An optional classifier language with D: Seram Timur

Overview This paper investigates the classifiers and DP structure of Seram Timur, an under-documented language spoken in Eastern Seram Island, Maluku, Indonesia. I argue that Seram Timur is an optional classifier language with an overt definite determiner (D) *-(r)a*. To account for the observed N-D-(Clf)-Num word order, I propose a post-syntactic D-lowering analysis. Crucially, the novel data from Seram Timur suggest the need for a refinement to Jiang's (2018) generalization, specifically in further distinguishing between obligatory (e.g., Nuosu Yi) and optional classifier languages (e.g., Seram Timur), and challenge the reliability of classifier-for-noun diagnostics (cf. Little et al. 2022).

Background OPTIONAL CLASSIFIERS. Seram Timur only contains a very small number of classifiers, which optionally occur with selected nouns. Based on the data I have collected, I work with the hypothesis that there are two classifiers in Seram Timur, *kusa* and *woi*. The former is the classifier for animals; the latter is for fruits, see (1). In other words, Seram Timur classifiers are always optional and numerals can directly combine with a count noun. The majority of Seram Timur count nouns cannot occur with any classifier, for the simple reason that they are not animals or fruits (2).

(1) a. kafuna (**kusa**) tolu
dog CLF three
'three dogs'

b. ayai (**woi**) lim
mango CLF five
'five mangoes'

(2) guru (*kusa) tolu
teacher CLF three
'three teachers'

(3) ayai-**ra** (woi) lim
mango-DEF CLF five
'the five mangoes'

DEFINITENESS MARKING. Generally, Seram Timur employs *-(r)a* in marking definites, with *-ra* attaching to nouns that end in a vowel and *-a* to the noun ending in a consonant. Specifically, *-(r)a* linearizes after the noun, turning an indefinite DP into a definite one, as illustrated in (3) (cf. (1b)). Moreover, *-(r)a* is not a demonstrative as Seram Timur has the proximal demonstrative *iwa* and the distal demonstrative *ira*. *-(r)a* is obligatory for unique entities, such as *the sun* (4a), whereas using *iwa* or *ira* in these contexts is infelicitous due to the anti-uniqueness requirement of demonstratives (Robinson 2005), see (4b). In addition, *-(r)a* cannot attach to numerals nor classifiers (5), contrasting with demonstratives, which follow numerals in the nominal domain (i.e., N-(Clf)-Num-Dem). It is noted that Seram Timur allows *-(r)a* and a demonstrative to co-occur within the same nominal domain (6). In particular, (6) is used in a context where there are multiple cups of water.

(4) a. Ku-suka utal-*(**a**).
1SG-like sun-DEF
'I like the sun.'

b. #Ku-suka utal iwa/ira.
1SG-like sun this/that
'I like this/that sun.'

(5) boi kusa(*-ra) roti(*-ra)
pig CLF(*-DEF) two(*-DEF)
intended: 'the two pigs'

(6) Ku-suka ar-**a** **iwa/ira**.
1SG-like water-DEF this/that
'I like this/that water.'

Furthermore, evidence from the consistency test (Löbner 1985), which is used to distinguish demonstrative from definite determiners, shows that *-(r)a* leads to a contradiction in (7) and thus confirms its role as a definite determiner. On the contrary, use of demonstratives in the same context is felicitous.

(7) #Ali na-suka kafuna-**ra** tapi na-suka kafuna-**ra** tei.
Ali 3SG-like dog-DEF but 3SG-like dog-DEF NEG
Lit. 'Ali likes the dog but doesn't like the dog.'

Analysis I take DP as the maximal extension of the nominal projection in Seram Timur given that it features the definite determiner *-(r)a*. Despite the fact that Seram Timur has SVO word order, I propose that the parameter of the noun phrase in Seram Timur is head-final. [CLF-NUM] CONSTITUENT. **a.** Seram Timur classifiers form a constituent with the numeral they co-occur with. We see in (8) that [Clf-Num] phrases can be used pronominally, and we see in (10) that [Clf-Num] phrases can be used as answer fragments. **b.** Classifiers must co-occur with a numeral. Classifiers appearing without a numeral yields ungrammaticality (9).

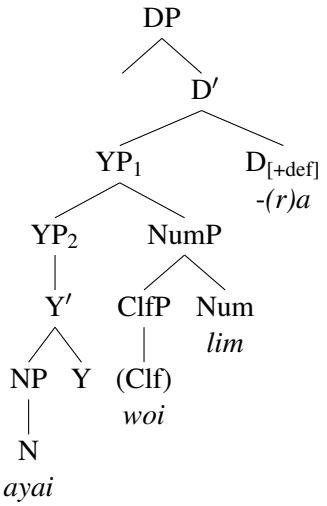
(8) Ku-langa boi (kusa) lim loka. Kusa tolu da-kifit si.
1SG-see pig (CLF) five PERF CLF three 3PL-sleep PROG
'I saw five pigs. Three are sleeping.'

(9) *Ku-langa boi kusa loka.
1SG-see pig CLF PERF
intended: 'I saw a pig.'

(10) Q: How many pigs are there?
A_1: Boi kusa tolu.
pig CLF three
'Three pigs.'
A_2: Kusa tolu.
A_3: Tolu.

POST-SYNTACTIC D LOWERING. To account for the word order and [Clf-Num] constituency, I propose that Seram Timur features a functional projection (YP) within the nominal spine (Cinque 2005; Little et al. 2022), with classifiers occupying the specifier position of the Numeral Phrase (NumP). The Num head selects a Classifier Phrase (ClfP) and the NumP adjoins to the right of YP. The higher YP then merges with D, as illustrated in (11). Adopting the functional projection YP yields a welcome result with respect to plural marking, as Seram Timur allows the co-occurrence of classifiers and the plural marking *-si*, as shown in (12). (11) allows the functional head Y to host the plural marker. Following the D-lowering approach proposed by Embick & Noyer (2001), I argue that the definite [N-D-Clf-Num] construction in Seram Timur is also best captured by a D-lowering analysis, specifically, $D_{[+def]}$ lowers to N (when there is no adjective in between) prior to Vocabulary Insertion, thus explaining the fact that the phonological form of D varies according to properties of the host. ADJECTIVE MODIFICATION. In Seram Timur, adjectives strictly follow the noun, and the definite determiner *-(r)a* necessarily attaches to the adjective (13a). Marking definiteness on the head noun will lead to ungrammaticality, see (13b) and (13c). This restriction is captured by extending the D-lowering analysis: D lowers to A when adjectives are present. Here, I assume that Seram Timur exhibits an AP-over-NP structure, with D lowering to the head A.

(11)



(12) kafuna-**si kusa** tolu
dog-PL CLF three
'three dogs'

(13) a. bunga futi-ra
flower white-DEF
'white flowers'
b. *futih bunga-ra
c. *bunga-ra futih

Cross-linguistic implications (i)TYPOLOGIES OF CLASSIFIER LANGUAGES WITH D. Classifier languages with overt definite determiners are very rare across languages. However, Jiang (2012, 2018) proposes that Nuosu Yi (Tibeto-Burman, China) is such a language as it features both *obligatory* classifiers and an overt D, namely, *su*. Nuosu Yi exhibits the stringent [N-Num-Clf-D] linear word order in which the classifier is required. Given this, Jiang (2018:34) predicts that in classifier languages with an overt D, D should only apply at higher nominal levels that are property-denoting, such as numeral-classifier phrases, and not at the level of kind-referring bare nouns. On the other hand, Seram Timur challenges this prediction, as it is a classifier language with D, yet the co-occurrence of a bare noun and the definite determiner *-(r)a* is allowed. Novel data from Seram Timur thus suggests a refinement to Jiang's prediction: **a.** in *obligatory* classifier languages, D cannot combine with a bare noun; **b.** in *optional* classifier languages, D may combine with either a bare noun or a numeral-classifier phrase. (ii) CLASSIFIERS FOR NUMERALS OR NOUNS? A long-standing debate concerns whether classifiers are needed to allow numerals to count or to allow nouns to be counted (classifiers 'for' numerals or 'for' nouns) (Krifka 1995; Chierchia 1998; Bale & Coon 2014). Little et al. (2022) propose that the variation in whether a noun requires a classifier serves as a diagnostic for classifier-for-noun languages. Seram Timur classifiers are 'for' numerals; however only a small set of Seram Timur nouns may combine with classifiers, thus it appears to be a classifiers-for-noun language despite being the opposite. Similarly, Jinghpaw (Tibeto-Burman, Myanmar) is also argued to be a classifier-for-numeral language but with only two classifiers available – one for humans and one for animals (Gao *to appear*). The cases of Seram Timur and Jinghpaw suggest that exceptions to the diagnostic may not be as uncommon as previously assumed, calling for careful application.