

The functions of full nominal reduplication in Jakarta Indonesian: A corpus-based examination

Evelyn Elmer Fettes*

Abstract. Existing literature on full nominal reduplication in Indonesian describes the process as marking plurality or variation of type/kind (Sneddon et al. 2010). There are conflicting claims in the literature as to whether fully reduplicated nouns can cooccur with numerals and/or classifiers (Chung 2000; Dalrymple & Mofu 2012). This paper presents a corpus study of full nominal reduplication (FNR) in Jakarta Indonesian (JI), a recent variant of Malay which emerged as a blend of Jakarta Malay (JM) and Standard Indonesian (SI). In particular, I examine the cooccurrence and linear ordering of other nominal elements with fully reduplicated nouns (FRNs), including quantifiers, numerals, classifiers, demonstratives, possessive pronouns, and the definite article *-nya*. The corpus study found no instances of FRNs cooccurring with numerals, but one potential instance of an FRN cooccurring with a classifier. FRNs can cooccur with all other nominal elements. The linear ordering of these elements in JI more closely resembles the ordering in JM than in SI in that the FRNs may either precede or follow demonstratives. The corpus study also explores the interpretation of FNR as it pertains to plurality and variation of kind. It presents evidence that an additional function of FNR in JI is contrastive focus.

Keywords. (full nominal) reduplication; Jakarta Indonesian; corpus

1. Introduction. Reduplication, the process of fully or partially duplicating a base to make a lexical entity, is a ubiquitous phenomenon in Austronesian languages, and the Malayic subgroup in particular. In Indonesian, reduplication may be used for multiple purposes: a verbal base may be reduplicated to indicate a continuous action, an action performed repeatedly, or an action done in a leisurely manner (1a). An adjective may be reduplicated so that it may function as an adverb (1b). A pronoun may be reduplicated to indicate disparagement or emphasis (1c). Nouns may be reduplicated as part of a word-formation process with unpredictable semantics (1d) and some words are formed from a bound base that has no independent meaning (1e).

- (1) Standard Indonesian (Sneddon et al. 2010: 20, 22–23)¹
- | | |
|----------------------------------------------------------|---------------------------------------------------------------------------------------------------------------|
| <p>a. melihat-lihat
RED-MEN.see
'browse' (V)</p> | <p>b. Anak itu berteriak keras-keras.
child that BER.scream RED-loud
'The child screamed loudly.'</p> |
|----------------------------------------------------------|---------------------------------------------------------------------------------------------------------------|

* A version of this paper (one far longer and more rambling) was written for advancement to PhD candidacy at Cornell University. I am indebted to my committee: Sarah Murray (chair), Abigail Cohn, Molly Diesing, and Amalia Skilton for their guidance. I am also thankful to my current dissertation chair, John Whitman, for his support.

Author: Evelyn Elmer Fettes, Cornell University (ee55@cornell.edu).

¹ I have added interlinear glosses to data taken from Sneddon et al. (2010) and Ikaranga (1980), altered glosses for Muhadjir (1981) and Winarto (2016), and occasionally altered the free translations of all for clarity. All affixes (nominal and verbal) except personal pronouns (and sometimes *-NYA*) are glossed verbatim, *yang* is glossed REL (relative clause marker), *di* as LOC (locative), V is for verb, N is for noun, CL is for classifier, and RED (reduplication) precedes the reduplicated base. Many of these choices were made for the sake of consistency, as this is how the JI corpus is glossed, not to hint at any ideological position regarding underlying structures.

- c. Mengapa hanya saya-saya yang selalu diberi tugas yang berat ini?
 why just RED-1SG REL always DI.give task REL heavy this
 ‘Why is it always poor old me who gets the hard jobs?’
- d. mata-mata e. kupu-kupu
 RED-eye RED-KUPU
 ‘spy’ (N) ‘butterfly/butterflies’

Perhaps the most productive form of reduplication is full nominal reduplication (FNR). FNR occurs for the purpose of expressing plurality. There is some controversy in the literature over whether this is FNR’s sole function or whether it may also be done to express “variety of kind” (Sneddon et al. 2010: 20–21). Most simplex and complex nouns may undergo FNR (2).¹

- (2) Standard Indonesian (Sneddon et al. 2010:20)
- | | |
|---------------------|------------------------|
| a. rumah | b. rumah-rumah |
| house | RED-house |
| ‘house/houses’ | ‘(types of) houses’ |
| c. perubahan | d. perubahan-perubahan |
| PER.change.AN | RED-[PER.change.AN] |
| ‘change/changes’(N) | ‘(types of) changes’ |

Typologically, these are two of the most common functions of nominal reduplication as noted in Mattiola & Barotto (2023), who state that functions of nominal reduplication cross-linguistically include: a. plurality (additive plurality, greater plurality, distributive plurality, and collectivity), b. related variety and taxonomy (“all kinds of _N_”), c. set construction and associativity (“_N_ and the like”), d. denoting new referents, i.e. to form to new but related lexical items (as in (1d,e), e. expressing evaluative meanings such as diminutive (as in (1c)), endearment, intensification/attenuation, and authenticity/prototypicality, f. exclusivity, and g. indeterminacy.²

There is additional controversy over whether fully reduplicated nouns (FRNs) may be preceded by numerals and/or classifiers in Indonesian. While some claim it is not possible (such as Carson 2000), both Chung (2000) and Dalrymple & Mofu (2012) claim that it is possible, albeit such constructions being rare and dispreferred. Notably, Chung’s sole example is taken from the *Hikayat Abdullah*, a 19th century Malay text, while Dalrymple & Mofu’s examples were gathered on the internet, which leads to uncertainty about their authors L1 and L2 influences.

This paper is an attempt to explore these controversies in Jakarta Indonesian, a prominent dialect of Indonesian, using a corpus compiled by the Max Planck Institute in Jakarta (Gil & Tradmor 2015). Specifically, it aims to answer two research questions: 1. What nominal

¹ Nouns whose lexical form already has a fully reduplicated base, as in (1e), may not be reduplicated (**kupu-kupu-kupu-kupu*). With regards to the coding of these bound bases in the JI corpus discussed in sections 3 and 4, it appears that words like ‘butterfly’ were not glossed as containing the morpheme RED, as KUPU has no independent meaning. For example, compare the glossing of *gorong-gorong* ‘water channel’ (where GORONG has no independent meaning) to *got-got* ‘gutters’ (from *got* ‘drain’) in the following:

(BTW-010307, SELBTW, 1525-1526)
 la é mang nyeburnya di gorong-gorong, di got-got.
 LA indeed N-plunge-NYA LOC water.channel LOC RED-gutter
 ‘we plunged ourselves into the water channels, into the gutters.’

² English notably utilizes reduplication for authenticity/prototypicality as described in “The Salad-Salad Paper” (Ghomeshi et.al. 2004). This fact is of particular relevance when it comes to section 4.3.

elements cooccur in a noun phrase with an FRN and in what order? and 2. Is there any indication that FNR is done to express meanings beyond plurality and variation of kind?

Section 2 will introduce reduplication and the word order of noun phrases in Standard Indonesian (SI), Jakarta Malay (JM), and (to some extent) Jakarta Indonesian (JI). Section 3 will describe the methodology of the corpus study into FNR in JI. Section 4 will present the results of the study. Section 5 will summarize and discuss the implications of this study for future research.

2. Noun phrases in Standard Indonesian, Jakarta Malay, and Jakarta Indonesian. A

difficulty in conducting literature-based research on “Indonesian” is that the term Indonesian may refer to either A. Standard Indonesian (SI) (*Bahasa Indonesia (BI)*), which is the standardized national language taught in all public schools or B. a closely related colloquial dialect/language (*bahasa sehari-hari*). Jakarta Indonesian (JI) is such a language; it is a newly emerging variant in Jakarta, the current capital of Indonesia, and may be conceptualized of as a blend of SI and Jakarta Malay (JM), a Malayic variant spoken by the local *anak Betawi* residents. This section will examine noun phrases in SI and JM, and how differences between them led to the formation of my research questions for the following corpus study.

2.1. STANDARD INDONESIAN (SI). As shown in (2) above, bare nouns in SI may be interpreted as singular or plural depending on context. Thus, reduplication is often considered an “optional” process. Sneddon et al. (2010: 20–21) states “A noun is not usually reduplicated unless it is unclear from context whether one or more than one is referred to and then only if this is important to what the speaker wishes to convey... Sometimes, however, a speaker does use reduplication even though plurality is clear from context.”

In order to specify a specific numeric quantity, bare nouns may be preceded by numerals and/or classifiers (3). Classifiers too are often “optional”, unless the numeral is the affixal form of *satu* ‘one’ i.e. *se-*, which must cliticize to the classifier.³ Quantifiers (*semua* ‘all’, *segala* ‘all’, *(se)tiap* ‘each, every’, *banyak* ‘many, a lot, much’, *beberapa* ‘several’, *sedikit* ‘few’) uniformly precede the noun, while demonstratives (*ini* ‘this’, *itu* ‘that’) uniformly follow the noun (4) unless they are functioning as demonstrative pronouns (5). Some quantifiers may cooccur with classifiers (6). Possessors, both in free and clitic form, immediately follow the noun and precede the demonstrative (5,7). Sneddon et.al. (2010:139) claims that classifiers cannot cooccur with demonstratives or possessives. Note that in (7) an FRN appears with a demonstrative. No explicit mention is made in Sneddon et.al. (2010) if FRNs may cooccur with quantifiers, classifiers, and/or numerals.

(3) Sneddon et al. (2010: 139)

- | | |
|--------------------|-----------------|
| a. dua (ekor) kuda | b. seorang guru |
| two (CL) horse | one.CL teacher |
| ‘two horses’ | ‘a/one teacher’ |

(4) Sneddon et al. (2010: 138)

- Semua orang itu bekerja di kota.
all people that BER.work LOC city
‘All those people work in town.’

³ The three most common classifiers in SI are *orang* (for people), *ekor* (for animals), and *buah* (for objects). These classifiers are homophonous with nominal lexical items, i.e. *orang* ‘people’, *ekor* ‘tail’, and *buah* ‘fruit’.

- (5) Sneddon et al. (2010: 164) (6) Sneddon et al. (2010: 141)
 Itu teman saya. beberapa buah buku
 that friend 1SG several CL book
 ‘That is my friend.’ ‘several books’
- (7) Sneddon et al. (2010: 171)
 Mahasiswa-mahasiswa ini akan melanjutkan pendidikannya di luar
 RED-student this will MEN.continue.KAN PEN.educate.AN-NYA LOC outside
 negeri
 country
 ‘Most students will continue their education overseas.’

2.2. JAKARTA MALAY (JM). Similarly to SI, bare nouns in JM may be interpreted as singular or plural depending on context (Ikaranga 1980: 67), and hence we may assume that reduplication is viewed as “optional”. Ikaranga (1980: 67) also explicitly states that when nouns are reduplicated, they may not cooccur with numerals (8). No explicit mention is made of FRNs being able to cooccur with quantifiers and/or classifiers. There is an example of an FRN cooccurring with a demonstrative (9.) Ikaranga (1980: 63) also states that “for certain nouns, there are corresponding derived [via reduplication] nouns meaning ‘various types of (N)’” (10). Muhadjir (1981: 77) gives a slightly different description of the semantic function of FNR, stating that it changes the lexical meaning of nouns to expressing “(indefinitely) many” (11a) or “a group of kinds of things or people” (11b). Note in (11a) the reduplicated noun may take the possessive pronoun clitic *-ku* and in (11b) the reduplicated noun is translated as singular.

- (8) Ikaranga (1980: 63)
- | | |
|-------------------------------------------|---------------------------------------------------------------|
| a. oraŋ
person
‘person/people’ | b. oraŋ-oraŋ
RED-person
‘people’ |
| c. due oraŋ
two person
‘two people’ | d. *due oraŋ-oraŋ
two RED-person
Intended: ‘two people’ |
- (9) Ikaranga (1980: 67)
 Tu bir-bir əngəʔ dikəluarin.
 that RED-beer no DI.outside.IN
 ‘You didn’t put out those beers.’
- (10) Ikaranga (1980: 63–64)
- | | |
|---------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------|
| a. Jual telur-telur.
sell RED-egg
‘He sells various kinds of egg dishes.’ | b. Ini buntut-buntut ape ni?
this RED-tailbones what this
‘What kind of tailbones are these?’ |
|---------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------|
- (11) Muhadjir (1981: 77)
- | | |
|--------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------|
| a. ampunilah dosa-dosaku
forgive-LAH RED-sin-1SG
‘Forgive my sins’ | b. kəsip-kəsip kaya kura-kura sədaŋ ɲərəm
RED-wink similar-to RED-turtle PROG sit.on.eggs
‘Winking like a turtle sitting on its eggs.’ |
|--------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------|

Word order and cooccurrence of nominal elements appear to be more fluid in JM than SI. Note that in (9) and (10b), the demonstrative precedes the noun. Ikaranga (1980: 17) states that demonstratives may precede, follow, or precede and follow the noun. They may also cooccur with possessive pronouns (12). Possessive pronouns may cliticize to the FRN as in (11a) and (12c). Quantifiers precede the noun (13), as do numerals (8c) – however if a noun occurs with a numeral AND a classifier, they will follow the noun (14).

(12) Ikaranga (1980: 17)

- | | | |
|------------------|-----------------|----------------------|
| a. anak lu (itu) | b. ni anak ni | c. anakñe itu |
| child 2 (that) | this child this | child-NYA that |
| ‘your child’ | ‘this child’ | ‘that/the/his child’ |

(13) Ikaranga (1980: 18)

Saben ari die datəŋ.
every day 3SG come
‘S/he comes every day.’

(14) Ikaranga (1980: 26)

ayam due biji
chicken two CL (thing)
‘two chickens’

2.3. JAKARTA INDONESIAN (JI) AND RESEARCH QUESTIONS. The contrast between SI and JM naturally begs the question of how noun phrases in JI are structured. While there has been much work done on (Colloquial) Jakarta Indonesian (e.g. Ewing 2005; Sneddon 2006), relatively little attention is paid to how nominal structures in JI compare to those of SI/JM. However, one particular point of interest noted by both Ewing and Sneddon is that the morpheme *-nya*, which functions as a third-person possessive in SI, may also function as an anaphoric definite article in JI (15). (This appears to be true in JM as well, see (12c).) Note that in (15) *-nya* and the demonstrative may cooccur, in contrast to SI. Winarto (2016) expands upon this, showing that in the “Indonesian DP”, the definite article *-nya* can cooccur with the demonstrative, but may not cooccur with a possessive or with a numeral, with or without the classifier present. However, classifiers and possessors/demonstratives can cooccur (16).

(15) Sneddon (2006: 39)

Ngobrol-ngobrol dan itu obrolannya direkam.
N.RED-talk and that talk.AN-DEF DI.record
‘Talk away and the talk will be recorded.’

(16) Winarto (2016: 223)

- | |
|------------------------------------------------|
| a. *Lima (buah) buku-nya mahal sekali |
| five (CL) book-DEF expensive very |
| Intended: ‘The five books are very expensive.’ |
| b. Tiga (buah) bola merah saya itu. |
| three (CL) ball RED 1SG that |
| ‘Those three balls of mine’ |

No overt evidence (to my knowledge) has been presented for the sole purpose of demonstrating the ability of a FRN to cooccur with quantifiers, numerals, classifiers, demonstratives, possessors, and/or the definite article in JI. Given this, as well as the typological range of possible functions for FNR and the controversies mentioned in section 1, I developed two research questions, repeated here:

- 1. What nominal elements cooccur in a noun phrase with an FRN and in what order?

- 2. Is there any indication that FNR is done to express meanings beyond plurality and variation of kind?

3. Methodology of corpus study. In order to address the above research questions, I used a subset of data from the Jakarta Indonesian corpus compiled by researchers at the Max Plank Institute at the Jakarta Field Station, (Gil & Tradmor 2015). This subset included all available ELAN files with native Jakarta Indonesian speakers using JI, 40 .eaf files in total.⁴ These files feature 128 participants, who were typically recorded in their or another participant's home having casual conversations. 15 of these speakers were over thirty years old and ethnically Betawi; the rest were from younger generation and a variety of ethnicities. Descriptions of these participants which include their name, speaker id code, ethnicity, gender, exact age, education level, and additional languages that they speak are available via the open access corpus.

Transcripts from each of the 40 files were exported as .tsv files and read into R. These transcripts included information such as the utterance turn in JI, its English gloss, a free translation, annotators' comments, speaker id, and the begin/end time of the utterance turn. The data frames containing this information were manipulated so that each contained a file id code, and each row corresponded to a single utterance turn. These data frames were then bound together and filtered so any row which did not contain any verbal content (notated as "0.") was excluded, resulting in a single tibble containing 59,357 rows. This tibble was further filtered for strings containing RED in the English gloss column, resulting in a tibble containing of 3328 rows which contained one or more tokens of a word glossed with the morpheme RED. An unnesting code was run on this tibble so that each row contained additional columns for individual words and their glosses, then filtered again to only contain word glosses containing the morpheme RED.⁵ The resulting tibble was then exported as a .csv file and handcoded.

As mentioned in section 1, nearly all parts of speech may be reduplicated for various reasons. As P.O.S. was not encoded in the corpus, I went through each row that contained a reduplicated base and handcoded whether the reduplicated base corresponded to a N(oun) or O(ther). The P.O.S. of the majority of words were easily identifiable. For example, most reduplicated verbs contained additional morphemes, e.g. *N-*, *di-*, *-in*. Morphologically, adjectives do not usually differ from simplex nouns (although nouns may contain additional nominal morphology such as *pe(N)-* and *-an*), so deciding on the P.O.S. was also informed by the English translation, the annotator's notes, my non-native knowledge of SI, the online Indonesian dictionary KBBI, and the structural position of the reduplicated base within the utterance. If the reduplicated base was a noun, I then handcoded the part of speech of the base as N(oun), V(erb), or ADJ(ective).⁶ Also, columns for the exact form the base and for a standardized spelling of the base were added in order to be able to tally the total number of unique bases. The handcoded tibble was then read into R and furthered filtered to contain only instances of both the base and the reduplicated base coded as N(oun)s. In total, the corpus contained 788 tokens of fully

⁴ The corpus contains an additional 7 files in ELAN where speakers used a mixture of JI and either Cirebon or *Bahasa Indonesia*, and several files that were only available in .wav not .eaf. These files were not analyzed.

⁵ I am particularly grateful to Amalia Skilton for instruction in R-coding and providing me with this portion of code, which I would have been unable to write by myself.

⁶ I also coded an additional column for whether the morpheme *-an* was attached AFTER the base was reduplicated, e.g. *mobil-mobilan* 'toy cars' as the form (RED-base).AN can have unique (diminutive!) meanings distinct from FRNs. ***This data and pictures of the tibbles are not included in this version of the paper due to space constraints. Further information/coding is available upon request.***

reduplicated nouns (FRNs). From here, the tibble containing all 788 tokens was again exported as a .csv file and two copies were made.

On the first .csv copy was handcoded relevant information for research question 1 (“What nominal elements cooccur in a noun phrase with an FRN and in what order?”) namely:

- whether the FRN was linearly preceded by a numeral, classifier, and/or a quantifier
- whether the FRN was linearly followed by a demonstrative, possessive pronouns, and/or the suffix -NYA, (functioning as either a 3rd person possessive pronoun or the definite article) and
- whether there was any ordering of words that would be unexpected in SI, i.e. a quantifier following the FRN rather than preceding it and/or a demonstrative preceding the FRN rather than following it.⁷

(By linearly precede/follow, I mean that for all non-affixal words I coded their presence if and only if there were no verbs, verbal TAM particles, adverbs, conjunctions, prepositions, or the relative clause marker *yang* intervening between them and the FRN. The suffix -NYA was coded as present if and only if cliticized to the FRN.)

On the second .csv copy was handcoded for semantically relevant information to research question 2 (“Is there any indication that FNR is done to express meanings beyond plurality and variation of kind?”) namely:

- whether the FRN’s base was a count or non-count (mass/aggregate) noun, and⁸
- if count, whether the FRN’s referent was interpreted as plural,
- whether the FRN appeared to have a “kind” interpretation, and
- whether the FRN appeared to have an additional function such as focus.⁹

Both handcoded .csv files were read into R and manipulated to get the results presented in the following section.

4. Results. This section describes the results of the data manipulation process described above in section 3. 788 tokens of fully reduplicated nouns (FRNs) were found in a total of 759 utterance turns (out of ≈60,000) made by 64 speakers.¹⁰ The FRNs were formed from 264 unique bases – 237 if phonetic variation is ignored/the spelling of bases is standardized.¹¹ Each primary research

⁷ I additionally coded information about whether the FRN was preceded by *para* (an SI particle used to denote a group of people), followed by *pada* (an overt plural marking in JM and JI which may itself be reduplicated), and the linear presence of other nouns and/or adjectives. *See footnote 7.*

⁸ It has been claimed by many (e.g. Dalrymple & Mofu 2012) that Indonesian does not have a count/mass distinction; at the moment I take no firm position on this issue. Judgements on whether a nominal base was coded as a count noun or not were made based on the English translation and an assessment of whether the nominal base was cumulative and/or divisive as defined by Deal 2017, i.e. if the nominal base was neither cumulative nor divisive then it was glossed as a count noun and if not, then not. Notably both count and not-count nouns can be reduplicated, as discussed in section 4.2.

⁹ I additionally coded information about the animacy of the FRN’s referents and whether or not the FRN’s base was a proper or common noun. *See footnote 7.*

¹⁰ While only half (64) of the speakers made utterances containing FRNs, the number of utterance turns that each speaker made varied wildly: 27 made ≤ 10, 31 made 11-100, 34 made 101-500, 20 made 501 to 1000, 15 made 1001 to 3055, and 1 made 11345. Thus, this statistic does not seem unusual, i.e. indicate in any way that FNR is an unfamiliar process to the other half of JI speakers.

¹¹ For example, the base ‘teacher’ appeared in the corpus alternately spelled as *guru* and *guruq* to indicate a final [ʔ].

question is addressed in a subsection below, with an additional subsection addressing the novel find that FRN may function as a marker of contrastive focus.

4.1. WHAT NOMINAL ELEMENTS COOCCUR IN A NOUN PHRASE WITH A FULLY REDUPLICATED NOUN (FRN) AND IN WHAT ORDER? Most crucially, regarding the controversy over numerals being able to cooccur with numerals, no instances were found of numerals cooccurring with FRNs, with or without classifiers. There was only one potential instance of an FRN being preceded by a classifier (17).¹² There were no instances of a numeral and or classifier following an FRN. (There were several instances of *orang* following an FRN, but in all of these cases, *orang* seems to be acting as a possessor/modifier (18).)

(17) (BTW-310807, JAKBTW, 1756-7)¹³
 orang temen-temennya yang lain ada, cuma diaq doang ngga ada.
 person RED-friend-NYA REL other exist only 3 just NEG exist
 ‘her friends were all there, but she wasn’t there.’

(18) (BTJ-170709, BTJABH, 1449-1451)
 makanya aib, aib-aib orang tu saya ogah gitu.
 that.is.why-NYA fault, RED-fault person that 1SG unwilling like.that
 ‘you know, I hate insulting other people's mistakes.’

There were 80 instances of -NYA appearing with an FRN. It appears to both function as a 3rd person possessor (as in (17)) and as the definite article (19). There were 36 instances of an FRN being followed by a non-clitic possessive pronoun: 19 with *saya* (1SG) (20), 1 with *gua(q)* (1SG), 4 with *lu* (2SG/PL), 5 with *dia(q)* (3SG), 6 with *kita(q)* (1PL), and 1 with *mereka* (3PL). There were no instances of an FRN occurring with other non-*nya* clitic possessive pronouns, e.g. *-ku* (1SG), but as possessive *-nya* may cliticize to the FRN, it seems likely that this is not a grammatically motivated gap.

(19) (BTW-240707, EXPOKK, 667-8)
 guruq-guruqnya masi ada kaliq yaq?
 RED-teacher-NYA still exist maybe yes
 ‘maybe the teachers are still there, right?’

(20) (BTJ-140408, BTJROH, 2699-2700)
 sampéq ruma saya critaq amaq anak-anak saya gitu.
 arrive house 1SG story with RED-child 1SG like.that
 ‘when I went home, I shared the story to my children.’

There were 82 instances of an FRN being followed by a demonstrative (n.b. the SI pattern): 72 with a variant of *itu* (*tu*, *tuh*, *tuq*) ‘that/those’ (as in (18)) and 10 with a variant of *ini* (*ni*) ‘this/these’ (21). There were 24 instances of a demonstrative preceding the FRN (n.b. the JM pattern). While in some utterances, it seems that the demonstrative is modifying the FRN (22), in other utterances its role is less clear, e.g. in (23) *ini* is perhaps a demonstrative pronoun in a separate nominal projection.

¹² It is possible that *orang* here is functioning as a full noun. If it is a CL, perhaps *orang*’s function here is to distinguish between the near homophones of ‘friend’ and ‘park’ (*teman* and *taman* in SI).

¹³ All data from the corpus is copied verbatim. The citation line notes (FILENAME, SPEAKERID, TIMESTAMP).

- (21) (BTJ-250508, BTJMAT, 450-1)
 orang-orang ini masi dapet jatah tu.
 RED-person this still get allotment that
 ‘these people still got allotment at that time.’
- (22) (BTJ-170709, BTJABH, 1870-1873)
 mang sini ada tu ayam-ayam jago kan lagi dibersihin ayam.
 to here exist that RED-chicken master KAN prog DI.clean.IN chicken
 ‘there’s a place here where roosters being cleaned.’
- (23) (BTJ-230109, EXPERN, 96-98)
 trus ini daerah rawa-rawa gitu?
 continue this region RED-swamp like.that
 ‘was it a swamp area?’

Lastly, there were 14 instances of an FRN being preceded by a quantifier. 11 were instances of an FRN preceded by *banyak* ‘many/a lot’ (24) and 3 were instances of an FRN preceded by *segala* ‘all’ (25).¹⁴ Surprisingly, there were also 17 instances of quantifiers following FRNs: 13 instances of *banyak* and 4 of *semua(nya)* ‘all (of them)’. In some utterances, the quantifier was clearly not referring to the FRN, but to an elided noun, e.g. in (26) *banyak* is not referring to the cottages, but to another referent likely already established in the common ground (with the original annotator commenting “possibly referring to ‘*gembili*’”). However in the majority of utterances, the quantifier clearly refers to the FRN (27).

- (24) (BTJ-20070513, EXPERN, 1126-7)
 tapiq kéyéqnya kan banyak pohon-pohon, tu banyak rumput-rumput.
 but like-NYA KAN a.lot RED-tree that a.lot RED-grass
 ‘but there are many trees and lots of grass there, right?’
- (25) (BTW-260507, XXX, 1472-3)
 iyaq, petéq, timun, segala daun-daun ada, duh!
 yes k.o.green.bean cucumber all RED-leaf exist EXCL
 ‘right, ‘petéq’, cucumber, and kind of leaf, oh my God!’
- (26) (BTJ-080509, BTJMIS, 3303-3305)
 sono di pondok-pondok itu banyak.
 there LOC RED-cottage that a.lot
 ‘there are many there at the cottages.’
- (27) (BTJ-170709, BTJABH, 2421-2424)
 kalo bis-bis SMS mang banyak dulu mah.
 TOP RED-bus SMS really a.lot formerly MAH
 ‘years ago, there were lots of SMS public transportations.’

4.2. IS THERE ANY INDICATION THAT FNR IS DONE TO EXPRESS MEANINGS BEYOND PLURALITY AND VARIATION OF KIND? Of the 788 FRN tokens, 723 were count nouns. These 723 were formed from 214 bases / 189 bases with standardized spelling. 65 tokens were mass/aggregate nouns formed from 50 bases /48 bases with standardized spelling.

¹⁴ Often *segala* has a “stronger” connotation of ‘each and every’ according to my SI Professor Ibu Jolanda Pandin, whom I am also incredibly indebted to for her language instruction and support.

Of the 723 instances of FRNs corresponding to count nouns, there were only 587 instances in which the translation unambiguously indicated that the FRN was interpreted as plural, i.e. the translation of the FRN contained an English plural marking, such as ‘teachers’ in (19) and ‘children’ in (20).

In some instances, the translation unambiguously refers to variation of kind (28). In other instances, although the translation does not specifically refer to variation of kind, an interpretation of that type could be inferred, e.g. in (29-30), the speaker may be referencing specific kinds of social activities / burial mats. However, in other instances, it is unclear what function FNR is performing. In (31) the base *paser* lexically already refers to a specific type of outdoor market and in (32) if the speaker is addressing a singular referent, it is impossible that *bentuk-bentuk* ‘shape’ be interpreted as plural.

- (28) (BTJ-040808, BTJMAS, 1146-7)
 iya, kuéq-kuéq gitu iya.
 yes RED-cake like.that yes
 ‘yes, right, kinds of snacks.’
- (29) (BTJ-230109, EXPERN, 329-335)
 kalo di sini masih ada kegiatan-kegiatan itu gak Bu sosial gitu yang, yang... apa ya namanya?
 TOP LOC here still exist RED-KE.active.AN that NEG TRU.mother social like.that REL REL what yes name-NYA
 ‘is there any social activity in this place that... er... what is it?’
- (30) (BTJ-250508, BTJINA, 2315-6)
 iya, maq tiker-tikernya dikubur.
 yes with RED-mat-NYA DI.grave
 ‘yes, it was buried together with the mat.’
- (31) (BTJ-080509, BTJMAI, 2425-2426)
 karang mah gak ke pasar-pasar.
 now MAH NEG to RED-market
 ‘now I don’t go to the market anymore.’
- (32) (BTJ-010707, BTJMIN, 1859-60)
 ‘ah bentuk-bentuk lu gua ngga tauq.’
 EXCL RED-form 2 1SG NEG know
 ‘hey, even I don’t know your shape.’

Of the 65 instances of non-count FRN tokens, there were no overt implications that FNR was done for the purpose of expressing “plurality” in terms of great(er) quantity/volume. In some instances, FNR appears to be indicating variation of kind. In (33) the FRN appears to be referencing a specific kind of hairstyle and in (34) the FRN appears to be referencing a kind of sweet that has prepared in a specific style. In (35) *musik-musik* could be referencing genres of music. However, parallel to *paser* in (31), the base *dangdut* lexically already references a specific genre of music, hence the function of FNR here is unclear. Also in (36-37) there is no obvious difference between the meaning of the FRNs and their nominal bases.

- (33) (BTJ-111109, BTJSIA, 1320-1323)
 kalu rambut-rambut nama jenggot kalu orang sini.
 TOP RED-hair name beard TOP person here
 ‘the hair is called ‘jenggot’ by the native speaker of this place.’

- (34) (BTW-260604, UDIBTW, 608-9)
 dodol si dodol cuman bukan dodol-dodol kaya orang dodol Betawi paké apiq.
 k.o.sweet SIH k.o.sweet only NEG RED-k.o.sweet like person k.o.sweet Betawi use fire
 ‘it’s dodol but it’s not dodol like Betawi dodol that uses fire.’¹⁵
- (35) (BTJ-170709, BTJABH, 1347-1352)
 kalo musik-musik... memang sih dangdut-dangdut saya jelas, saya jelas saya denger gitu.
 TOP RED-music indeed SIH RED-k.o.music 1SG clear 1SG clear 1SG hear like.that
 ‘talking about music... yeah, actually I like to listen to dangdut music.’
- (36) (BTJ-170709, EXPERN, 2369-2371)
 dari bambu-bambu gitu, bukan?
 from RED-bamboo like.that NEG
 ‘wasn’t it made of bamboo?’
- (37) (BTJ-040808, BTJMAS, 212-3)
 aér-aér ajaq.
 RED-water just
 ‘it’s just mineral water.’¹⁶

Given the numerous instances of FNR not indicating plurality or variation of kind, I posit that FNR must be able to serve an additional function in JI. This function fits neatly into one of Mattiola & Barotto (2023)’s typological categories, repeated here from section 1: FNR may express “evaluative meanings such as diminutive, endearment, intensification/attenuation, and authenticity/prototypicality”. The clearest example that FNR may perform this function is described in the following section.

4.3. INSTANCE OF CONTRASTIVE FOCUS! In (38) below, the speaker appears to be emphasizing the authentic origins of the people they are discussing. The first argument Pasar Minggu (a district of South Jakarta) is not reduplicated, whereas the second argument Lampung (a province of Southern Sumatra) is. The free English translation uses the word ‘actually’ to describe the origins of the people who are the topic of conversation, despite no independent lexical item in the JI utterance having that meaning. This suggests that the reduplication of Lampung marks it as these people’s actual place of origin, in contrast to Pasar Minggu. Perhaps the “they” have lived in Pasar Minggu for several years despite being born in Lampung – much as I might say in English “I’m from Ithaca, but REALLY I’m from Fair Haven”, prosodically indicating that Fair Haven is where I was born.

- (38) (BTW-010307, JAKBTW, 2100-1)
 dari Pasar.Minggu jugaq dari Lampung-Lampung jugaq itu.
 from Pasar.Minggu also from RED-Lampung also that
 ‘they are from Pasar Minggu, but they are actually from Lampung.’

In light of this example, we can reanalyze the previously confusing data in section 4.2: In (31) the speaker may be implicitly contrasting *paser* ‘outdoor market’ with a supermarket. In (32) the speaker could be implicitly contrasting the addressee’s shape with another person’s shape. In

¹⁵ The annotator’s comment reads “referring to the other kind of dodol which is cooked on stove while Betawi dodol is cooked on firewood.”

¹⁶ The annotator’s comment reads “implicitly telling there’s only mineral water that prepared for EXPERN without snack that’s usually given to a guest.”

(36) the speaker could be correcting a previous speaker's assertion that the object in question was made out of a different material. Lastly, in (37) the speaker is implying that there is only water available as opposed to water and a snack (as stated in the annotator's comment). These examples have been repeated in (39-42) below for convenience.

- (39) (BTJ-080509, BTJMAI, 2425-2426)
karang mah gak ke pasar-pasar.
now MAH NEG to RED -market
'now I don't go to the market anymore.'
- (40) (BTJ-010707, BTJMIN, 1859-60)
'ah bentuk-bentuk lu gua ngga tauq.'
EXCL RED-form 2 1SG NEG know
'hey, even I don't know your shape.'
- (41) (BTJ-170709, EXPERN, 2369-2371)
dari bambu-bambu gitu, bukan?
from RED-bamboo like.that NEG
'wasn't it made of bamboo?'
- (42) (BTJ-040808, BTJMAS, 212-3)
aér-aér ajaq.
RED-water just
'it's just mineral water.'

5. Summary and directions for future research. The JI corpus study has shown:

- No evidence that FRNs can cooccur with numerals.
- Possible evidence that FRNs can be preceded by classifiers.
- Evidence that FRNs can cooccur with quantifiers, demonstratives, possessive pronouns, and the definite marker *-nya*.
- That JI appears to pattern closer to JM than SI with regards to the position of demonstratives.
- That JI allows quantifiers to both linearly proceed and follow FRNs, unlike in SI/JM.
- Evidence that plurality and related variety are functions of FNR.
- Evidence that count nouns can be reduplicated to express plurality, but no evidence that non-count nouns may be.
- Evidence that FNR has an additional evaluative function, indicating authenticity/prototypicality/contrastive focus.

While giving my LSA talk, I was asked why I didn't just do fieldwork with native Jakarta Indonesian speakers in order to answer my research questions. This was likely due to the fact that, unlike a consultant, a corpus cannot provide negative data, i.e. tell us what isn't possible, such as whether FRNs can cooccur with numerals. However, my answer then, and now, is that doing a corpus study – if possible – is a valuable preliminary step prior to conducting fieldwork. The information gained from such a study may help guide a researcher in their development of future research questions. If I had not conducted a corpus study, I wouldn't have imagined that contrastive focus, despite being a function of FNR in English, was a function of FNR in Jakarta Indonesian. Now that I am aware of the possibility, in future fieldwork with JI consultants and

speakers of other Malayic variants, I can attempt to elicit whether this function exists in their grammar.

At this stage, I cannot posit what the underlying syntactic structure of the Jakarta Indonesian DP (or NP, or NumP) looks like. However, I suspect that there is a reduplicative morpheme (RED1) which represents an indefinite number and lives in the same syntactic position as a numeral would (hence numerals and RED1 cannot cooccur), and that there is a separate reduplicative morpheme (RED2) which is associated with contrastive focus (/potentially other evaluative meanings). The rare/dispreferred nature of the combination of a numeral and a noun + RED2 may stem from a listener's initial confusion of RED2 with RED1. FRIN!

References

- Carson, Jana Cécile. 2000. The semantics of number in Malay noun phrases. Calgary: University of Calgary MA thesis.
- Chung, Sandra. 2000. On reference to kinds in Indonesian. *Natural Language Semantics* 8. 157–171.
- Dalrymple, Mary & Suriel Mofu. 2012. Plural semantics, reduplication, and numeral modification in Indonesian. *Journal of Semantics* 29 (2). 229–60.
<https://doi.org/10.1093/jos/ffr015>.
- Deal, Amy Rose. 2017. Countability distinctions and semantic variation. *Natural Language Semantics* 25(2). 125–171. <https://doi.org/10.1007/s11050-017-9132-0>.
- Ewing, Michael C. 2005. Colloquial Indonesian. In K. Alexander Adelaar & Nikolaus Himmelmann (eds.), *The Austronesian languages of Asia and Madagascar*, 227–258. London: Routledge. <https://doi.org/10.4324/9780203821121>.
- Gil, David & Uri Tadmor, 2015. *Jakarta Indonesian*. 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.
<https://hdl.handle.net/1839/00-0000-0000-0022-5AC9-0>.
- Ghomeshi, Jila, Ray Jackendoff, Nicole Rosen & Kevin Russell. 2004. Contrastive focus reduplication in English (the salad-salad paper). *Natural Language & Linguistic Theory* 22(2). 307–357. <https://doi.org/10.1023/B:NALA.0000015789.98638.f9>.
- Ikaranga, Kay. 1980. *Melayu Betawi grammar* (NUSA: Linguistic Studies in Indonesian and Languages in Indonesia Volume 9). Jakarta: Universitas Atma Jaya.
- Mattiola, Simone & Alessandra Barotto. 2023. Nominal reduplication in cross-linguistic perspective: From plurality to change of referents' specificity. *Studies in Language* 47(1). 135–189. <https://doi.org/10.1075/sl.21050.mat>.
- Muhadjir. 1981. (Translated by Kay Ikaranga.) *Morphology of Jakarta dialect: Affixation and reduplication*. Jakarta: Badan Penyelenggara Seri NUSA, Universitas Atma Jaya.
- Setiawan, Ebta. Arti kata – Kamus Besar Bahasa Indonesia (KBBI). Website.
<https://www.kbbi.web.id/>.
- Sneddon, James N. 2006. *Colloquial Jakartan Indonesian*. Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies, The Australian National University.
- Sneddon, James N., K. Alexander Adelaar, Dwi Djenar & Michael Ewing. 2010. *Indonesian: A comprehensive grammar* (2nd edn.). Routledge comprehensive grammars. London: Routledge.

Winarto, Ekarina. 2016. The Indonesian DP. In Henrison Hsieh (ed.), *AFLA 23: The Proceedings of the 23rd meeting of the Austronesian Formal Linguistics Association*, 220–237. Canberra: The Australian National University. <http://hdl.handle.net/1885/101155>.