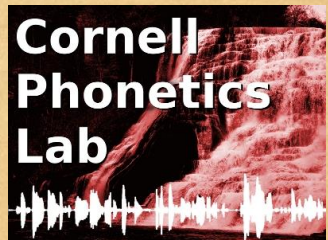


Intergenerational language transmission: evidence from Jakarta Indonesian speech corpora

Ferdinan Okki Kurniawan
fok3@cornell.edu

Department of Linguistics, Cornell University



Purpose of this study

- Investigates intergenerational language transmission in Jakarta Indonesian (JI) using large corpora of three generations of JI speakers:
 - Adults in 1970s
 - Adults in 2000s
 - Children in 2000s
- Through phonological patterns which are transmitted over three generation of speakers, this investigation attempts to shed light on:
 - The development of JI.
 - Its relationships with another variety spoken in Jakarta, namely Jakarta/Betawi Malay.

Main Finding

The findings show that patterns of use of some phonological evidence have changed across three generations of JI speakers.

Outline of the talk

1. Importance of the study
2. Phonological variable
3. Corpora
4. Methodology
5. Findings
6. Summary

Importance of the Study

Importance of the Study

- Most historical linguistics investigations rely on analysis of language change and language reconstruction to discover changes that happen across centuries.
- However, change in progress which involves speakers across generations is less commonly studied, especially with understudied languages such as Jakarta Indonesian.
- It is at this time scale that we can start to understand the mechanism of change.

Importance of the Study

- To investigate linguistic change in progress, Sankoff (2006) proposed two types of longitudinal studies: trend studies and panel studies.
 - Trend study applies data collected at different times but not necessary from the same speakers. Resampling from language community is required for trend studies.
 - Panel study tracks linguistic evidence from the same individual(s) across times.

Importance of the Study

- Using large naturalistic corpora collected from three generations of JI speakers, it is possible to conduct trend and panel study to investigate change in progress.
- Using this naturalistic data, this current study offers new evidence of how phonological patterns of use are transmitted across generations.

Phonological Variables

Phonological Variables

- The patterns of use of:
 - Final vowels [e], correspond to Standard Indonesian (SI) final vowel [a], such as in [ape] ~ SI form [apa] ‘what’, [die] ~ SI form [dia] ‘3rd person pronoun’.
 - Final laryngeals: glottal stop [-ʔ] and glottal fricatives [-h] in phrase final position, such as in [pagiʔ] ~ [pagi] ‘morning’, [ini] ~ [inih] ‘this’.
 - Vowel quality in non-lax and lax vowels: [i] ~ [ɪ], [u] ~ [ʊ], [e] ~ [ɛ], [o] ~ [ɔ].

Final vowel: [e]

Historical account of the final vowel [e]

The emergence of Jakarta (Betawi) Malay

- Two varieties of Malay in early Jakarta (Wallace 1976):
 - Urban Jakarta Malay
 - Rural Jakarta Malay
 - Ikranagara (1981), Muhadjir (1980), Chaer (1976) among others termed it Betawi 'Ora'

The emergence of urban Jakarta Malay

- Portuguese-based creole was used as lingua franca in Jakarta until the 17th century (Ikranagara 1981).
- Urban Jakarta Malay emerged around the 17th – 18th centuries (Wallace 1976); a Malay variety with influence from Arabic, Chinese, Dutch, Portuguese and surrounding local languages.
- Phonological property that distinguish urban Jakarta Malay from other variety of Malay: final vowel [e] as in:
 - [ape] ‘what’, corresponds with SI [apa];
 - [aye] ‘1st person pronoun’, corresponds with SI [saya].

Historical account of final vowel [e]

- It was adopted from the Arabs from Hadhramaut region (Yemen nowadays) who took part in trading in Southeast Asia and eventually settled in Jakarta urban area.
- This Arabic forms were imitated by Jakarta inhabitants and becoming a prestigious markers around the early 19th century.

Historical account of final vowel [e]

- Urban Jakarta Malay speakers uses final vowel [e] across function words and content words

Function words	Urban Jakarta Malay	Indonesian/Malay
Personal pronouns	[die] '3SG', [aye] '1SG'.	[dia], [saya]
Interrogatives	[ape] 'what', [mane] 'where', etc.	[apa], [mana]
Prepositions	[ame] 'with', [daripade] 'instead of'	[ama], [daripada]
Content Words		
Noun	[mate] 'eye', [(kə)pale] 'head'	[mata], [kəpala]
Verb	[taje] 'ask', [puje] 'posses'	[taja], [puja]
adjective	[suse] 'difficult', [mude] 'young'	[susah], [muda]

The emergence of rural Jakarta Malay

- This important garrison caused the surrounding area to be developed into the first suburb of Jakarta.
- Wallace considered the variety spoken in this area as transitional rural-urban blending variety and suggested that it might be the possible origin of Modern Jakarta Malay (MJM) or Jakarta Indonesian (JI; following Grijn's 1991 term of MJM)

Historical account of final vowel [e]

- Rural Jakarta Malay speakers uses final vowel [e] in function words but not in content words

Function words	Rural Jakarta Malay	Indonesian/Malay
Personal pronouns	[die] ~ [dia(h)] '3SG', [aye]/[gue] ~ [gua(h)] '1SG'.	[dia], [saya]
Interrogatives	[ape] ~ [apa(h)] 'what', [mane] ~ [mana(h)] 'where', etc.	[apa], [mana]
Prepositions	[ame] ~ [ama] 'with', [daripade] ~ [daripada] 'instead of'	[ama], [daripada]
Content Words		
Noun	[mataʔ] 'eye', [(kə)palaʔ] 'head'	[mata], [kəpala]
Verb	[tanaʔ] 'ask', [puʔnaʔ] 'posses'	[tana], [puʔna]
adjective	[susah] 'difficult', [mudaʔ] 'young'	[susah], [muda]

To summarize...

- Urban Jakarta Malay:
 - Across function and content words: $a > e / _ \#$
- Rural Jakarta Malay:
 - In function words:
 - $a > \{e, a\} / _ \#$;
 - $\emptyset > h / a _ \#$
 - In content words:
 - $a > a / _ \#$;
 - $\emptyset > \text{ʔ} / a _ \#$

The questions...

- What is the status of these rules in the vernacular spoken in Jakarta nowadays?
- How can we provide a careful and systematic study of these rules so that the evidence from it could help us to have a better understanding about the emergence and development of JI?
- In order to address these questions, an investigation across generations is needed.
- We need corpora that could describe actual language use and development in each generation.

Corpora

Corpora

1. A corpus of adults speech collected in early-mid 1970s (Wallace 1976)
2. A corpus of adults speech collected in early 2000s (Gil and Tadmor 2015)
3. A corpus of children speech collected in early 2000s (Gil and Tadmor 2015)

Corpus of Adults Speech (1970s)

- Wallace (1976)
 - Documented in Jakarta the early-middle of 1970s as the basis for his doctoral thesis at Cornell University, Department of Linguistics.
 - Investigated socio-phonological aspects of Jakarta Malay.
 - The corpus comprised data from around 35 hours of recordings and involved around 250 adult speakers in informal settings.

Corpus of Adults Speech (2000s)

- Gil and Tadmor (2015)
 - Collected and transcribed at MPI, Jakarta Field Station between 2004-2012, and consists of adult-to-adult conversations in informal settings.
 - This corpus involves 69 adults from various socio-economic backgrounds.
 - There is a total of 75,079 transcribed utterances in this corpus.

Corpus of Children Speech (2000s)

- Gil and Tadmor (2015)
 - Collected longitudinally over the course of four years; and transcribed at MPI, Jakarta Field Station between 2000-2012.
 - Consists of children speech in daily settings.
 - Involves 10 target children from various socio-economic backgrounds.
 - There is a total of 915,182 transcribed utterances in this corpus.

Methodology

Methodology: Speakers' ethnic classification

- Wallace classified his speakers in his 1970s data into:
 - Traditional Jakarta Malay (TJM): speakers of Betawi ethnicity.
 - Modern Jakarta Malay (MJM)/JI: speakers of those who have ethnic background other than Betawi (their parents are not of Betawi ethnicity), but were born and grew up in Jakarta.
- For this current study, I follow Wallace's classification for MJM/JI speakers in 2000s data.

Methodology: number of speakers

	Adults 1970s	Adults 2000s	Pre-adolescence 2000s (from children corpus)
Lower Socio-Economic Status (SES)	6	2	2
Middle SES	4	2	2
Upper SES	2	excluded; very limited tokens	excluded; very limited tokens

Search criteria

- Limited only to function words uttered by speakers.
 - Examples of function words found in the corpus:
 - gua ~ gue ‘1SG’, iya ~ iye ‘yes’, ya ~ ye ‘yes’, -ɲa ~ ɲe ‘possessive, article’, apa ~ ape ‘what’, kənapa ~ kənape ‘why’, bərapa ~ bərape ‘how.much’, and so forth.
- Identify the varied patterns of $a > \{a, e\} / __\#$ in function words
- Excluding content words: $a > a / __\#$

Findings

Wallace's (1976) findings

- From his 1970s data, Wallace reported similar findings with the situation in the early 19th centuries:
 - TJM speakers' used of final vowel [e] across function words and content words.
 - MJM speakers used final vowel [a] ~ [ah] ~ [aʔ] in content words and final vowel [a] ~ [e] in function words.

Wallace's (1976) findings

- In his 1970s data, the patterns of use of final vowel [e] in function words among MJM speakers were quite high.
 - Speakers from lower SES show higher distribution: 98%, $n = 6$
 - Speakers from middle SES show higher distribution: 90%, $n = 4$
 - Speakers from upper SES show much lower distribution: 22%, $n = 2$

The current findings: Adult speakers in 2000s corpus

- Adult speakers in 2000s corpus produce lower distribution of final vowel [e] in function words than 1970s adult speakers:
 - Speakers from lower SES show low distribution: 11% $n = 2$
 - Final vowel [a]: 786/1024
 - Final vowel [e]: 108/1024
 - Speakers from middle SES show low distribution: 4% $n = 2$
 - Final vowel [a]: 860/888
 - Final vowel [e]: 28/888

The current findings: pre-adolescent speakers in 2000s corpus

- Pre-adolescent speakers in 2000s corpus produce even lower than 2000s adult speakers
 - Speakers from lower SES show lower distribution: less than 1%, $n = 2$;
 - Final vowel [a]: 360/362
 - Final vowel [e]: 2/362
 - Speakers from middle SES show lower distribution: 1%, $n = 2$
 - Final vowel [a]: 512/517
 - Final vowel [e]: 5/517

Examples of the tokens: an adult speaker from middle SES (2000s corpus)

Function Words	Gloss	Final vowel -a	Final vowel -e
gua	1SG		4
iya	yes		47
ya	yes		104
nya	DET; POSS		97
apa	what		50
kenapa	why		14
brapa	how.much		13
siapa	who		6
mana	where		16
gimana	how		18
kita	1PL.incl		7
dia	3SG		25
aja	just; only		19
ada	exist		50
tiga	three		6
dua	two		10
lima	five		3
dah	PFCT		1
udah	PFCT		33
ama	with		4
suka	often; sometimes		1
			528
			8

Examples of the tokens: an adult speaker from lower SES (2000s corpus)

Function Words	Gloss	Final vowel -a	Final vowel -e
iya	yes	62	4
ya	yes	152	34
nya	DET; POSS	156	60
apa	what	13	
kenapa	why	4	
brapa	how.much	4	
siapa	who	7	
mana	where	2	
gimana	how	4	
kita	2PL.INCL	3	
gua	1SG	4	
dia	3SG	72	3
aja	just; only	23	3
ada	exist	77	
tiga	three	4	
dua	two	21	
lima	five	5	
dah	PFCT	2	
udah	PFCT	20	3
ama	with	1	
suka	often; sometimes	10	
selama	as.long.as	1	
		647	107

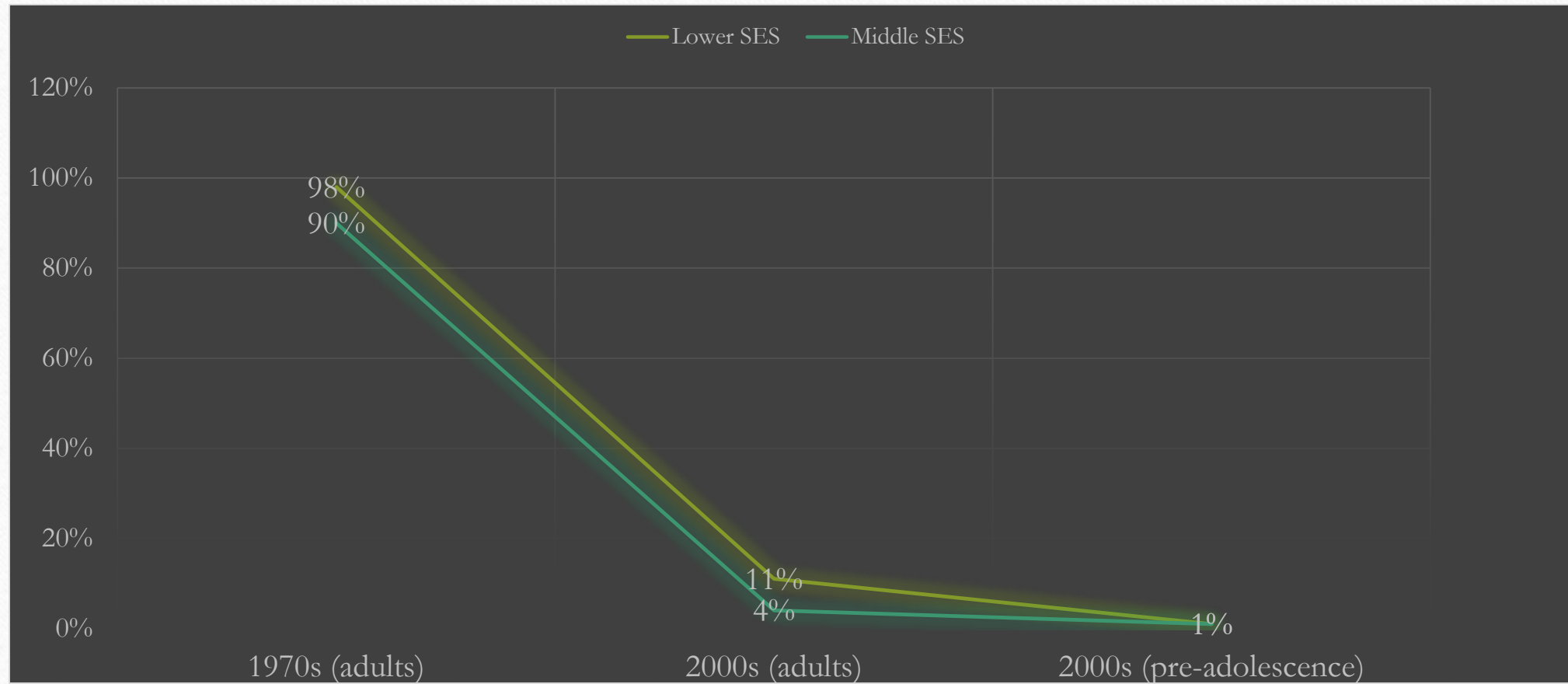
Examples of the tokens: a pre-adolescent speaker from middle SES (2000s corpus)

Function Words	Gloss	Final vowel -a	Final vowel -e
aja	just; only	39	
dah	PFCT	13	3
udah	PFCT	18	
nya	DET; POSS	131	1
ya	yes	90	1
iya	yes	23	
ada	exist	36	
dua	two	28	
lima	five	15	
tiga	three	16	
pa-pa (apa-apa)	RED-what	3	
apa	what	38	
mana	where	27	
sapa	who	1	
siapa	who	9	
berapa	how.much	13	
kenapa	why	1	
gimana	how	3	
dia	3SG	5	
ama	with	1	
sama	with	1	
daripada	instead.of	1	
		512	5

Examples of the tokens: a pre-adolescent speaker from lower SES (2000s corpus)

Function Words	Gloss	Final vowel -a	Final Vowel -e
ya	yes	60	
iya	yes	27	
kenapa	why	1	
ngapa	why	1	
napa	why	2	
gimana	how	1	
berapa	how.much	1	
siapa	who	10	
mana	where	23	
pa-pa	RED-what	1	
apa	what	20	
nya	DET; POSS	116	
dah	PFCT	3	2
udah	PFCT	13	
ama	with	15	
aja	only; just	24	
dia	3SG	4	
ada	exist	23	
dua	two	7	
lima	five	3	
tiga	three	5	
		360	2

Distribution of final vowel [e] in function words across generations



Distribution of final vowel [e] in function words across generations

- The patterns of use of final vowel [e] in function words were not faithfully transmitted from parents generation (adults in 1970s) to their children generation (adults in 2000s).
- The patterns of use of final vowel [e] in function words were faithfully transmitted from parents generation (adults in 2000s) to their children generation (children in 2000s); however, it shows slightly decreasing patterns of use in children generation.

Final laryngeals

Final laryngeals

- The patterns of use of glottal stop [-ʔ] and glottal fricatives [-h] in phrase final position. Examples:
 - dəkət apan waruŋ jambuʔ (JFS text ID: 957007105517080704)
 - Possible variation: [jam.buʔ] ~ [jam.bu]. To formulize: .(C)uʔ ~ .(C)u/___#
 - kaya ʔiku:t ʔabri gituh (JFS text ID: 337010154057160804)
 - Possible variation: [gituh] ~ [gitu]. To formulize: .(C)uh ~ .(C)u/___#

Final laryngeals: Phonetic environment where variation might occur

Vowel quality	Glottal stop class		Glottal fricative class	
	Non-lax	Lax	Non-lax	Lax
High	$.(C)iʔ \sim .(C)i / __\#$	$.(C)ɪʔ \sim .(C)ɪ / __\#$	$.(C)ih \sim .(C)i / __\#$	$.(C)ɪh \sim .(C)ɪ / __\#$
High	$.(C)uʔ \sim .(C)u / __\#$	$.(C)ʊʔ \sim .(C)ʊ / __\#$	$.Cuh \sim .(C)u / __\#$	$.(C)ʊh \sim .(C)ʊ / __\#$
Mid-back	$.(C)oʔ \sim .(C)o / __\#$	$.(C)ɔʔ \sim .(C)ɔ / __\#$		
Mid-front	$.(C)eʔ \sim .(C)e / __\#$	$.(C)ɛʔ \sim .(C)ɛ / __\#$		

Final laryngeals

- Glottal fricative class is found in deictic forms such as [inɪh], [ituh]; personal pronoun [luh], [gua].
- Glottal stop class are found in other forms: verb, nouns, numerals, etc.
- Pro-clitics such as [di], [ke] are never been laryngealized.

Historical Perspective of the Final Laryngeals

- In the early formation of Jakarta/Betawi Malay in the 17th - 18th centuries (Wallace 1976):
 - Javanese and Balinese formed the urban Jakarta/Betawi Malay; thus no laryngealization.
 - Sundanese formed the rural Jakarta/Betawi Malay; thus laryngealization in most .CV in word and phrase final position.;
 - Migration of rural speakers to urban (inner city) area caused variation (emergence of innovative speakers) in non-low vowel forms: [i, u, e, o].

How do these patterns of use of final laryngeals in 1970s and 2000s data look like?

Wallace's (1976) findings

- He classified his MJM/JI speakers into innovative vs conservative ones
 - These innovative vs conservative speakers are under the same ethnic classification as MJM/JI (first generation of immigrants).
 - Conservative JI speakers:
 - Laryngealization occurs very rarely in non-low vowels, lexically determined, mostly males from low SES, mostly live in Betawi neighborhood.
 - Innovative JI speakers:
 - Laryngealization is irregular and inconsistent, mostly females no matter their SES (if males, they are from middle-high SES), persons live in non-Betawi neighborhood.
- Unfortunately, there is no distribution percentage provided in his study.

Current findings: Adult speakers in 2000s corpus

- Adult speakers in 2000s corpus produce high distribution for final laryngeals
 - Males from lower SES show high distribution: 92 % $n = 2$
 - Final zero laryngeals: 34/419
 - Final laryngeals: 385/419
 - Speakers (1 male and 1 female) from middle SES show high distribution 89% $n = 2$
 - Final zero laryngeals: 60/493
 - Final laryngeals: 433/493

Current findings: Pre-adolescent speakers in 2000s corpus

- Pre-adolescent speakers in 2000s corpus also produce high distribution for final laryngeals
 - Speakers from lower SES show high distribution: 98% $n = 2$
 - Final zero laryngeals: 5/361
 - Final laryngeals: 356/361
 - Speakers from middle SES show high distribution 96% $n = 2$
 - Final zero laryngeals: 18/584
 - Final laryngeals: 566/584

Final Laryngeals: general picture

17th – 18th Centuries:
Urban Jakarta/Betawi Malay: no final laryngealization

19th centuries:
Migration from rural to urban (inner city) brought final laryngealizations

1970s: conservative MJM/JI speakers produced very rare laryngealization, resembling Jakarta/Betawi Malay; innovative speakers emerged with inconsistent use of final laryngealization.

2000s: both adult and pre-adolescent speakers show robust distribution of final laryngealization

Conclusion & orientation for further research

Conclusion & orientation for further research

- Conclusion:
 - These findings suggest that patterns of use of final vowel [e] and final laryngeals have changed across these three generations of speakers, if the data from these speakers are representative.
 - The patterns of use of final vowel [e] that had been faithfully transmitted from generation to generation for more than two centuries (from the early 19th centuries – 1970s), has drastically changed within these three generations.
 - The patterns of use of final laryngeals produced by the innovative speakers in 1970s have been faithfully transmitted to the 2000s adult and pre-adolescent speakers.
- Orientation for further research:
 - More speakers need to be involved.
 - Parallel with this study, I am also conducting an acoustic study (speech production task) to investigate vowel quality.
 - Probabilistic (formal or statistical?) modeling should be further developed.

Thank you!

Acknowledgement

- I would like to thank: Abby Cohn, John Wolff, and Draga Zec for their valuable input. All errors in this talk are my responsibility.

References

- Chaer, A. 2008. *Morfologi Bahasa Indonesia (Pendekatan Proses)*. Jakarta: Rineka Cipta.
- Gil, D. and U. Tadmor. 2015. The MPI-EVA Language Database. A joint project of the Department of Linguistics, Max Planck Institute for Evolutionary Anthropology and the Center for Language and Culture Studies, Atma Jaya Catholic University.
- Ikranagara, K. 1980. *Melayu Betawi Grammar*. *NUSA Linguistic Studies in Indonesian and Languages in Indonesia*, Vol. 9, Atma Jaya University, Jakarta.
- Miesel, J., M. Elsig and E. Rinke. 2013. *Language Acquisition and Change: A Morphosyntactic Perspective*. Edinburgh: Edinburgh University Press.
- Wallace, S. 1976. *Linguistic and Social Dimensions of Phonological Variation in Jakarta Malay*. Doctoral Dissertation and transcribed recordings. Cornell University, Ithaca, NY.