Usage of Indonesian Possessive Verbal Predicates:

A Statistical Analysis Based on Storytelling Survey[1]

David MOELJADI

davidmoeljadi@yahoo.com

Abstract

This paper deals with possessive verbal predicates in Indonesian, both the present high variety which is originally based on Riau Malay, and the present low variety, which is called "Colloquial Jakartan Indonesian" in Sneddon (2006). The eight predicates in Moeljadi (2010) are chosen as the object of discussion: three possessive verb predicates, *memiliki, mempunyai,* and *punya*; two existential verb predicates, *ada* and *ada ...=nya*; and three predicates with the denominal affixes: *ber-, ber-...-kan*, and *-an*. This paper tries to answer whether the frequency of occurrence of each possessive verbal predicate differs according to whether it appears in the high or low variety. A storytelling survey was conducted in Malang and Tokyo in 2011, in order to determine the speakers' choice of predicates in both varieties, based on the assumption that the speakers choose different possessive verbal predicates for different varieties. Quantitative analysis of variables and correlation coefficient are employed to investigate the statistical relationship between the low variety's tokens and the predicate tokens. The main result is that *memiliki* is primarily used in the high variety, while *punya* is very frequently used in the low variety.

1. Overview of Indonesian

Indonesian, which is called *bahasa Indonesia* by its speakers, is classified as a Malayic language of Malayo-Polynesian branch of Austronesian language family (Austronesian > Malayo-Polynesian > Malayo-Sumbawan > North and East > Malayic > Malay > Indonesian) (Lewis ed. 2009). It is spoken mainly in the Republic of Indonesia, as the sole official and national language and as the common language for hundreds of ethnic groups living there (Alwi et al. 2000: 1-2). Indonesian is a diglossic language (Alwi et al. 2000: 10-11, Sneddon 2006: 3-4). The high

---

[1] Previous studies, overview of possessive verbal predicates, and questionnaire survey for clustering possessive verbal predicates were presented at the fifteenth International Symposium on Malay/Indonesian Linguistics (ISMIL 15) on June 25, 2011 and published in Tokyo University Linguistic Papers (TULIP) 31, page 117-133. The original version of this paper was presented at the fifth Austronesian and Papuan Languages and Linguistics conference (APLL5) on May 5, 2012.

variety of Indonesian, also called *bahasa resmi* 'official language' or *bahasa baku* 'standard language', is based on Riau Malay of northeast Sumatra (Alwi et al. 2000: 12, 15). It is the language of government, law, administration, formal situations (such as speeches and lectures), mass media, literature, and education. In contrast, the low variety, also called *bahasa informal* 'informal language' or *bahasa takbaku* 'non-standard language', is based on the colloquial variety in Jakarta, the capital city of Indonesia (Sneddon 2006). It is the language of everyday communication between Indonesians. However, there is no clear dividing line between the high and the low varieties of Indonesian. There are intermediate forms, associated with semi-formal context, between the two varieties (Sneddon 2006: 6-7). The present paper deals with both the high and the low varieties of Indonesian.

2. Overview of possessive verbal predicates in Indonesian

For the overview of possessive verbal predicates in Indonesian, I mainly refer to Moeljadi (2010) since it deals with the largest number of possessive verbal predicates.

2.1   Constructions with possessive verbs *memiliki*, *mempunyai*, and *punya*

Possessive verbs *memiliki* and *mempunyai* are derived from the roots *milik* and *punya* respectively, with the actor voice prefix *me-* and the suffix *-i* attached to the roots. The applicative suffix *-i* with a nominal base generally forms transitive verbs which has many meanings such as 'to put the nominal base to the object' (Alwi et al. 2000: 124) and 'to act as, or to be the nominal base with reference to the object' (Sneddon 1996: 86). *Milik* originally means 'property' or 'possessions'; thus, *milik-i* may have the meaning 'to ascribe the sense of property/possessions to the object mentioned'. *Punya* has the original meaning 'master', 'lord', or 'possessor'; *punya-i* may have the original meaning 'to be the possessor of the object mentioned'.

Possessive verbs *memiliki* and *mempunyai* can be passivized with an undergoer voice prefix *di-* as in example (2-3) and (2-6) and can be changed to imperative with an imperative suffix *-lah* as in (2-4) and (2-7), but there are some syntactic restrictions for *mempunyai*[2]. Personal pronouns and personal proper names can be the direct object in the *memiliki* construction, as in (2-2) but not in the *mempunyai* construction, as in (2-5). The possessive verb *punya* cannot be passivized, and there is no imperative construction with it, as in (2-9). It cannot take personal pronouns or personal proper names as the direct object, as in (2-8).

According to Moeljadi (2010: 45), the possessive verbs *memiliki* and *mempunyai* tend to appear in the high variety of Indonesian, while *punya* tends to appear in the low variety of Indonesian.

---

[2] The possessive verb *me-miliki* alternates with *di-miliki* in passive and with *miliki-lah* in imperative. The possessive verb *mem-punyai* is changed to *di-punyai* in passive and to *punyai-lah* in imperative. However, while all speakers accept *di-miliki* and *miliki-lah*, a few speakers do not accept *di-punyai* and some speakers do not accept *punyai-lah*.

(2-1) Dia   ingin **me-milik-i**   tas=ku.
      3SG   want   AV-milik-TR   bag=1SG
      'S/he wants to possess my bag.' (own data)

(2-2) Aku   ingin **me-milik-i**=mu,   Indah.
      1SG   want   AV-milik-TR=2SG
      'I want to possess you, Indah.' (own data)

(2-3) Enam model   jeans yang   harus **di-milik-i**
      six   type   jeans REL   must   UV-milik-TR
      'Six types of jeans which must be owned' (title of an article in *KOMPAS*.com 2009/10/29)

(2-4) **Milik-i**-lah   anak   yang   banyak.
      milik-TR-IMP child   REL   many
      'Have many children.' (*Lagu kesunyian di penghujung hari* in *KOMPAS*.com 2009/08/29)

(2-5) *Orang   gila   itu   ingin **mem-punya-i**=ku.
      person   crazy that   want   AV-punya-TR=1SG
      'That crazy person wants to possess me.' (own data)

(2-6) Hak   untuk meng-ekspresi-kan   diri   **di-punya-i**   oleh   setiap orang.
      right to   AV-expression-APP   self   UV-punya-TR by   every person
      'The right to express oneself is possesed by everyone.' (own data, based on Sneddon 2006: 208)

(2-7) ?**Punya-i**-lah   impian![3]   (2-8) *Gua pingin[4] **punya** elu.
      punya-TR-IMP   dream   1SG want   punya   2SG
      'Have a dream!' (own data)   'I want to possess you.' (own data)

(2-9) ***Punya**-lah   impian!
      punya-IMP   dream (own data)

---

[3] In Moeljadi (2010), I regarded that this imperative sentence was acceptable but later I found that some speakers do not accept such imperative sentences with *punyailah*. A survey for this variation needs to be done.
[4] The word *pingin* is the low variant of *ingin* 'want'.

2.2 Constructions with the existential verb *ada*

The existential verb *ada*, which functions as the predicate in the existential (X *ada* 'X exists', *ada* X *di* Y 'there is X in Y') and locative constructions (X *ada di* Y 'X is in Y'), also functions as the predicate in the possessive verbal predicate constructions (X *ada* Y, X *ada* Y=*nya*). These constructions (X *ada* Y, X *ada* Y=*nya*) do not have passive and imperative counterparts and cannot take personal pronouns and/or proper names as Y. Both constructions tend to appear in the low variety of Indonesian (Moeljadi 2010: 58).

(2-10) Saya **ada** uang untuk mem-beli ini.
 1SG EXIST money to AV-buy this
 'I have money to buy this.' (Alieva 1992: 15)

(2-11) Lu **ada** kunci=**nya**?
 2SG EXIST key=nya
 'Do you have the key?' (own data)

(2-12) (while looking at a photo which has been modified)
 Kok gua kagak **ada** mata=**nya** di foto ini?
 DP 1SG NEG EXIST eye=nya LOC photo this
 'Why don't I have eyes in this photo?' (own data)

2.3 Constructions with denominal verbal affixes *ber-*, *ber-...-kan*, and *-an*

The prefix *ber-*, the circumfix *ber-...-kan*, and the suffix *-an* denominalize the base so that it becomes a verbal predicate in each construction. None of these three constructions can be passivized. Alwi et al. (2000: 121) notes that the suffix *-kan* with a nominal base derives a verb which means 'to regard something as stated/expressed by the nominal base'. Thus, the suffixal part *-kan* in the X *ber*-Y-*kan* Z may have the meaning 'to regard Z as Y'; the construction, on the whole, means 'X has Z as Y' (Sneddon 1996: 110). It triggers the presence of an obligatory noun complement (Z), as in example (2-15) and (2-16). As for the X *ber*-Y construction it may take an optional noun complement which specifies Y, as in example (2-13) and (2-14). Both X *ber*-Y and X *ber*-Y-*kan* Z tend to appear in the high variety of Indonesian, while X Y-*an* tends to appear in the low variety of Indonesian (Moeljadi 2010: 77).

(2-13) Dia sudah **ber**-istri. (2-14) Dia sudah **ber**-istri orang Minang.
 3SG PERF ber-wife 3SG PERF ber-wife person
 'He already has a wife.' 'He already has a Minang person as his wife.'
 (Alwi et al. 2000: 142) (Alwi et al. 2000: 142)

(2-15) *Dia   sudah **ber**-istri-**kan**.
       3SG   PERF ber-wife-APP
       (Alwi et al. 2000: 142)

(2-16) Dia   sudah **ber**-istri-**kan** orang Minang.
       3GS   PERF ber-wife-APP person
       'He already has a Minang person as his wife.'
       (Alwi et al. 2000: 142)

Regarding the X Y-*an* construction, Sneddon (1996: 53) states that from some noun bases –*an* derives adjectives, meaning 'having many [base], containing many [base]' or 'suffering from [base]', as in example (2-17) and (2-18).

(2-17) Orang   itu   uban-**an**.
       person   that   grey.hair-an
       'That person is grey-haired.' (own data)
       (lit: 'That person has lots of grey hair.')

(2-18) Pipi   lu   jerawat-**an**.
       cheek 2SG   pimple-an
       'Your cheek is pimpled.', (own data)
       'Your cheek is covered with many pimples.'

3. Previous studies

I mainly refer to Hopper (1972), Alieva (1992), and Moeljadi (2010) for previous studies of possessive verbal predicates in Indonesian.

Hopper (1972: 137-140) states that in formal written Indonesian, *mempunyai* has come into general use which corresponds in most usages to English *have* in the sense of ' to own' or 'to possess'. The colloquial equivalent of *mempunyai*, i.e. *punya*, serves as a general equivalent of Western *have*-like[5] verbs. *Ada* is said to bear the same relation to *punya*, as *have* does to *own* in English. However, the possessive use of *ada* is not considered acceptable by all speakers.

Alieva (1992: 15-19) mentions that Malay[6] verbs such as –*punyai* 'to have' and –*miliki* 'to have, to possess' are special possessive verbs with their proper voice forms, but all of them belong to modern educated speech and are secondary in origin, i.e. they are originally not indigenous Malay words. *Punya*, which is common in everyday speech, is not primary either. Instead of clauses with a lexeme 'to have', the following two synthetic clause models –the one with pronominal clitics, and the other with verb-deriving prefixes– can be considered as the primary forms (originally Malay) and the basis of possessivity in Malay:

1. Real topic clauses with *ada* + possessee noun phrase In this clause, a possessor noun phrase in the initial position is cross-referenced by a pronominal enclitic and the possessive meaning disappears from *ada*, being expressed in the possessive noun phrase, as in (3-1).

---

[5] In Western (European) languages, possessive verbal predicate is encoded by way of a *have*-verb (to be found in the Germanic, Romance, Baltic subfamilies, and also in West and South Slavonic, Modern Greek, Albanian, and Armenian) or by a *be*-verb (Celtic, East Slavonic, also in Estonian, Latvian, Hungarian, and Finnish). Polish, Ukrainian, and Belorussian are said to represent a transitional stage from *be*-verb to *have*-verb (Stassen 2009: 8, Heine 1997: 211).

[6] Alieva (1992) does not mention that she also deals with Indonesian possessive verbs in her paper. However, I consider that the Malay verbs in Alieva (1992) are relevant to our discussion in this paper.

(3-1) Rumah   ini   **ada**   beranda=**nya**[7].

       house   this   EXIST   verandah=nya

       'This house has its verandah.' (Alieva 1992: 16)

2. Clauses with predicates expressed by *ber-* verbs (also *ber-...-kan* verbs) They are real and original devices for rendering the meaning 'to have, to possess', but in a peculiar synthetic form. The relation of possession and the possessed object are both expressed by one and the same word. For instance, in (3-2) 'to have a basis' is expressed by one word as ***ber-dasar***, while using possessive verbs or the existential verb *ada*, it is expressed by two words, i.e. ***memiliki*** *dasar*, ***mempunyai*** *dasar*, ***punya*** *dasar*, ***ada*** *dasar*, and ***ada*** *dasar=nya*.

(3-2) Pendapat=nya   tidak   **ber-**dasar.

       opinion=nya   NEG   ber-basis

       'His opinion has no basis.' (Sneddon 1996: 111)

Alieva (1992: 19) also notes that these two kinds of clauses are prevalent in texts, while the special possessive verbs, though rather differentiated in meaning and capable of voice alternation, are used only rarely. In a corpus containing samples of various texts, each of which consists of 120 sentences, there was one use of *punya* or *mempunyai*, but from five to seven instances of *ber-* verbs with possessive meaning.

    Concerning clauses with *ada*, Alieva (1992: 15-16) states that the verb *ada* 'to be' can itself express the meaning 'to have' (but not 'to own', 'to possess'). In an *ada* sentence, possession is expressed through the idea of existence, and the meaning 'to have' is secondary to the meaning 'to be, to exist'. An *Ada* sentence can be analyzed as having possessor topics in the initial position (possessors in focus) that can be omitted without the effect of ellipsis.

    In Moeljadi (2010: 34), I found that there are eight possessive verbal predicate constructions in Indonesian, which can be classified into three groups based on the form of the main verb as follows (X represents 'possessor', Y represents 'possessee' or 'possessum', and Z represents a complement):

    I. Constructions with possessive verbs: (1) X *memiliki* Y, (2) X *mempunyai* Y, (3) X *punya* Y

    II. Constructions with the existential verb *ada*: (4) X *ada* Y, (5) X *ada* Y*=nya*

    III. Constructions with denominal affixes: (6) X *ber-*Y, (7) X *ber-*Y*-kan* Z, (8) X Y*-an*

    Based on my intuition, I concluded that the register (i.e. the high and the low varieties of Indonesian), and the '(in)alienability' notion (see Table 1) play important roles in the encoding

---

[7] The enclitic *=nya* is originally derived from *ia* '3SG' and retains the original function as the third person singular pronominal enclitic.

process (Moeljadi 2010: 93-102). In addition, I hypothesized that the enclitic =*nya* in X *ada* Y=*nya* functions as an 'inalienability marker'. However, it has later become clear that the '(in)alienability' notion is merely useful to explain the difference between X *ada* Y and X *ada* Y=*nya*. This problem was discussed in Moeljadi (2011).

Table 1. Result of analysis of possessive verbal predicate constructions in Indonesian (Moeljadi 2010: 93)

Constructions	Parameters	Possessive predicate		Possessee (Y)		
		High/ Low variety	Passivization and imperative	Pers.Pron, Pers.Proper Names	Alienable	Inalienable
possessive verbs	X memiliki Y	H	+	+	+	+
	X mempunyai Y					
	X punya Y					
existential verb *ada*	X ada Y	L	−	−	−	−
	X ada Y=nya				−	−
denominal affixes	X ber-Y	H			+	+
	X ber-Y-kan Z					
	X Y-an	L			−	

4. Storytelling survey for identifying factors in speakers' choice of possessive verbal predicates in the high and low varieties

4.1 Storytelling survey

The storytelling survey was conducted in Malang in August 2011 and in Tokyo in September and October 2011. The consultants were from different backgrounds (as for age, place of birth, place of growing up, and mother tongue). All had received a high school-level or university-level education. All could be classified as belonging to the middle class socially[8]. The backgrounds of the consultants are given in Table 2.

The consultants are divided into two groups, i.e. the 'speaking' and 'writing' groups, mainly based on the free choice given to each consultant. The percentages of male and female consultants in the 'speaking' group are 23.08% and 76.92% respectively, while the percentages of male and female consultants in the 'writing' group are 54.55% and 45.45% respectively. The places of birth and the places of growing up are shown in Map 1.

---

[8] I also did this survey with two female consultants whom I regard to be the representatives of the low class. One consultant, BA, is 57 years old, born and grew up in Turen, a small town in East Java, and speaks Javanese as the mother language. Her educational background is up to junior high school-level and now she is working as a housemaid in Malang, East Java. Although I gave some explanation about the pictures beforehand, she told the story with much difficulty and I decided that her spoken data is not suitable to be included in this survey. The other consultant, DA, is 21 years old, born and grew up in Kupang, a city in East Nusa Tenggara, and speaks Kupang Malay as the mother language. She did not continue her study to junior high school after graduating from elementary school and now she is working as a housemaid in Malang, East Java. Although I gave some explanation about the pictures beforehand, she seemed very unwilling to tell the story and seemed afraid to make any mistakes. I could not get her spoken data.

Table 2. Backgrounds of the consultants in the storytelling survey

	Consultant	Sex	Age	Place of birth	Place of growing up	Mother language	Place of survey
Speaking	AL	F	40	Magelang	Magelang	Javanese	Malang
	B	F	46	Jakarta	Jakarta	CJI	Malang
	BEW	M	17	Balikpapan	Balikpapan	CI, Mandarin	Malang
	CWS	F	18	Jayapura	Jayapura	Papuan Malay	Malang
	EW	F	50	Malang	Malang	Peranakan Javanese	Malang
	JC	F	27	Pontianak	Pontianak	Teochew	Malang
	KA	M	54	Tabanan	Tabanan	Balinese	Tokyo
	NDL	F	13	Jakarta	Malang	CI, Javanese mixed	Malang
	P	F	17	Pontianak	Jakarta	CJI	Malang
	PL	F	16	Jayapura	Jayapura	Papuan Malay	Malang
	PN	F	19	Malang	Malang	Peranakan Javanese	Malang
	SS	F	35	Jakarta	Jakarta	CJI	Malang
	YOK	M	32	Yogyakarta	Jakarta	CJI	Malang
Writing	AM	F	24	Jakarta	Jakarta	CJI	Tokyo
	AP	F	25	Jakarta	Jakarta	CI	Tokyo
	AW	M	31	Palembang	Palembang	Palembang Malay	Tokyo
	EIS	F	41	Malang	Malang	Javanese	Malang
	J	M	22	Jakarta	Jakarta	Hokkien, CJI	Tokyo
	JSP	F	16	Makassar	Jayapura	Papuan Malay	Malang
	JSS	M	25	Malang	Malang	Peranakan Javanese	Malang
	R	M	36	Long Iram	Long Iram	CI, Banjar	Tokyo
	RAS	M	25	Solo	Bekasi	CI	Tokyo
	SKR	F	15	Tangerang	Tangerang	CI, Javanese mixed	Malang
	TTAW	M	16	Sumba Barat	Sumba Barat	Sumba (Kambera)	Malang

Map 1. The location of the places of birth and the places of growing up of the consultants in the 'storytelling' survey (Source: http://www.indonesiamatters.com/86/indonesian-provinces-map/, accessed in November 2011) (1: Palembang, 2: Tangerang, 3: Jakarta, 4: Bekasi, 5: Magelang, 6: Yogyakarta, 7: Solo, 8: Malang, 9: Tabanan, 10: Sumba Barat, 11: Pontianak, 12: Long Iram, 13: Balikpapan, 14: Makassar, 15: Jayapura)

The place of growing up in Table 2 refers to the place where the consultant was living during the upper grades of the elementary school, i.e. about ten to twelve years old. 'Mother language' refers

to the language used for communication at home, particularly for communicating with parents. The language abbreviated to CJI is Colloquial Jakartan Indonesian and the one abbreviated to CI is Colloquial Indonesian. I have to admit here that the difference I made in Table 2 is rather obscure, based on the answer given by the consultant. If, towards the question *bahasa yang dipake sehari-hari di rumah kalo ngomong sama ayah ibu apa?* 'what language do you use in everyday communication at home, especially when you talk with your father and mother?', the consultant answers like *bahasa Indonesia, tapi logat Jakarta* 'Indonesian, but in Jakartan dialect', I regard Colloquial Jakartan Indonesian (CJI) as the mother language of the consultant. If the consultant answers like *bahasa Indonesia* 'Indonesian', then I regard Colloquial Indonesian (CI) as his/her mother language. In brief, the difference between CJI and CI in this survey depends on the awareness of each consultant about whether s/he speaks Jakartan Indonesian or Indonesian.

A story which contains many possessive verbal predicates is prepared to gather the corpus (spoken and written data) in this survey. The story script is presented in Table 3. During the process of making a part of the story I referred to the story script of *Frog, Where Are You?* by Mercer Mayer on page 5 of the Narrative Story Retell Reference Database.

Table 3. Story script for the storytelling survey

Picture	Script
1	There once were two boys who were friends, named A and B.
2	A was a rich boy. He had a very big house with a swimming pool inside and a limousine with a chauffeur. He also had a lot of expensive clothes and money. He wore glasses and had pimples on his cheeks.
3	A was the only child. He had neither brothers and sisters, nor grandfather and grandmother, but he had a father and a mother who loved him very much. He also had many friends.
4	On the other hand, B was a poor boy. His house was made of bamboos with a zinc roof. His bicycle was very old and rusty. He did not have a lot of money and his clothes were very few.
5	B was the only child, too. He had neither brothers and sisters, nor father and mother, but he still had a grandmother who suffered from bronchitis. He also had a dog and a frog as pets and many friends.
6	His pet frog was very unique. It had two big eyes and a long tail. He kept the frog in a jar.
7	One day, when A and B were going to the same school, they met and greeted each other.
8	In the classroom, they sat side by side. B asked A, "A, do you have a red pencil?"
9	A answered, "Yes, I do. Here, you can borrow it."
10	A said to B, "B, what's that inside your bag?" B answered, "My pet frog. I kept it in a jar."
11	Then B showed his pet frog in a jar to A. "Look! My pet frog", said B. "I like your frog!" said A.
12	A said, "I want to buy your frog. Here are some money to buy it.", but B said, "No, I don't want to sell it. I won't sell it to anyone. Even plenty money cannot buy my frog."
13	A became angry and said, "This frog is mine. You may not possess it." "No, it's mine. You can't have it.", B said angrily. They were fighting and scrambling for the frog. The teacher warned them, "A and B, don't fight in the classroom!"
14	Suddenly the bottle cap was opened and the frog jumped out of the jar.
15	The frog landed on the teacher's hair and jumped outside the classroom, holding the teacher's hair. It turned out that the teacher wore a wig.

16	Realizing the teacher was actually bald and wore a wig, all the students in the class laughed very loudly. The teacher was angry at A and B.
17	After school, it was raining heavily but B did not go home. He looked everywhere for the frog. He called out, "Frog, where are you?"
18	At night, he arrived at a pond with many frogs sitting on lotus leaves. There he found his pet frog with a long tail.
19	He went into the pond and put his pet frog into the jar.
20	Sopping wet, he went home. His grandmother and pet dog were waiting for him outside the house.
21	In the middle of the night, he had a fever, caught a cold, and was sweating. His grandmother covered up his body with a blanket.
22	The next day, there was a mathematics exam. B was absent and A felt guilty and regretted what he had done to B the day before.
23	After school, A went to B's house with his parents. He apologized to him and said, "I'm really sorry for what I did yesterday. By the way, I have a really good personal doctor. Do you want me to call him for you?"
24	A phoned his personal doctor, "Hello, Mr. Doctor, I have a friend who has a fever. If you have time, please come to his place soon." The doctor replied, "OK! I'll be there soon."
25	The doctor soon came to B's house and examined his fever and cold. B was still sweating very much. The doctor had a heavy beard and a tooth with a gold crown.
26	The doctor then gave a cure-all/panacea with great efficacy to cure B's illness.
27	B drank the cure-all and became healthy. B's grandmother who was suffering from bronchitis, also drank the cure-all and became healthy.
28	A and B became friends again.

Twenty-eight pictures are drawn based on the story (some of them can be seen in Appendix). Various words expressing possessees are written on the pictures. Some symbols, such as ○, ×, ?, =, ≠, and ⇒, are also added to the picture.

Before showing those pictures to the consultants, I told them that this survey was a 'storytelling' survey in Indonesian (*penelitian mengenai bercerita dalam bahasa Indonesia*). Then I showed those twenty-eight pictures one by one to the consultants and added some explanation, such as instructions to use all the words written on the pictures in the story if possible, and explanation about the symbols (for instance, ○ means 'yes', 'positive', or 'okay', × means 'no', 'negative', or 'not okay', etc.). After that, I asked them to tell a story based on those pictures in their own daily Indonesian. I also emphasized that there was no need to use 'good and correct' Indonesian (*tolong ceritakan dalam bahasa Indonesia sehari-hari, nggak perlu pakai bahasa Indonesia yang 'baik dan benar'*).

I left to the consultants a choice between making her/his story by 'speaking' or 'writing'. Looking at the pictures, the consultants who chose 'writing' wrote the story with a word processor, while the consultants who chose 'speaking' told the story, which were recorded.

The spoken data consist of thirteen monologues recorded in Malang and Tokyo between August and October 2011. They were recorded in the consultants' houses, a place of work (an office), or in public places, such as a school canteen and a church. At all recordings, the consultants were first requested to consent to be recorded. The spoken data, containing 14,082 words in total, were

relatively spontaneous monologues mostly with considerable background noise. Fillers, ellipsis, slips of the tongue, repetition, and obscuring of sentence boundaries are all common features of the spoken data.

The written data consist of eleven texts typed by the consultants in Malang and Tokyo between August and October 2011. The data contain 8,225 words. Unlike the spoken data, sentences in the written data are generally well-formed.

4.2   Quantitative study of variables

Sneddon (2006: 10) states that by a quantitative study of variables we can show the variation in the speech of individuals and by aggregating percentages we can make a statement about general tendencies in the community's linguistic behaviour.

I assume that more features of the low variety appear in the spoken data than in the written data and I therefore consider it of great worth to count tokens of each variable of the low variety in the spoken and written data, respectively. Since the low variety also tend to be used in informal speech, it is also necessary to count tokens of each variable of the low variety in the indirect and the direct narration of each spoken and written data.

The criteria for deciding the low variants are mainly based on Sneddon (2006: 15-136). Various linguistic variables were chosen to determine the degree of lowness of each indirect and direct narration in the spoken and written data, as follows:

I. Phonological features

1. Word-final diphthongs in the high variety *ai* /ay/ and *au* /aw/ are pronounced with single vowels *e* /e/ and *o* /o/ in the low variety respectively. For instance, a conjunction *kalau* which introduces a subordinate clause in the high variety, is usually pronounced *kalo* in the low variety (Sneddon 2006: 15, 80-82).

2. The vowel *a* /a/ in a final closed syllable of some words in the high variety is usually pronounced with the mid-central vowel schwa /ə/ in the low variety (Sneddon 2006: 17-18).

3. Some words with the initial *s* in the high variety are favourably pronounced without *s* in the low variety. For instance, a perfective marker *sudah* in the high variety is usually pronounced *udah* in the low variety (Sneddon 2006: 18-20, 50-53, 55).

II. Morphological features

4. The prefix *meN-*[9] indicating the active voice in the high variety may be completely lost, leaving just the base; or partially lost, leaving only *N-*; or may be replaced by the prefix *nge-* (Sneddon 2006: 20-24).

5. The prefix *ter-* indicating uncontrolled events in the high variety is usually replaced by *ke-* in the low variety (Sneddon 2006: 25-27).

---

[9] *N* symbolizes a nasal which shares the point of articulation with the first sound of the verb base.

6. The suffixes *-kan* and *-i* are usually replaced by *-in* in the low variety (Sneddon 2006: 30-34).

III. Lexical features

7. The negative marker *tidak*, used with non-nominal predicates in the high variety, is usually replaced by *enggak~nggak~gak~ga~ndak*[10] in the low variety (Sneddon 2006: 56-57).

8. The word *buat*, which originally means 'to make', is often used to replace the preposition *untuk* 'to/for', indicating purpose or the recipient in the high variety. *Kamus besar bahasa Indonesia* (a comprehensive dictionary of Indonesian) notes this word as a colloquial preposition.

9. A single preposition *sama~ama* in the low variety usually corresponds to more than one preposition in the high variety, such as *dengan* 'with' indicating accompaniment or reciprocal relationship, *terhadap* 'towards', *kepada* 'to' marking the recipient, and *oleh* 'by' indicating the agent in a passive construction (Sneddon 2006: 50-53).

10. The first person singular pronoun *gua~gue*, which is derived from Hokkien Chinese, is typically associated with youth and very informal situations (Sneddon 2006: 59-60).

11. The second person singular pronoun *lu*, like *gua~gue*, is derived from Hokkien Chinese and typically associated with the colloquial Indonesian of young people (Sneddon 2006: 64-66).

12. The word *apaan*, which is derived from *apa* 'what' and means 'what kind of', is frequently used in the low variety. *Kamus besar bahasa Indonesia* (a comprehensive dictionary of Indonesian) notes this word as a colloquial word.

13. The word *banget* 'very' is the most common word in the low variety. Its counterparts in the high variety are *amat*, *sangat*, and *sekali*. Young people use *banget* very frequently in conversations (Sneddon 2006: 83-84).

14. In the low variety, *bilang* 'to say' is one of the most frequently occurring verbs. It is strongly associated with informal speech and rarely occurs in purely formal language. Its counterparts in the high variety are *berkata* and *mengatakan* (Sneddon 2006: 101-102).

15. The word *btw*, pronounced /be.te.we/, originally stands for 'by the way' and still maintains its original meaning, is frequently used in the low variety.

16. The word *bikin* 'to make' is usually used in the low variety, instead of *membuat* which is used in the high variety (Sneddon 2003: 532; 2006: 17).

17. The word *cowok* 'male, man, guy' is highly marked as the low variant. Its counterparts are *laki-laki* and *pria* (Sneddon 2006: 105-106).

18. The word *curhat* stands for *curahan hati*, which means 'to have a heart-to-heart talk'. It

---

[10] In this section the following convention is employed: x~y indicates that x and y are different pronunciations of a single form. For instance, *nggak~gak* 'NEG' indicates free variation between *enggak* and *gak*.

is used in the low variety. In the high variety, *mencurahkan isi hati* is usually used instead.

19. The word *doang* 'only, just' sometimes occurs in the low variety and is very strongly marked for informality. While this word does not occur often, in the usage of younger speakers it is nevertheless much more frequent than *saja*, its counterpart in the high variety (Sneddon 2006: 82).

20. The common word for 'money' in the low variety is *duit*, while in the high variety *uang* is frequently used. *Kamus besar bahasa Indonesia* (a comprehensive dictionary of Indonesian) notes this word as a colloquial word.

21. The common word for 'big' in the low variety, especially among young people, is *gede*, while in the high variety *besar.* (Sneddon 2006: 104-105)

22. The word *bagaimana* 'how?' in the high variety is frequently contracted to *gimana* in the low variety. (Sneddon 2006: 90-91)

23. The word *kali* is short for *barangkali* 'perhaps, possibly'. *Barangkali* is the high variant, while *kali* is the low variety counterpart, strictly confined to informal speech (Sneddon 2006: 96).

24. In the high variety, *memberi* is the general word for 'to give'. In the low variety the more common word is *kasi~kasih* (Sneddon 2006: 100-101).

25. The word *kayak* 'like, resembling' occurs frequently in the low variety. Its high variant is *seperti*, being the only one to occur in the high variety (Sneddon 2006: 98-99).

26. The word *kenapa* 'why?, what's the matter with?' is often used in the low variety. In the high variety, *mengapa* is usually used (Sneddon 2006: 92).

27. The word *makanya* 'that's why, because of that' is usually used in the low variety, while in the high variety *maka dari itu* or *oleh sebab itu* is used instead. *Kamus besar bahasa Indonesia* (a comprehensive dictionary of Indonesian) notes this word as a colloquial word.

28. The word *terima kasih* 'thank you' in the high variety is sometimes contracted to *makasih* in the low variety.

29. Besides *btw*, mentioned in 14, *ngomong-ngomong* 'by the way' is also frequently used in the low variety (Sneddon 2006: 104).

30. The word *pada* 'altogether' is highly informal. It occurs in *Kamus besar bahasa Indonesia* (a comprehensive dictionary of Indonesian), being identified as conversational/colloquial (Sneddon 2006: 84-86). Its counterpart in the high variety is *semuanya* or *sama-sama*.

31. The word *pas* 'when (of past events)' is confined to informal language, while the word *waktu* is neutral as to formality and *ketika* is confined to formal language (Sneddon 2006: 92-94).

32. The word *soalnya* 'because' is often used in the low variety, while its counterpart,

karena is usually used in the high variety.

33. The word *sori*, originally from 'sorry', is frequently used in the low variety. Its counterpart in the high variety is *maaf*.

IV. Pragmatic features

34. Discourse particles or discourse markers: *deh, dong, kan, kok, lho~loh~lo, nah~na, nih~ni, tuh~tu, sih,* and *yah~ya* are very frequent in the low variety, especially in the discourse (Sneddon 2006: 117-131).

35. The adverbs of manner *begini* 'like this' and *begitu* 'like that' in the high variety are usually shortened to *gini* and *gitu* in the low variety, respectively. These shortened forms are very commonly used in the low variety as pragmatic devices similar to discourse particles (Sneddon 2006: 132-136).

V. Others

In determining the degree of lowness in the spoken data, I also include some spoken variants, as follows:

36. Fillers, such as *e:, o:,* and *a:*.

37. Filler *apa* 'what?' and *apa namanya* 'what's the name?'.

38. Click, written as *ck* in the transcripts.

39. Repetition.

40. Slip of the tongue.

4.3   Results and discussion

Referring to the forty variables mentioned above, the total number of the low variety's tokens and their percentages in the indirect and direct narration of the spoken and written data of each consultant were counted. For example, in the spoken data of consultant AL the total number of the low variety's tokens in the indirect narration is 6 and the total number of the word tokens in the

corresponding indirect narration is 796, hence the percentage is $\frac{6}{796} \times 100\% = 0.75\%$ (see Table 4).

In this way the percentages of the low variety's tokens, as well as the percentages of the eight possessive verbal predicates, in the indirect and direct narration of the spoken and written data of each consultant are counted and presented in Table 4 and Table 5. Table 4 and Table 5 show us that the spoken data in total has more percentage of the low variety's tokens (11.01%) than the written data in total (2.10%) and the direct narration has more percentage of the low variety's tokens (12.75% for spoken data and 18.54% for written data) than the indirect narration (10.70% for spoken data and 0.20% for written data). Hence we may presume that more features of the low variety appear in the spoken data than in the written data, as well as in the direct narration than in the indirect narration. Figure 1 and Figure 2 show us that there is more gap between the indirect and direct narration in the written data than in the spoken data.

Table 4. The number of possessive verbal predicates' tokens and their percentages in the spoken data (L: Low variety, In: Indirect narration, Dn: Direct narration)

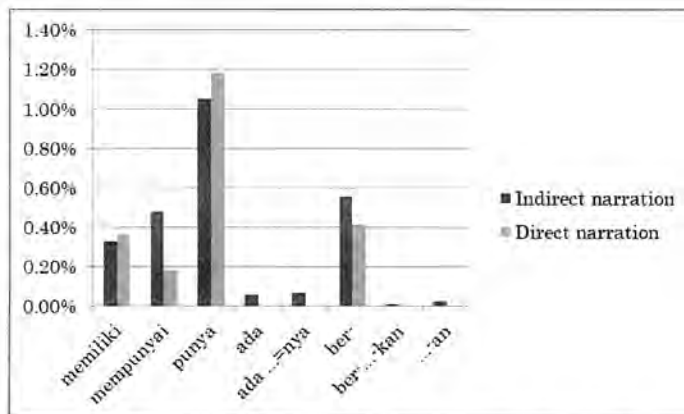| Consultant	Word tokens		L tokens		memiliki		mempunyai		punya		ada		ada ...-nya		ber-		ber-...-kan		-an	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
AL (In)	796	87.96%	6	0.75%	6	0.75%	1	0.13%	3	0.38%	2	0.25%	1	0.13%	6	0.75%	0	0.00%	0	0.00%
AL (Dn)	109	12.04%	1	0.92%	0	0.00%	1	0.92%	1	0.92%	0	0.00%	0	0.00%	1	0.92%	0	0.00%	0	0.00%
AL (In+Dn)	905	100.00%	7	0.77%	6	0.66%	2	0.22%	4	0.44%	2	0.22%	1	0.11%	7	0.77%	0	0.00%	0	0.00%
B (In)	1049	67.90%	41	3.91%	5	0.48%	3	0.29%	5	0.48%	0	0.00%	0	0.00%	3	0.29%	0	0.00%	0	0.00%
B (Dn)	496	32.10%	94	18.95%	1	0.20%	0	0.00%	3	0.60%	0	0.00%	0	0.00%	5	1.01%	0	0.00%	0	0.00%
B (In+Dn)	1545	100.00%	135	8.74%	6	0.39%	3	0.19%	8	0.52%	0	0.00%	0	0.00%	8	0.52%	0	0.00%	0	0.00%
BEW (In)	719	88.66%	45	6.26%	14	1.95%	3	0.42%	3	0.42%	0	0.00%	0	0.00%	5	0.70%	1	0.14%	0	0.00%
BEW (Dn)	92	11.34%	6	6.52%	4	4.35%	1	1.09%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%
BEW (In+Dn)	811	100.00%	51	6.29%	18	2.22%	4	0.49%	3	0.37%	0	0.00%	0	0.00%	5	0.62%	1	0.12%	0	0.00%
CWS (In)	672	99.70%	41	6.10%	1	0.15%	10	1.49%	5	0.74%	0	0.00%	1	0.15%	4	0.60%	0	0.00%	0	0.00%
CWS (Dn)	2	0.30%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%
CWS (In+Dn)	674	100.00%	41	6.08%	1	0.15%	10	1.48%	5	0.74%	0	0.00%	1	0.15%	4	0.59%	0	0.00%	0	0.00%
EW (In)	1654	87.89%	254	15.36%	0	0.00%	1	0.06%	27	1.63%	0	0.00%	3	0.18%	5	0.30%	0	0.00%	1	0.06%
EW (Dn)	228	12.11%	47	20.61%	1	0.44%	0	0.00%	7	3.07%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%
EW (In+Dn)	1882	100.00%	301	15.99%	1	0.05%	1	0.05%	34	1.81%	0	0.00%	3	0.16%	5	0.27%	0	0.00%	1	0.05%
JC (In)	1785	86.57%	83	4.65%	1	0.06%	0	0.00%	26	1.46%	0	0.00%	0	0.00%	2	0.11%	0	0.00%	0	0.00%
JC (Dn)	277	13.43%	12	4.33%	0	0.00%	0	0.00%	1	0.36%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%
JC (In+Dn)	2062	100.00%	95	4.61%	1	0.05%	0	0.00%	27	1.31%	0	0.00%	0	0.00%	2	0.10%	0	0.00%	0	0.00%
KA (In)	684	89.30%	55	8.04%	0	0.00%	3	0.44%	13	1.90%	0	0.00%	0	0.00%	14	2.05%	0	0.00%	0	0.00%
KA (Dn)	82	10.70%	6	7.32%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%
KA (In+Dn)	766	100.00%	61	7.96%	0	0.00%	3	0.39%	13	1.70%	0	0.00%	0	0.00%	14	1.83%	0	0.00%	0	0.00%
NDL (In)	408	100.00%	36	8.82%	0	0.00%	9	2.21%	1	0.25%	0	0.00%	0	0.00%	3	0.74%	0	0.00%	0	0.00%
NDL (Dn)	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%
NDL (In+Dn)	408	100.00%	36	8.82%	0	0.00%	9	2.21%	1	0.25%	0	0.00%	0	0.00%	3	0.74%	0	0.00%	0	0.00%
P (In)	756	92.31%	244	32.28%	0	0.00%	0	0.00%	22	2.91%	0	0.00%	0	0.00%	2	0.26%	0	0.00%	1	0.13%
P (Dn)	63	7.69%	13	20.63%	0	0.00%	0	0.00%	3	4.76%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%
P (In+Dn)	819	100.00%	257	31.38%	0	0.00%	0	0.00%	25	3.05%	0	0.00%	0	0.00%	2	0.24%	0	0.00%	1	0.12%
PL (In)	661	93.76%	69	10.44%	9	1.36%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	5	0.76%	0	0.00%	0	0.00%
PL (Dn)	44	6.24%	3	6.82%	0	0.00%	0	0.00%	1	2.27%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%
PL (In+Dn)	705	100.00%	72	10.21%	9	1.28%	0	0.00%	1	0.14%	0	0.00%	0	0.00%	5	0.71%	0	0.00%	0	0.00%
PN (In)	705	89.02%	60	8.51%	3	0.43%	10	1.42%	2	0.28%	0	0.00%	1	0.14%	5	0.71%	0	0.00%	0	0.00%
PN (Dn)	87	10.98%	9	10.34%	2	2.30%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%
PN (In+Dn)	792	100.00%	69	8.71%	5	0.63%	10	1.26%	2	0.25%	0	0.00%	1	0.13%	5	0.63%	0	0.00%	0	0.00%
SS (In)	1039	70.87%	213	20.50%	0	0.00%	0	0.00%	12	1.15%	3	0.29%	0	0.00%	6	0.58%	0	0.00%	1	0.10%
SS (Dn)	427	29.13%	74	17.33%	0	0.00%	0	0.00%	7	1.64%	0	0.00%	0	0.00%	3	0.70%	0	0.00%	0	0.00%
SS (In+Dn)	1466	100.00%	287	19.58%	0	0.00%	0	0.00%	19	1.30%	3	0.20%	0	0.00%	9	0.61%	0	0.00%	1	0.07%
YOK (In)	965	77.39%	125	12.95%	0	0.00%	17	1.76%	6	0.62%	2	0.21%	2	0.21%	6	0.62%	0	0.00%	0	0.00%
YOK (Dn)	282	22.61%	14	4.96%	0	0.00%	2	0.71%	3	1.06%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%
YOK (In+Dn)	1247	100.00%	139	11.15%	0	0.00%	19	1.52%	9	0.72%	2	0.16%	2	0.16%	6	0.48%	0	0.00%	0	0.00%
Total (In)	11893	84.46%	1272	10.70%	39	0.33%	57	0.48%	125	1.05%	7	0.06%	8	0.07%	66	0.55%	1	0.01%	3	0.03%
Total (Dn)	2189	15.54%	279	12.75%	8	0.37%	4	0.18%	26	1.19%	0	0.00%	0	0.00%	9	0.41%	0	0.00%	0	0.00%
Total (In+Dn)	14082	100.00%	1551	11.01%	47	0.33%	61	0.43%	151	1.07%	7	0.05%	8	0.06%	75	0.53%	1	0.01%	3	0.02%

Figure 1. The percentages of eight possessive verbal predicates in the indirect and direct narrations of the spoken data

Table 5. The number of possessive verbal predicates' tokens and their percentages in the written data

| Consultant	Word tokens		L tokens		memiliki		mempunyai		punya		ada		ada ...-nya		ber-		ber-...-kan		-an	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
AM (In)	523	76.91%	2	0.38%	8	1.53%	1	0.19%	1	0.19%	0	0.00%	0	0.00%	5	0.96%	0	0.00%	0	0.00%
AM (Dn)	157	23.09%	55	35.03%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%
AM (In-Dn)	680	100.00%	57	8.38%	8	1.18%	1	0.15%	1	0.15%	0	0.00%	0	0.00%	5	0.74%	0	0.00%	0	0.00%
AP (In)	1011	93.35%	0	0.00%	9	0.89%	0	0.00%	1	0.10%	0	0.00%	0	0.00%	11	1.09%	0	0.00%	0	0.00%
AP (Dn)	72	6.65%	0	0.00%	2	2.78%	0	0.00%	1	1.39%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%
AP (In-Dn)	1083	100.00%	0	0.00%	11	1.02%	0	0.00%	2	0.18%	0	0.00%	0	0.00%	11	1.02%	0	0.00%	0	0.00%
AW (In)	953	80.63%	2	0.21%	9	0.94%	0	0.00%	0	0.00%	1	0.10%	0	0.00%	8	0.84%	0	0.00%	0	0.00%
AW (Dn)	229	19.37%	35	15.28%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%
AW (In-Dn)	1182	100.00%	37	3.13%	9	0.76%	0	0.00%	0	0.00%	1	0.08%	0	0.00%	8	0.68%	0	0.00%	0	0.00%
EIS (In)	937	100.00%	0	0.00%	11	1.17%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	16	1.71%	1	0.11%	0	0.00%
EIS (Dn)	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%
EIS (In-Dn)	937	100.00%	0	0.00%	11	1.17%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	16	1.71%	1	0.11%	0	0.00%
J (In)	471	78.24%	10	2.12%	3	0.64%	0	0.00%	7	1.49%	0	0.00%	0	0.00%	4	0.85%	1	0.21%	1	0.21%
J (Dn)	131	21.76%	45	34.35%	0	0.00%	0	0.00%	1	0.76%	2	1.53%	0	0.00%	0	0.00%	0	0.00%	0	0.00%
J (In-Dn)	602	100.00%	55	9.14%	3	0.50%	0	0.00%	8	1.33%	2	0.33%	0	0.00%	4	0.66%	1	0.17%	1	0.17%
JSP (In)	509	91.38%	0	0.00%	11	2.16%	0	0.00%	1	0.20%	0	0.00%	0	0.00%	4	0.79%	1	0.20%	0	0.00%
JSP (Dn)	48	8.62%	1	2.08%	0	0.00%	1	2.08%	1	2.08%	0	0.00%	0	0.00%	1	2.08%	0	0.00%	0	0.00%
JSP (In-Dn)	557	100.00%	1	0.18%	11	1.97%	1	0.18%	2	0.36%	0	0.00%	0	0.00%	5	0.90%	1	0.18%	0	0.00%
JSS (In)	581	81.37%	0	0.00%	8	1.38%	0	0.00%	4	0.69%	0	0.00%	0	0.00%	4	0.69%	0	0.00%	0	0.00%
JSS (Dn)	133	18.63%	20	15.04%	0	0.00%	0	0.00%	1	0.75%	1	0.75%	0	0.00%	0	0.00%	0	0.00%	0	0.00%
JSS (In-Dn)	714	100.00%	20	2.80%	8	1.12%	0	0.00%	5	0.70%	1	0.14%	0	0.00%	4	0.56%	0	0.00%	0	0.00%
R (In)	769	92.54%	0	0.00%	8	1.04%	0	0.00%	3	0.39%	0	0.00%	0	0.00%	9	1.17%	0	0.00%	0	0.00%
R (Dn)	62	7.46%	2	3.23%	1	1.61%	0	0.00%	2	3.23%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%
R (In-Dn)	831	100.00%	2	0.24%	9	1.08%	0	0.00%	5	0.60%	0	0.00%	0	0.00%	9	1.08%	0	0.00%	0	0.00%
RAS (In)	557	96.53%	0	0.00%	4	0.72%	0	0.00%	4	0.72%	1	0.18%	0	0.00%	5	0.90%	0	0.00%	1	0.18%
RAS (Dn)	20	3.47%	0	0.00%	0	0.00%	0	0.00%	1	5.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%
RAS (In-Dn)	577	100.00%	0	0.00%	4	0.69%	0	0.00%	5	0.87%	1	0.17%	0	0.00%	5	0.87%	0	0.00%	1	0.17%
SKR (In)	501	100.00%	0	0.00%	13	2.59%	0	0.00%	1	0.20%	0	0.00%	0	0.00%	3	0.60%	1	0.20%	0	0.00%
SKR (Dn)	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%
SKR (In-Dn)	501	100.00%	0	0.00%	13	2.59%	0	0.00%	1	0.20%	0	0.00%	0	0.00%	3	0.60%	1	0.20%	0	0.00%
TTAW (In)	561	100.00%	1	0.18%	10	1.78%	1	0.18%	0	0.00%	0	0.00%	0	0.00%	3	0.53%	0	0.00%	0	0.00%
TTAW (Dn)	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%
TTAW (In-Dn)	561	100.00%	1	0.18%	10	1.78%	1	0.18%	0	0.00%	0	0.00%	0	0.00%	3	0.53%	0	0.00%	0	0.00%
Total (In)	7373	89.64%	15	0.20%	94	1.27%	2	0.03%	22	0.30%	2	0.03%	0	0.00%	72	0.98%	4	0.05%	2	0.03%
Total (Dn)	852	10.36%	158	18.54%	3	0.35%	1	0.12%	7	0.82%	3	0.35%	0	0.00%	1	0.12%	0	0.00%	0	0.00%
Total (In-Dn)	8225	100.00%	173	2.10%	97	1.18%	3	0.04%	29	0.35%	5	0.06%	0	0.00%	73	0.89%	4	0.05%	2	0.02%
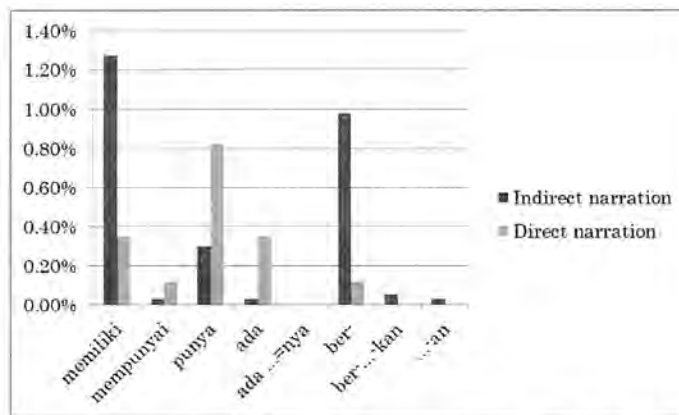
Figure 2. The percentages of eight possessive verbal predicates in the indirect and direct narrations of the written data

In the indirect narration of the written data, *memiliki* occurs with a high frequency (1.27%, 94 tokens out of 7373), followed by *ber-* (0.98%, 72 tokens out of 7373). However, in the direct

narration of the written data, *punya* occurs most frequently (0.82%, 7 tokens out of 852) among the possessive verbs. Actually, *punya* occurs with the highest frequency, both in the indirect (1.05%, 125 tokens out of 11,893) and the direct narration (1.19%, 26 tokens out of 2,189) of the spoken data. Thus, we may say that *punya* occurs more frequently in the spoken than in the written data, and more frequently in the direct than in the indirect narration. It seems to suggest that the choice of *punya* as the possessive verbal predicate would correlate with the lowness of the variety in which it occurs.

In order to verify such correlation between the speakers' choice of a specific possessive verbal predicate and the degree of lowness, the relation needs to be examined statistically. In this case, we have to check how and to what extent the number of the low variety's tokens and the number of each possessive verbal predicate's tokens are correlated. The degree of correlation between these two variables can be measured using the Pearson's correlation coefficient, the value of which ranges from -1 to +1 (Young 1966: 308-317). Positive correlation coefficients (approach +1) indicate that both variables increase or decrease together, whereas negative correlation coefficients (approach -1) indicate that as one variable increases, the other decreases, and vice versa. If coefficients approach zero, correlation gets weaker. The closer the coefficient is to either -1 or +1, the stronger the correlation between the variables.

Specifically the correlation between the total number of the low variety's tokens and the total number of each possessive predicate's token can be measured with the following steps. For example, if we want to measure the correlation coefficient between *memiliki* and the degree of lowness in the written data, first we make some calculations as in Table 6. In Table 6, the first column contains the consultants, the second column the respective L tokens in the written data (X), the third column the corresponding possessive predicate *memiliki* tokens (Y), the fourth column represents the XY values, the fifth column contains the X^2 values, and the sixth column the Y^2 values. The sums of the values in each column are next determined.

Table 6. Calculation of the Pearson's correlation coefficient between the low variety's tokens and the possessive verbal predicate *memiliki* in the written data

Consultant	L tokens (X)	*memiliki* tokens (Y)	XY	X^2	Y^2
AM	57	8	456	3249	64
AP	0	11	0	0	121
AW	37	9	333	1369	81
EIS	0	11	0	0	121
J	55	3	165	3025	9
JSP	1	11	11	1	121
JSS	20	8	160	400	64
R	2	9	18	4	81
RAS	0	4	0	0	16
SKR	0	13	0	0	169
TTAW	1	10	10	1	100
Totals (Σ)	173	97	1,153	8,049	947

Subsequently, proper substitutions are made in the formula below (Young 1966: 316) and the correlation coefficient calculated. In this formula, N represents the number of pairs of X and Y.

$$r = \frac{N\sum XY - \left(\sum X\right)\left(\sum Y\right)}{\sqrt{\left[N\sum X^2 - \left(\sum X\right)^2\right]\left[N\sum Y^2 - \left(\sum Y\right)^2\right]}}$$

$$= \frac{11(1,153) - (173)(97)}{\sqrt{\left[11(8,049) - (173)^2\right]\left[11(947) - (97)^2\right]}}$$

$$= \frac{12,683 - 16,781}{\sqrt{(88,539 - 29,929)(10,417 - 9,409)}}$$

$$= \frac{-4,098}{\sqrt{59,078,880}}$$

$$= \frac{-4,098}{7,686.278683472}$$

$$= -0.533157874 \approx -0.53$$

We get the correlation coefficient, $r = -0.53$. The same way of calculation can be applied to the other possessive verbal predicates in the corresponding spoken, written, and both data. The result can be seen in Table 7.

Table 7. The Pearson's correlation coefficients between the low variety's tokens and each possessive verbal predicate's tokens in the spoken, written, and both data

Possessive verbs	Both	Spoken	Written
memiliki	-0.57	-0.38	-0.53
ber-...-kan	-0.29	-0.21	-0.06
ber-	-0.12	-0.03	-0.26
mempunyai	0.06	-0.31	0.11
ada	0.25	0.25	0.56
ada ...=nya	0.46	0.30	(absent)
-an	0.63	0.93	0.25
punya	0.82	0.78	0.25

We may interpret that *punya* and *-an* have strong positive correlations with the low variety's tokens and thus are frequently used in the low variety, although in the written data the degree of correlation is much lower than in the spoken data. The possessive verb *memiliki* has a quite strong negative correlation with the low variety's tokens and are thus considered to be used in the high variety, although in the spoken data correlation is less significant than in the written data. The existential verb *ada* has a quite strong positive correlation in the written data, but not in the spoken and both data, thus we may presume that *ada* has no significant correlation with the register. However, since *ada* is used more in the spoken data and in the direct narration, than in the indirect narration of the

written data, we may assume that it tends to be present more in the low variety (see Table 4 and Table 5). Lastly, because the absolute values of the coefficients of the other possessive verbal predicates, i.e. *mempunyai*, *ber-*, *ber-...-kan*, and *ada ...=nya*, are lower than those of the above-mentioned predicates, we may thus presume that those possessive verbal predicates have no significant correlations with either the low or the high variety.[11] Considering the data in Table 4 and Table 5, we know that the expression *ada ...=nya* is present only in the spoken data and we may assume that it tends to occur more in the low variety.

Based on Table 4, Table 5, and Table 7, we may summarize our findings with the results of previous studies in Table 8.

Table 8. The summary of the relationship between possessive verbal predicates and the register

Possessive verbal predicate	Register			
	Hopper (1972)	Alieva (1992)	Moeljadi (2010)	2011 Survey
memiliki		High	High	High
mempunyai	High	High	High	High, Low
punya	Low	Low	Low	Low
ada	-	-	Low	(High,) Low
ada ...=nya		-	Low	(High,) Low
ber-		-	High	High, Low
ber-...-kan		-	High	High, Low
-an			Low	Low

5. Summary of results

As discussed above, we may conclude that the possessive verb *memiliki* tends to be mainly used in the high variety, while the possessive verb *punya* tends to be very frequently used in the low variety. The suffix *-an* tends to appear in the low variety. The prefix *ber-*, *ber-...-kan*, and the possessive verb *mempunyai* tend to occur both in the high and low varieties. The existential verb *ada* and *ada ...=nya*, though tend to occur in both varieties, are used more in the low variety.

Abbreviations: 1, 2, 3 : First, second, and third person, APP : Applicative suffix, AV : Actor voice prefix, CL : Classifier, DP : discourse particle, EXIST : Existential verb, IMP : Imperative, LOC : Locative preposition, NEG : Negative, PERF : Perfective aspect, PL : Plural number, PROG : Progressive aspect, QW : Question word, REL : Relativizer, SG : Singular number, TR : Transitivizer, UV : Undergoer voice prefix, x~y : x and y are forms which freely alternate (in a particular context), e.g. *sama~ama*, '=' indicates a clitic boundary, '-' indicates a bound morpheme boundary

---

[11] Looking at the usage of each possessive verbal predicate more detailed, it is later concluded that the affixes *ber-...-kan* and *-an* take a very limited number of possessees, thus we may presume that they are peripheral to the domain of possession in Indonesian and cannot be regarded as typical possessive verbal predicates. *Ada ...=nya* takes inalienable (and definite) possessees and can be included in *ada*. This accords with the conclusion mentioned in Moeljadi 2011: 131, i.e. only five possessive verbal predicates (*memiliki*, *mempunyai*, *punya*, *ada*, and *ber-*) can be regarded as the real possessive verbal predicates in Indonesian.

References

Alieva, Natalia F. (1992) Malay and Cham possession compared. *Oceanic Linguistics* 31: 13-21.

Alwi, Hasan et al. (2000) *Tata bahasa baku bahasa Indonesia* (A standard grammar of the Indonesian language). Third edition. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa dan Balai Pustaka.

Heine, Bernd (1997) *Possession: cognitive sources, forces, and grammaticalization.* Cambridge: Cambridge University Press.

Hopper, Paul J. (1972) Verbless stative sentences in Indonesian. In: Verhaar (ed.) *The verb 'be' and its synonyms*, 115-152. Dordrecht: D. Reidel.

Kamus Besar Bahasa Indonesia dalam jaringan (KBBI Daring). Based on *Kamus besar bahasa Indonesia* (A compherensive dictionary of Indonesian). Third edition. Available online at http://pusatbahasa.kemdiknas.go.id/kbbi/ (Accessed in November 2011)

KOMPAS.com. Available online at http://www.kompas.com (Accessed in October and December 2009)

Lewis, M. Paul (ed.) (2009) Indonesian. In: *Ethnologue: languages of the world, sixteenth edition.* Dallas: SIL International. Available online at http://www.ethnologue.com/ show_language.asp?code=ind (Accessed in November 2011)

Moeljadi, David (2010) *Possessive predicate constructions in Indonesian.* Unpublished graduation thesis. Tokyo: The University of Tokyo.

_____ (2011) Possessive verbal predicate constructions in Indonesian. *Tokyo University Linguistic Papers (TULIP)* 31: 117-133.

Narrative Story Retell Reference Database page 5 (Story Script for *Frog, Where Are You?* by Mercer Mayer, 1969). In: *Systematic Analysis of Language Transcripts (SALT) software.* Madison: University of Wisconsin-Madison. Available online at http://www.saltsoftware.com/salt/ downloads/NarStoryRetellRDBDoc.pdf (Accessed in July 2011)

Sneddon, James N. (1996) *Indonesian: a compherensive grammar.* London: Routledge.

_____ (2003) Diglossia in Indonesian. *Bijdragen tot de Taal-, Land- en Volkenkunde* 159/4: 519-549.

_____ (2006) *Colloquial Jakartan Indonesian.* Canberra: Pacific Linguistics 581.

Stassen, Leon (2009) *Predicative possession.* Oxford: Oxford University Press.

Young, Pauline V. (1966) *Scientific social surveys and research: an introduction to the background, content, methods, principles, and analysis of social studies.* Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Appendix: Some of the pictures used in the storytelling survey
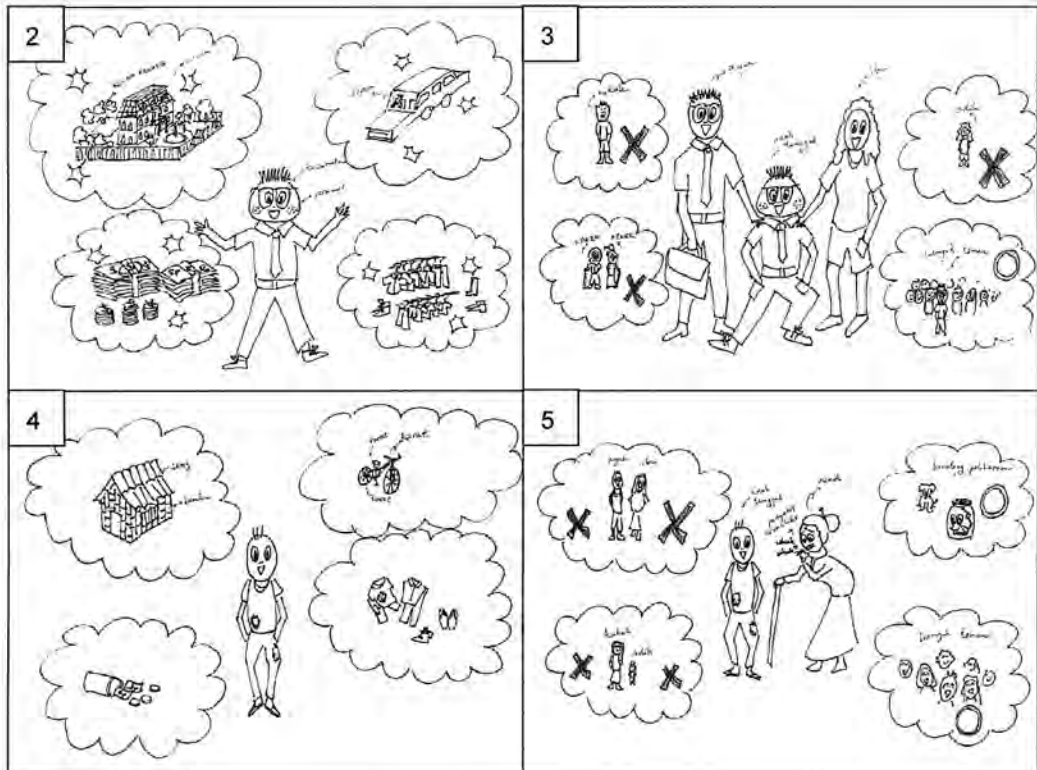
Translations of the words and phrases written on the pictures:

Pict.	Word/Phrase	English translation	Pict.	Word/Phrase	English translation
2	kolam renang	swimming pool	4	seng	zinc
	rumah	house		bambu	bamboo
	sopir	chauffeur		karat	rust
	kacamata	glasses	5	ayah	father
	jerawat	pimples		ibu	mother
3	kakak	elder sibling		kakak	elder sibling
	kakek	grandfather		adik	younger sibling
	nenek	grandmother		anak tunggal	the only child
	ayah	father		nenek	grandmother
	ibu	mother		penyakit	disease
	anak tunggal	the only child		binatang peliharaan	pet animal
	adik	younger sibling		banyak teman	many friends
	banyak teman	many friends		uhuk uhuk	cough cough

インドネシア語の所有動詞述語の使用

—ストーリーテリング調査に基づいた統計的分析—

ダヴィド・ムルヤディ

davidmoeljadi@yahoo.com

キーワード: インドネシア語、H 変種、L 変種、所有、ストーリーテリング、
変種の定量分析、相関係数

要旨

本研究では、インドネシア語（H 変種及び L 変種）の所有動詞述語構文について、その使い分けを考察する。Moeljadi (2010)では、インドネシア語に 8 つの所有動詞述語構文（X *memiliki* Y, X *mempunyai* Y, X *punya* Y, X *ada* Y, X *ada* Y=*nya*, X *ber*-Y, X *ber*-Y-*kan* Z, X Y-*an*）があると述べた。H 変種及び L 変種、それぞれの変種でよく使用される所有述語があるかを調べるために、筆者は 2011 年にインドネシアの東ジャワ州のマラン及び東京でストーリーテリング調査を行った。その調査から得られたデータは定量分析及び相関係数で分析した。主な結論として、所有構文 X *memiliki* Y は H 変種でよく使われており、一方 X *punya* Y は L 変種でよく使われていることが分かった。

（ムルヤディ・ダヴィド）