

Syntax and Semantics of Adjectives in Indonesian

Yosephine, Hannah Choi, David Moeljadi, and Francis Bond

The Twenty-Second International Symposium on Malay/Indonesian Linguistics (ISMIL 22)
University of California, Los Angeles
11-12 May 2018



Objectives (on progress)

- To study the **syntactic** patterns and **semantic** functions of **Indonesian adjectives**
- A special emphasis on **emotion adjectives** taking the active voice prefix ***meN-*** and its interaction with the suffixes ***-i*** and ***-kan***
- **Adjective classification/subcategorization** based on the syntactic and/or semantic patterns
- For various purposes such as in computational grammar

Outline

- Adjectives in Indonesia
- Methodology and the tests
- Results (syntactic patterns)
- Discussion (semantic roles)
- Summary and future research

Adjectives

- adjectives = varied and hard to recognize (Dixon, 2004, p. 9)
 - They can fill in the intransitive predicate position
 - They can fill in the copula complement position

Adjectives in Indonesian

Two functions of adjectives in Indonesian (Alwi et al, 2014, p. 177):

- **Predicative**

(1) Dia marah

3SG angry

“He’s angry.”

(intransitive predicate position)

(2) Anak yang memakai kemeja merah itu (adalah) pintar

kid REL MEN-wear shirt red that (COP) smart

“The kid who is wearing a red shirt is smart.” (copula complement position)

- **Attributive**

(3) Anak cantik itu

kid pretty that

“That pretty kid”

What they are not

- Adjectives ≠ Nouns

(5) *Dia tidak/*bukan cantik*

3SG NEG pretty

“She is not pretty.”

- Adjectives ≠ Verbs

(6) *Anak itu sangat kuat*

kid that very strong

“The kid is very strong.”

Adjectives = Verbs

- Adjectives in Indonesian can take verbal affixes like verbs such as *meN* and *beR-*, to form *menyedihkan* ‘to make someone/something sad’ and *bersedih* ‘to be sad’.

(7) *Berita itu menyedihkan*
news that MEN-sad-KAN
“The news is sad.”

KBBI (Kamus Besar Bahasa Indonesia)

← → ↻ Secure | <https://kbbi.kemdikbud.go.id/entri/menyenangkan>



KBBI Daring Cari Seputar Laman

 Informasi: Temukan bantuan menggunakan KBBI Daring [di sini](#).



senang » me.nye.nang.kan

1. ✓ menjadikan senang; membuat bersuka hati: *ia menyayangi adiknya dan selalu berusaha ~ hatinya*
2. ✓ membangkitkan rasa senang hati; memuaskan; menarik (hati): *semoga perjalanan ini ~ Tuan*
3. ✓ merasa senang (puas dan sebagainya) akan; menyukai

The confusion

Adjectives = Intransitive verbs?

(8) Permainan itu menyenangkan.
PER-play-AN that MEN-happy-APPL
“The game is fun.” (Predicative)

(9) Permainan menyenangkan itu adalah susah.
PER-play-AN MEN-happy-APPL that COP hard
“The fun game is hard.” (Attributive)

Methodology

Target roots: 10 common emotion words/roots
(Musgrave, 2001)

- benci 'hate'
- cinta 'love'
- malu 'embarrassed'
- senang 'glad'
- khawatir 'worry'
- takut 'afraid'
- kecewa 'disappointed'
- kasih 'love'
- marah 'anger'
- sedih 'sad'

Methodology

- LCC Corpus: Reclassified Indonesian Leipzig Corpora Collection (Quasthoff, 2006; Nomoto et al, to appear)
- Python script: Search for target roots, both upper and lower case
- Frequency count per target root and morphological pattern
- Classification by syntactic behavior

Methodology

	<i>MeN-</i>	<i>MeN-</i> and <i>-kan</i>	<i>MeN-</i> and <i>-i</i>
<i>cinta</i> 'love'	<i>mencinta</i>		<i>mencintai</i>
<i>benci</i> 'hate'	<i>membenci</i>		
<i>senang</i> 'happy'		<i>menyenangkan</i>	<i>menyenangi</i>
<i>takut</i> 'afraid'		<i>menakutkan</i>	<i>menakuti</i>
<i>sedih</i> 'sad'		<i>menyedihkan</i>	
<i>kecewa</i> 'disappointed'		<i>mengecewakan</i>	
<i>khawatir</i> 'worried'		<i>mengkhawatirkan</i>	
<i>malu</i> 'embarrassed'		<i>memalukan</i>	
<i>marah</i> 'angry'			<i>memarahi</i>
<i>kasih</i> 'love'			<i>mengasihi</i>

Methodology

- Python script: Search for 13 target words, both upper and lower case
- Extracted all sentences containing the target words into separate tsv files

Methodology

	<i>MeN-</i>	<i>MeN-</i> and <i>-kan</i>	<i>MeN-</i> and <i>-i</i>
<i>cinta</i> 'love'	<i>mencinta</i> (109)		<i>mencintai</i> (8420)
<i>benci</i> 'hate'	<i>membenci</i> (2439)		
<i>senang</i> 'happy'		<i>menyenangkan</i> (10205)	<i>menyenangi</i> (416)
<i>takut</i> 'afraid'		<i>menakutkan</i> (1440)	<i>menakuti</i> (179)
<i>sedih</i> 'sad'		<i>menyedihkan</i> (2195)	
<i>kecewa</i> 'disappointed'		<i>mengecewakan</i> (3040)	
<i>khawatir</i> 'worried'		<i>mengkhawatirkan</i> (4714)	
<i>malu</i> 'embarrassed'		<i>memalukan</i> (2386)	
<i>marah</i> 'angry'			<i>memarahi</i> (780)
<i>kasih</i> 'love'			<i>mengasihi</i> (2001)

- Collocate Search: AntConc

AntConc 3.5.6 (Macintosh OS X) 2018

Concordanc Concordance Pl: File Vie Clusters/N-Gram Collocate Word Li: Keyword Li:

Concordance Hits 4744

Hit	KWIC	File
1	asi Asuransi Umum Indonesia (AAUI) mengkhawatirkan , model pembukua	mengkhawatirk
2	36 Menteri BUMN Mustafa Abubakar mengkhawatirkan adanya masalah in	mengkhawatirk
3	— Menteri BUMN Mustafa Abubakar mengkhawatirkan evaluasi gaji direks	mengkhawatirk
4	— Menteri BUMN Mustafa Abubakar mengkhawatirkan evaluasi gaji direks	mengkhawatirk
5	at publik, apa yang dilakukan Aceng mengkhawatirkan . 18377 Jika tidak	mengkhawatirk
6	g meleset dari target. 16776 Adalah mengkhawatirkan bahwa kasus yang	mengkhawatirk
7	Maliki menyatakan: "Darurat adalah mengkhawatirkan diri dari kematian k	mengkhawatirk
8	macam kejelasan sejauh euro adalah mengkhawatirkan , (meskipun) masih	mengkhawatirk
9	katanya menegaskan. 15400 Adang mengkhawatirkan dampak dari DPT	mengkhawatirk
10	DPRD Kota Solok, Irman Yefri Adang , mengkhawatirkan terjadinya pengura	mengkhawatirk
11	asal Argentina itu. 17174 Adhyaksa mengkhawatirkan PON XVII di Kaltim	mengkhawatirk
12	ersitas Gajah Mada, Sri Adiningsih , mengkhawatirkan kerja sama dengar	mengkhawatirk
13	ahan di sini. 18575 Cuma yang agak mengkhawatirkan adalah faktor cuac	mengkhawatirk
14	per. perhatikan kesehatan anda agak mengkhawatirkan , cobalah pergi ke l	mengkhawatirk
15	664 Melihat posisi PSMS yang agak mengkhawatirkan ini , tak pelak memi	mengkhawatirk
16	di negara maju itu kan agak mengkhawatirkan jadinya kapasitas k	mengkhawatirk
17	petisi itu. 18976 Satu hal yang agak mengkhawatirkan , kalau dilihat deng	mengkhawatirk

Search Term Words Case Regex Search Window Size 50

 Advanced

Start Stop Sort Show Every Nth Row 1

Kwic Sort Level 1 1L Level 2 1R Level 3 0 Clone Results

Corpus Files

- memalukan.txt
- memarahi.txt
- membenci.txt
- menakuti.txt
- menakutkan.txt
- mencinta.txt
- mencintai.txt
- mengasih.txt
- mengecewakan.txt
- mengkhawatirkan.txt**
- menyedihkan.txt
- menyenangi.txt
- menyenangkan.txt

Total No. 13
Files Processed

Corpus Files

- memalukan.txt
- memarahi.txt
- membenci.txt
- menakuti.txt
- menakutkan.txt
- mencinta.txt
- mencintai.txt
- mengasahi.txt
- mengecewakan.txt
- mengkhawatirkan.txt
- menyedihkan.txt
- menyenangi.txt
- menyenangkan.txt

Total No.
13

Files Processed

Concordanc Concordance Plr File Vie Clusters/N-Granr Collocate Word Li: Keyword Li:

Total No. of Collocate Types: 1880 Total No. of Collocate Tokens: 18976

Rank	Freq	Freq(L)	Freq(R)	Stat	Collocate
1	1158	1156	2	4.96435	sangat
2	964	892	72	2.34011	yang
3	402	402	0	6.17095	cukup
4	364	328	36	2.53714	tidak
5	318	306	12	3.93669	juga
6	284	278	6	5.78601	semakin
7	238	232	6	4.68415	masih
8	248	228	20	3.63837	lebih
9	226	226	0	6.17113	terlalu
10	256	210	46	3.52345	ia
11	232	184	48	3.42890	dia
12	176	174	2	5.90160	perlu
13	376	158	218	1.47825	dan
14	120	120	0	3.33734	sudah
15	164	112	52	1.75238	ini
16	152	106	46	2.37626	mereka
17	114	104	10	5.96461	makin

Search Term Words Case Regex

mengkhawatirkan **Advanced** **Window Span** Same

From... 1L To... 1R

Start Stop Sort

Sort by Invert Order

Sort by Freq(L)

Min. Collocate Frequency 1

Clone Results

- POS count for first 100 sentences per target word
- Assigning semantic roles for the first 10 sentences in each group

Test 1: A relativizer is optional for attributive adjectives

(10) a. Anak yang cantik itu adalah teman saya
kid REL pretty that COP friend 1SG
“The pretty kid is my friend.”

b. Anak cantik itu adalah teman saya

(11) a. Anak yang sedang tidur itu adalah teman saya
kid REL PROG sleep that COP friend 1SG
“The kid who is sleeping is my friend.”

b. *Anak sedang tidur itu adalah teman saya.

Test 2: Adjectives can take intensifiers

(12) Anak sangat cantik itu adalah teman saya. (Adj)

kid very pretty that COP friend 1SG

“That kid who is very pretty is my friend.”

(13) *Anak sangat tidur itu adalah teman saya. (V)

kid very sleep that COP friend 1SG

Test 3: Adjectives cannot take *saling* “each other”

(14) *Mereka saling cantik. (Adj)

3PL RECP pretty

(15) Mereka saling memukul. (V)

3PL RECP MEM-hit

“They hit each other.”

Test 4: Verbs are preceded by preposition *untuk* “for”

(16) *Ini untuk cantik (Adj)
this for pretty

(17) Ini untuk makan (V)
this for eat
“This is for eating.”

Results - The collocates

- The collocates (meN-)

No.	Left (-1)	Node	Right (+1)
1	yang	membenci	orang
2	sangat		dan
3	dan		mereka
4	saling		sesuatu
5	untuk		Islam
6	ia		kamu
7	tidak		aku
8	dia		apa
9	akan		karena
10	semakin		ayahnya

- The collocates (meN- and -kan)

No.	Left (-1)	Node	Right (+1)
1	sangat	mengkhawatirkan	karena
2	yang		jika
3	cukup		kondisi
4	tidak		dan
5	juga		hal
6	semakin		bahwa
7	masih		adanya
8	lebih		akan
9	terlalu		bagi
10	ia		terjadinya

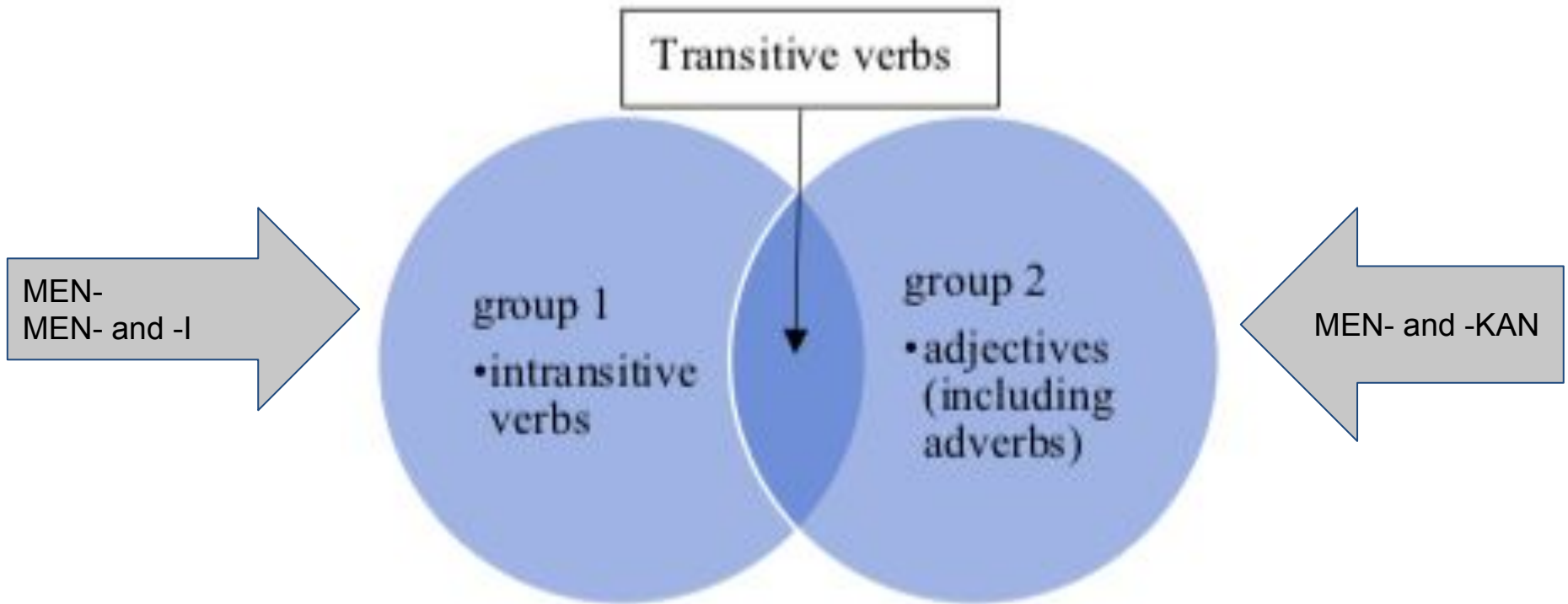
- The collocates (meN- and -i)

No.	Left (-1)	Node	Right (+1)
1	sangat	mencintai	dan
2	yang		orang
3	saling		allah
4	untuk		seseorang
5	dan		diri
6	ia		kita
7	tidak		dirinya
8	lebih		lingkungan
9	masih		mereka
10	saya		dunia

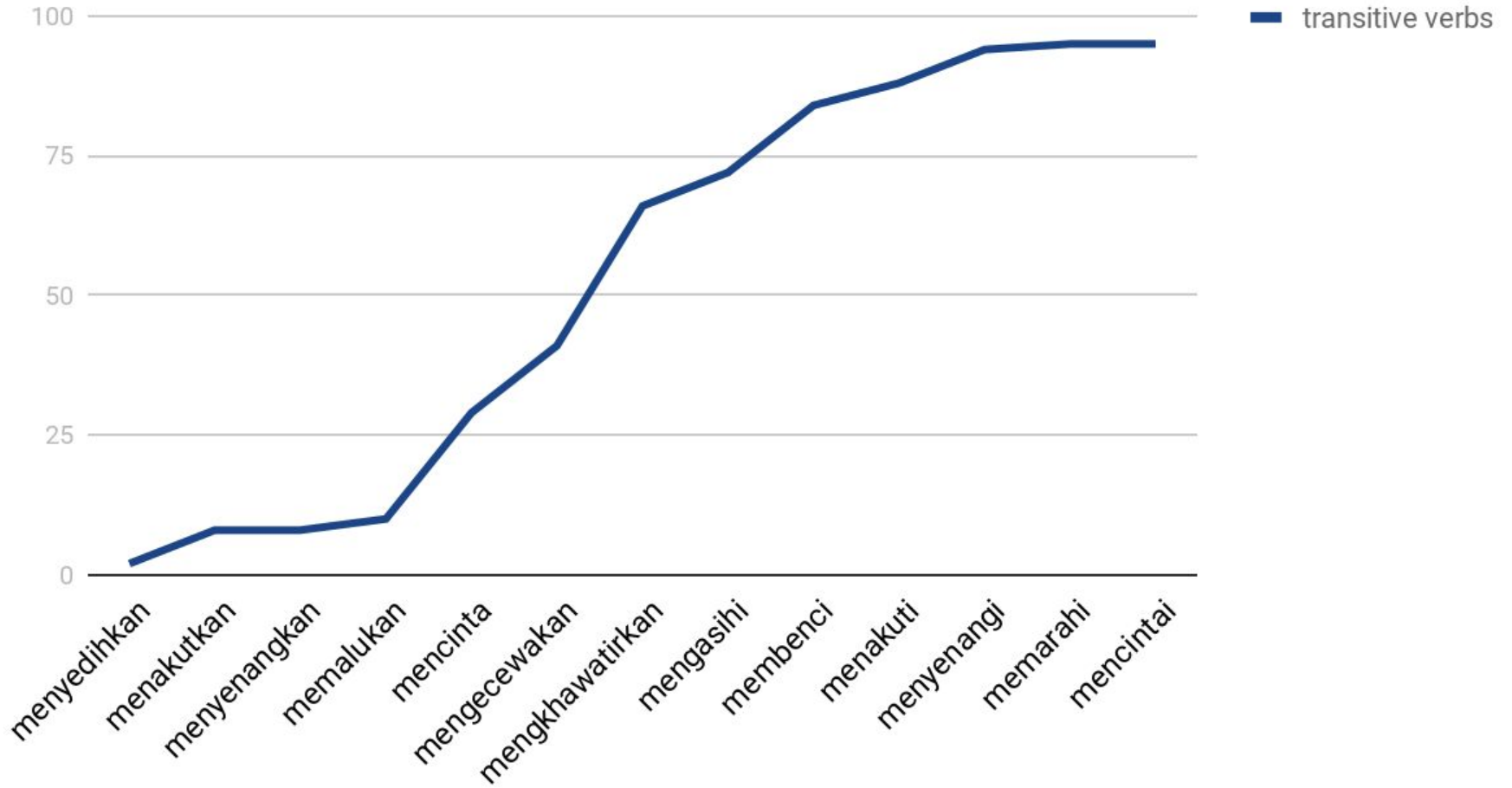
Results - The Distribution

	A	B	C	D	E	F	G
1	prefix- / -suffix	Surface form	Count (upper and	transitive	intransitive	adjective	adverbs
2	Mem-	membenci	2439	84	16	0	0
3		mencinta	109	29	71	0	0
4							
5	Men- and -kan	memalukan	2386	10	0	82	8
6		menyenangkan	10205	8	0	89	3
7		mengkhawatirkan	4714	66	0	34	0
8		menakutkan	1440	8	0	92	0
9		menyedihkan	2195	2	0	97	1
10		mengecewakan	3040	41	0	58	1
11							
12	Men- and -i	mencintai	8420	95	5	0	0
13		menakuti	179	88	12	0	0
14		menyenangi	416	94	6	0	0
15		memarahi	780	95	5	0	0
16		mengasihi	2001	72	28	0	0

Results - Distribution in Syntactic patterns



Transitive verbs



Discussion: Theta roles for NP in Adj constructions

(18) Itu agak memalukan.

theme?

that rather embarrassing

“That is rather embarrassing.”

(19) kekalahan yang memalukan

theme?

loss REL embarrassing

“embarrassing loss”

Result - Semantic patterns

- Group 1: meN- and meN-...-i

(20) Saya mengasihi tim Barca ini
agent experienter/theme
1SG MEN-love-I team barca this
“I love this Barca team.”

(21) Ajaran agama sangat membenci perceraian
theme theme
teaching religious very MEM-hate divorce
“Religious teachings hate divorce very much”

- Group 2: meN-...-kan

(22) kekalahan yang memalukan

theme

loss REL MEM-embarrassed-KAN

“embarrassing loss”

Summary: So far...

We found some syntactic patterns and classified the adjectives into 2 groups but we are still not sure if there's some patterns in the semantics

Further Research

- work on:
 - more emotion words, expand it to other adjectives
 - the distribution of semantic patterns
- contribute to:
 - KBBI (*Kamus Besar Bahasa Indonesia*)
 - INDRA (*Indonesian Resource Grammar*)

References

- Alwi, H., Dardjowidjojo, S., Lapoliwa, H., Moeliono, A. (2014). *Tata bahasa baku bahasa Indonesia (3rd edition)*. Jakarta: Balai Pustaka.
- Anthony, L. (2018). AntConc (Version 3.5.6) [Computer Software]. Tokyo, Japan: Waseda University. Available from <http://www.laurenceanthony.net/software>.
- Dora Amalia, ed. Kamus Besar Bahasa Indonesia. 5th ed. Jakarta: Badan Pengembangan dan Pembinaan Bahasa, 2016.
- Dixon, R. (2004). Adjective class in typological perspective. In R.M.W. Dixon & A. Y. Aikhenvald, *Adjective classes: A cross-linguistic typology*. Oxford: Oxford University Press.
- Musgrave, S. (2001). Non-subject arguments in Indonesian (Doctoral dissertation). Melbourne: The University of Melbourne.
- Nomoto, H., Akasegawa, S., & Shiohara, A. (to appear). Reclassifying the Leipzig corpora collection for Malay and Indonesian. *Nusa*.
- Quasthoff, U., Richter, M. and Biemann, C. (2006). Corpus portal for search in monolingual corpora. In *Proceedings of the Fifth International Conference on Language Resources and Evaluation*, 1799–1802. Genoa.
- Quasthoff, U., Richter, M. and Biemann, C. (2006-). *Corpora collection Leipzig university*. Retrieved March 31, 2018, from https://corpora.informatik.uni-leipzig.de/en?corpusId=ind_mixed_2013.
- Sneddon, J, Adelaar, A., Djenar, D.N., & Ewing, M. C. (2010). *Indonesian reference grammar 2nd edition*. Crows Nest, Australia: Allen and Unwin.

Thank You

Terima Kasih