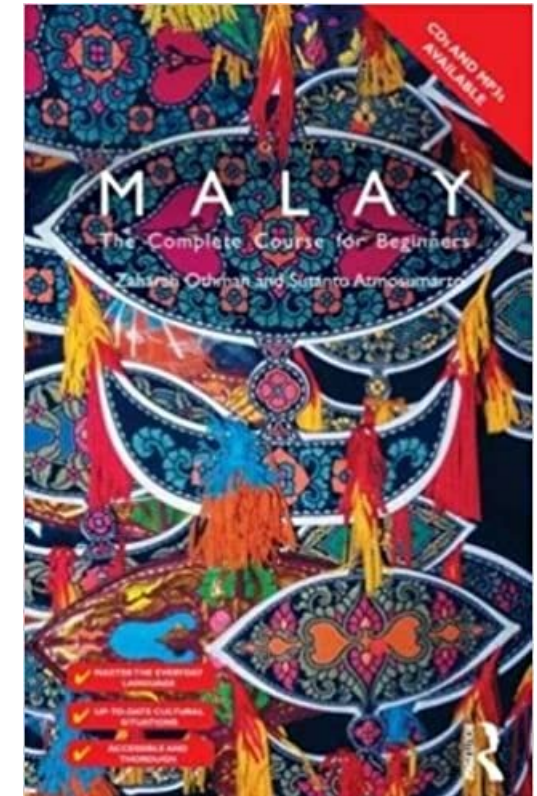
# Textbooks to be examined

1. Zaharah & Atmosmarto. 1995. *Colloquial Malay* (1st edn.)
2. Shaik Omar & Yamazaki. 1997. *Omaa-Azuu no Mareego Kouza [Omar & Azu's Malay Course]*
3. Faridah & Kondo. 1999. *Ekusupuresu Mareego [Express Malay]*
4. Zaharah. 2012. *Colloquial Malay* (2nd edn.)
5. Faridah & Kondo. 2019. *Nyuu Ekusupuresu Purasu Mareego [New Express Plus Malay]*



# [1] Zaharah & Atmosmarto. 1995. *Colloquial Malay* (1st edn.)

- Lesson 1
  - Formal personal pronouns: *saya* 'I', *anda* 'you'
  - These two pronouns are used throughout the book as the default 1st & 2nd person terms.
- Lesson 2 **Proper name**
  - "Saying 'you' and 'yours' is almost avoided"
  - **Annie** tak ada anak lelaki? 'Annie, do **you** have a son?'
- Lesson 3
  - Informal personal pronouns: *aku/ku* 'I', *kamu/mu* 'you', *engkau/kau* 'you'
- Lesson 14 **Common noun (occupation)**
  - Pada pendapat **Tuan Doktor**, penyakit ini boleh menjangkiti orang lain? 'Doctor, in **your** opinion, can this disease be communicated to others?'
- Lesson 16 **Common noun (kinship term)**
  - **Kakak** mesti telefon pejabat untuk memberitahu mereka **kakak** terpaksa pulang. 'Sister, **you** must call the office to tell them **you** have to go home.'



## [2] Shaik Omar & Yamazaki. 1997. *Omaa-AzUU no Mareego Kouza [Omar & Azu's Malay Course]*

- Lesson 1 **Title**
  - Formal personal pronouns: *saya* 'I', *anda* 'you'
  - These two pronouns are used throughout the book as the default 1st & 2nd person terms.
  - **Cik** 'Miss', **puan** 'Mrs.', **encik** 'Mr.'
- Lesson 3 **Common noun (kinship term, obedience)**
  - Informal personal pronouns: *aku* 'I', *awak* 'you', *kamu* 'you', *engkau* 'you'
  - **Saudara** 'relative, brother', **saudari** 'relative, sister', **pak cik** 'uncle', **mak cik** 'aunt', **abang** 'elder brother', **kakak** 'elder sister', **adik** 'younger sibling'
  - **Tuan** 'master'
  - "When foreigners speak Malay, it is safe to use *pak cik/mak cik*, *abang*, *kakak* and *adik* even for familiar people."



# [3] Faridah & Kondo. 1999. *Ekusupuresu Mareego* [Express Malay]

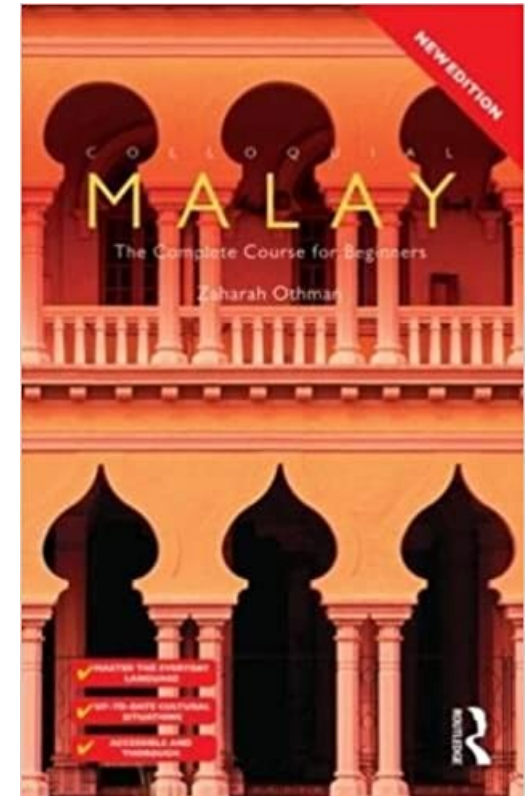
- Lesson 1
  - Formal and informal personal pronouns: *saya* 'I', *aku* 'I', *awak* 'you', *kamu* 'you', *engkau* 'you'
  - The default 1st & 2nd person terms are *saya* 'I' and 'Title + Proper name'.
  - *Anda* 'you' is never used.
  - Pronoun substitutes (presented as 2nd person personal pronouns): **encik** 'Mr.', **cik** 'Miss', **tuan** 'master', **puan** 'Mrs.'
- Lesson 2 **Title** **encik** 'Mr.'
- Lesson 3 **Title + Proper name** **Encik Sato** 'Mr. Sato'
- Lesson 10 **Common noun (kinship term)** **abang** 'brother'
- Lesson 12 **Common noun (occupation)** **cikgu** 'teacher'
- Lesson 13
  - 2nd person pronoun (*awak*) is actually used for the first time!



# [4] Zaharah. 2012. *Colloquial Malay* (2nd edn.)

- Unit 1

- Formal and informal personal pronouns: *saya* 'I', *aku* 'I', *anda* 'you', *kamu* 'you', *awak* 'you', *engkau* 'you'
- The default 1st & 2nd person terms are *saya* and *awak/kamu* (not *anda*!).
- Special emphasis on pronoun substitutes
- Maria: Mak, **Maria** nak makan. 'Mum, Maria/I want to eat.'  
Mother: **Maria** nak makan apa? **Mak** masak kari.  
'What do Maria/you want to eat? Mum/I will cook curry.'
- Exercise 4:  
How would each person address the other?
  1. a child speaking to an elderly woman at a fruit stall
  2. a woman speaking to a gentleman at a post-office counter
  3. ...



[5] Faridah & Kondo.

2019. *Nyuu Ekusupuresu Purasu Mareego* [New Express Plus Malay]

- Lesson 1

- Formal and informal personal pronouns: *saya* 'I', *anda* 'you', *awak* 'you' *kamu* 'you'
- The default 1st & 2nd person terms are *saya* 'I' and 'Title (+ Proper name)'
- Pronoun substitutes (presented as 2nd person personal pronouns): **encik** 'Mr.', **cik** 'Miss', **tuan** 'master', **puan** 'Mrs.', **cik puan** 'Mrs.', **saudara** 'relative, brother', **saudari** 'relative, sister', **pak cik** 'uncle', **mak cik** 'aunt', **abang** 'elder brother', **kakak** 'elder sister', **adik** 'younger sibling'



# How should we teach pronoun substitutes?

- Spread over multiple lessons
  - Too many to teach/learn all in Lesson 1
  - But how? Which ones to teach first?
- Limitation/bias of classroom situation
  - The interpersonal relationship is fixed to ‘teacher-student’
  - Need to provide conversation examples in many different settings
  - Which settings should be given priority?
    - University (preparation for study abroad) *Dr., prof., encik ‘Mr.’*, etc.
    - Shopping (for travel) *abang ‘brother’, mak cik ‘aunt’*, etc.
    - Business/Formal *tuan ‘master’, bos ‘boss’, YB (title)*, etc.

# What should we consult when teaching pronouns substitutes? --- Linguistic study

- “A Cross-Linguistic Study of Pronoun Substitutes and Address Terms” Project (JSPS KAKENHI Grant #20H01255)
- Database of expressions that can be used as pronoun substitutes, address terms and titles in 8 languages
  - Japanese, Korean, Malay, Indonesian, Javanese, Thai, Vietnamese, Burmese
  - Common questionnaire for investigating pronoun substitutes and address terms
- Annotate corpora with pronoun substitutes and address terms
- Nomoto et al. (2021)

# Our team



- Sunisa Wittayapanyanon (Saito) (TUFS) **Thai**
- Hiroki Nomoto (TUFS) **Malay, NLP**
- Virach Sornlertlamvanich (Musashino/Thammasat) **Thai, NLP**
- Ryuko Taniguchi (TUFS) **Japanese**
- Yunjin Nam (TUFS) **Korean**
- Yoshimi Miyake (Akita) **Indonesian, Javanese**
- Sri Budi Lestari (Ritsumeikan APU) **Indonesian, Javanese**
- Atsushi Kasuga (Kanda International U) **Vietnamese**
- Kenji Okano (TUFS) **Burmese**
- Thuzar Hlaing (TUFS) **Burmese**





# Database

## Word list of pronoun substitutes and address terms

[Export JSON](#)[Export CSV](#)

Word	Language	Creator	Editor	Action
Encik Isa	Malaysian	KOSEKI Ayumi 小関 歩 7419103	Poey BKT	<a href="#">Edit</a>
Encik Abe	Malaysian	KOSEKI Ayumi 小関 歩 7419103	KOSEKI Ayumi 小関 歩 7419103	<a href="#">Edit</a>
Puan Zeti	Malaysian	KOSEKI Ayumi 小関 歩 7419103	KOSEKI Ayumi 小関 歩 7419103	<a href="#">Edit</a>
Pak Mat	Malaysian	KOSEKI Ayumi 小関 歩 7419103	KOSEKI Ayumi 小関 歩 7419103	<a href="#">Edit</a>
cik puan	Malaysian	KOSEKI Ayumi 小関 歩 7419103	KOSEKI Ayumi 小関 歩 7419103	<a href="#">Edit</a>
pelayar dan pengikut media sosial Dewan Bahasa dan Pustaka	Malaysian	KOSEKI Ayumi 小関 歩 7419103	KOSEKI Ayumi 小関 歩 7419103	<a href="#">Edit</a>
นักแสดง	Thai	parkhan ngamwannakorn	Poey BKT	<a href="#">Edit</a>
นักร้อง	Thai	parkhan ngamwannakorn		<a href="#">Edit</a>
บุหรี	Thai	parkhan ngamwannakorn	Poey BKT	<a href="#">Edit</a>
พี่ชายตรงนั้น	Thai	sunisa@tufs.ac.jp	sunisa@tufs.ac.jp	<a href="#">Edit</a>

< 1 2 3 4 5 ... 150 > 10 / page

# Database

Example (original language):	<input type="text" value="Hmm, sebab itu Encik Isa menjauhi aktiviti kelab sukan, ya."/>
Japanese translation:	<input type="text" value="ふーん、だからイサさんはスポーツクラブ活動を敬遠しているんですね。"/>
English translation:	<input type="text" value="Hmm, that's why Mr. Isa [= you] avoids doing sports activities."/>
Source:	<input type="text" value="『ニューエクスプレス マレー語』 p. 90"/>
Note:	<input type="text"/>
⊖	
Example (original language):	<input type="text" value="Amboi, hebat betul Encik Isa, ya!"/>
Japanese translation:	<input type="text" value="まあ、イサさんは本当にすごいんですね。"/>
English translation:	<input type="text" value="Wow, Mr. Isa [= you] is really great."/>
Source:	<input type="text" value="『ニューエクスプレス マレー語』 p. 90"/>
Note:	<input type="text"/>
⊖	
<input type="button" value="+ Add an example"/>	

# Common questionnaire for investigating pronoun substitutes and address terms

- <https://github.com/matbahasa/ProSub>

147 lines (147 sloc) | 8.26 KB

Raw Blame

Search this file...

1	1	坊や (10歳くらいまでの年少者、男)	little boy	10285313-n	Age
2	2	お嬢ちゃん (10歳くらいまでの年少者、女)	little girl	10084295-n	Age
3	3	子供 (10歳くらいまでの年少者、男女問わず)	child	09917593-n	Age
4	4	お嬢さん (10代後半から20代前半、女)	miss	10084295-n	Age
5	5	若い男性 (10代後半～20代、男)	young man	07944050-n	Age
6	6	若い女性 (10代後半～20代、女)	young woman	10129825-n	Age
7	7	若者 (10代後半～20代、男女問わず)	youth	07944050-n	Age
8	8	父	father	10080869-n, 09988063-n	Family
9	9	母	mother	10332385-n, 10278128-n	Family
10	10	伯父 (父の兄)	uncle (father's elder brother)	10736091-n	Family
11	11	伯父 (母の兄)	uncle (mother's elder brother)	10736091-n	Family
12	12	叔父 (父の弟)	uncle (father's younger brother)	10736091-n	Family
13	13	叔父 (母の弟)	uncle (mother's younger brother)	10736091-n	Family
14	14	伯母 (父の姉)	aunt (father's elder sister)	09823502-n	Family

# Corpus annotation (AY2022-2023)

先生：小春さんのお母さんはいくつか先生に教えてください。

**ProSub.2nd**

**ProSub.1st**

Teacher: Tell teacher [= me] how old Miss Koharu's [= your] mother is.

小春：お父さん、お母さんの年っていくつだっけ？

**AddressTerm**

Koharu: Father, how old was mum?

父：父さんも忘れたよ。

**ProSub.1st**

Father: Father [= I] also forgot how old she was.

# References

- Faridah Mohamed and Yumi Kondo. 1999. *Ekusupuresu Mareego [Express Malay]*. Tokyo: Hakusuisha.
- Faridah Mohamed and Yumi Kondo. 2019. *Nyuu Ekusupuresu Purasu Mareego [New Express Plus Malay]*. Tokyo: Hakusuisha.
- Nomoto, Hiroki, Kenji Okano, David Moeljadi and Hideo Sawada. 2018. [TUFS Asian Language Parallel Corpus \(TALPCo\)](#). *Proceedings of the Twenty-Fourth Annual Meeting of the Association for Natural Language Processing*, 436-439.
- Nomoto, Hiroki, Kenji Okano, Sunisa Wittayapanyanon and Junta Nomura. 2019. [Interpersonal meaning annotation for Asian language corpora: The case of TUFS Asian Language Parallel Corpus \(TALPCo\)](#). *Proceedings of the Twenty-Fifth Annual Meeting of the Association for Natural Language Processing*, 846-849.
- Nomoto, Hiroki, Sunisa Wittayapanyanon (Saito), Kenji Okano, Thuzar Hlaing, Yunjin Nam and Sri Budi Lestari. 2021. Daimeishidaiyou, yobikake hyougen kenkyuu no genjou: Taigo, Birumago, Mareego, Indoneshiago, Jawago, Chousengo [Current state of studies on pronoun substitutes and address terms: Thai, Burmese, Malay, Indonesian, Javanese and Korean]. *Gogaku Kenkyuujo Ronshuu* 25: 63-78. Tokyo University of Foreign Studies. [English version]
- Shaik Omar Mohamed and Azusa Yamazaki. 1997. *Omaa-Azuu no Mareego Kouza [Omar & Azu's Malay Course]*. Tokyo: Mekon.
- Zaharah Othman. 2012. *Colloquial Malay: The Complete Course for Beginners*. London: Routledge, 2nd edn.
- Zaharah Othman and Sutanto Atmosmarto. 1995. *Colloquial Malay: The Complete Course for Beginners*. London: Routledge, 1st edn.