

Genders and classifiers in the perspective of dynamic functional typology

Masayoshi Shibatani
Deedee McMurtry Professor Emeritus of Linguistics, Rice University
Professor Emeritus, Kobe Univesity
Visiting Professor, National Intititue for Japanese Language and Linguistics

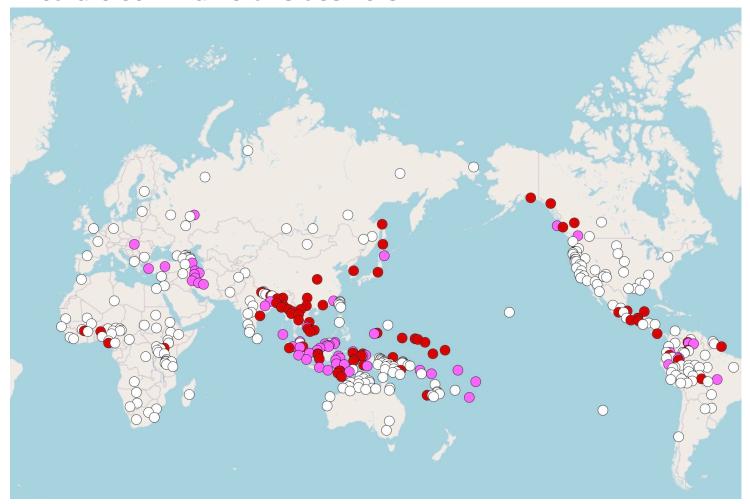
Introduction

Goals

- Motivate the tyoological framework dubbed "dynamic functional typology".
- Starting with a look at the problems of the kind of typology assumed by "The World Atlas of Language Structure" (WALS) by the (former) Leipzig group.
- Illustrate how the proposed typological framework improves our understanding of the phenomena surrounding classifiers and (grammatical) gender
 - by allowing a new analysis that (1) unifies classifiers and genders (2) as classifying nominalizations, as opposed to ordinary nominalizations that yeild nominal structures denoting "things" and thing-like objects without classification. CLFs/genders nominalize and then classify what the nominalized structures denote/refer.
 - by providing a theoretical framework that (1) constrains synchronic distributions of CLF-/gender-marked forms and that (2) predicts diachronic developments of classifying nominalization constructions.

The World Atlas of Language Structures Online

Feature 55A: Numeral Classifiers





David Gil. 2013. Numeral Classifiers. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) WALS Online (v2020.3) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.7385533 (Available online at http://wals.info/chapter/55, Accessed on 2023-06-14.)

Problem 1

 Isolating classifier-marking of numerals from other cases of CLF-marking, often involving the same CLFs.

Haas (1942) already points out that the term "numeral classifier" may not be appropriate in view of the fact that classifiers in languages like Thai mark other structures in addition to numeral.

Haas, Mary. 1942 The use of numeral classifiers in Thai. Language 18:201-215.

Northern Thai*

Context

Responses/follow ups

Numeral	NP use	Modification use
tǔa khây dâay bà-mûaŋ	(pôən khây dâay)	
2SG want take mango	[sǒoŋ <mark>kèn</mark>]	bà-mûaŋ [sŏɔŋ <mark>kèn</mark>]
kìi kèn? how.many CLF _{FRUIT}	two CLF '(I want) two.CLF'	mango two CLF 'two mangos'
'How many mangos do you	want?'	

Demonstrative

pôən mii bà-mûaŋ	sìi kèn.	[<mark>kèn</mark> níi] wǎan.	bà-mûaŋ	[(kèn) níi]
1SG have mango	four CLF	CLF this sweet	mango	CLF this
'I have four mangos.'		'This one is sweet.'	'this mang	50'

^{*}Research on Nothern Thai reported here has been supported by the JSPS project "Cross-linguistic studies on grammatidal nominalizations: with a focus on classifiers and gender markers" (22H00659 PI: Kazuhiro Kawachi)

Context

Responses/follow ups

Modification use

Genitive

tìn níi cín khỏoŋ pôən. CLF_{PIECE} this meat GEN 1SG 'This is my meat.'

Adjective

từa khây dâay măa từa day? 2SG want take dog CLF which 'Which dog do you want?'

Relative clause

cûay aw sîn tǔa mày
help take sarong CLF new
maa hûu nòy!
come give ADV
'Please bring me the new sarong!'

NP use

(tìn) [khỏoŋ tửa] cín [khỏoŋ tửa] CLF GEN 2SG meat GEN 2SG mii nǎy? 'your meat' have where 'Where is yours?'

[tǔa nốɔy]măa [(tǔa) nốɔy]CLF smalldog CLF small'(The) small (one).''small dog'

(tŭa) [tîi hôy

CLF NMLZR hang nay tûu kă]? in closet Q '(The one) hanging in the closet?'

sîn (ŭa) [tîi hôy sarong CLF NMLZR hang nay tûu kǎ]? in closet Q 'The sarong that is hanging in the closet?'

Position of Nothern Thai



(https://en.wikipedia.org/wiki/Thai_language)

Problem 1

- Isolating classifier marking of numerals from other cases of CLF marking, often involving the same CLFs.
- This is unfortunate because the fact that classifiers mark structures other than numerals is important in understanding the basic function of classifiers

The most widely-held idea on the function of NUM CLF claims that those languages without compulsory singular/plural morphological distinction need classifiers to individuate referred objects for the purpose of counting (Greenberg 1994, Chierchia 1998, etc.).

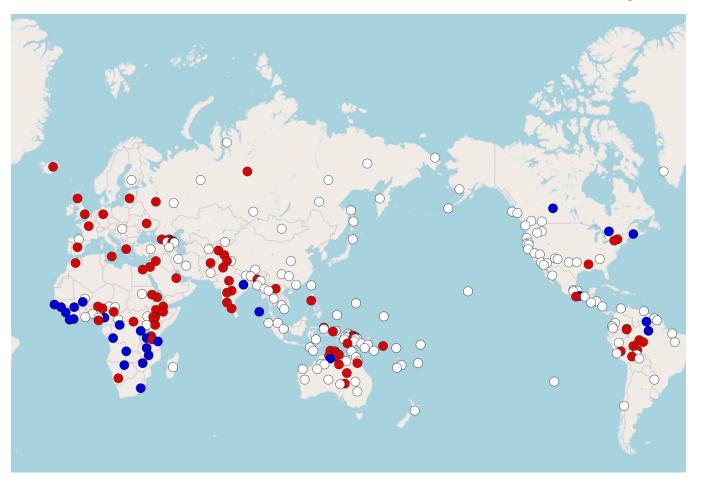
If so, why the same (numeral) classifiers are used to mark Demonstratives, Genitives, etc., as in Thai, that have nothing to do with counting?

Problem 2

Treating CLFs and grammatical Gender separately

The World Atlas of Language Structures Online

Feature 31A: Sex-based and Non-sex-based Gender Systems



Values		
• •	No gender	145
• >	Sex-based	84
• •	Non-sex-based	28

Greville G. Corbett. 2013. Sex-based and Non-sex-based Gender Systems. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) WALS Online (v2020.3) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.7385533 (Available online at http://wals.info/chapter/31, Accessed on 2023-06-14.)

But classifiers and genders are both classificatory systems

1. They both categorize nominal denotations/references

2. They show similar distributions

(Brazilian) Portuguese gender marking

NP use Modification use

Numeral

Quantos carros você vê? (Eu vejo) Um.

'How many cars do you see?' (I see) 'One.M.'

Uma. uma casa

um

este

'one.M car'

carro

carro

esta casa

'this.F house'

'How many houses do you see? 'One.F' 'one.F house'

Demonstrative

Qual é o seu carro? Este.

'Which is your car?'

'This.M'

'this.M car'

Qual é a sua casa? Esta. 'Which is your house?'
'This.F.'

M=masculine, F=femine

NP use Modification use

branco

Genitive

De quem é esse carro? Meu. meu carro 'Whose is this car?' 'my.M car' 'Mine.M'.

De quem é essa casa? Minha. minha casa 'Whose is this house?' 'my.F house' 'Mine.F.'

Adjective Qual é o seu carro? 0 branco. carro 'white.M car' 'Which is your car?' 'The.M white.M.'

Qual é a sua casa? branca. branca casa 'Which is your house?' 'The.M white.F.' 'white.F house'

Relative clause

NP use (Showing a family photo)

O [sentado na cadeira] é meu pai.
the.M sitting.M in.the chair is my father
'The.M [(one) sitting.M in the chair] is my father.'

A [sentada na cadeira] é minha mãe. 'The.F [(one) sitting.F in the chair] is my mother.'

Modification use

O homem [sentado no cadeira] é meu pai. the.M man sitting.M in.the chair is my father 'The.M man sitting.M in the chair is my father.'

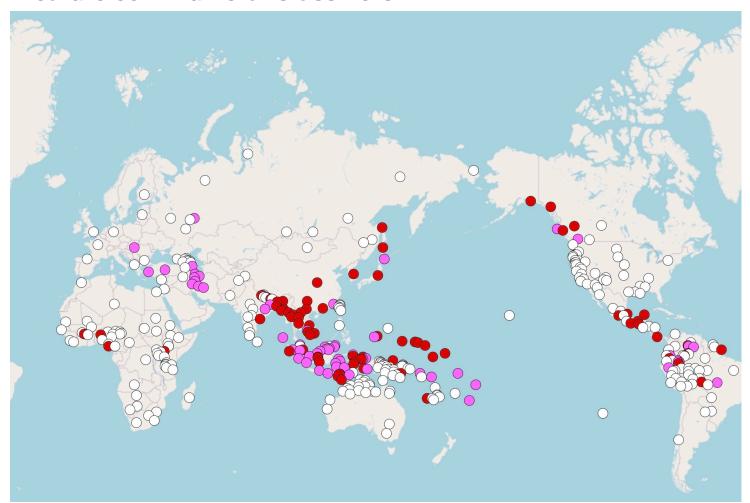
A mulher [sentada na cadeira] é minha mãe. 'The.F woman sitting.F in the chair is my mother.'

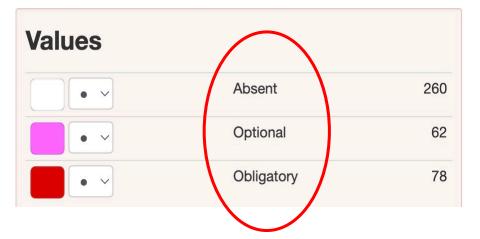
- But classifiers and genders are both classificatory systems
 - 1. They both categorize nominal denotations/references
 - 2. They show similar distributions
 - 3. They both show "agreement"

```
German
                               Mandarin Chinese
ein
      großer Mann
                               三 支 铅笔
one.M tall.M man(M)
                               sān zhī qiānbì
'a/one tall man'
                               three CLF pencil
*eine große Mann
                               *三 本 铅笔
 one.F tall.F man(M)
                               sān běn qiānbǐ
 'a/one tall man'
                               three CLF pencil
eine große Frau
one.F tall.F woman(F)
                               sān běn shū
 'a/one tall woman'
                               three CLF book
```

Back to classifiers

Feature 55A: Numeral Classifiers





Assumes holistic language typology: languages are (a) non-classifying, (b) optionally classifying or (3) classifying.

Such an approach entails erroneous/ imperfect characterizations of the use of CLFs/genders at large and even limiting attention to NUM CLFs.

David Gil. 2013. Numeral Classifiers. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) WALS Online (v2020.3) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.7385533 (Available online at http://wals.info/chapter/55, Accessed on 2023-06-14.)

At large,

CLF-/Gender marking is construction-specific. Even if it is obligatory with numerals, it may be optional or impossible with other structures within a single language.

Northern Thai

```
NUM
                     DEM
                                            GEN
                     bà-mûan [(kèn) níi]
                                            cín *tìn [khỏon tùa]
bà-mûan [sɔɔn kèn]
                                            meat CLF GEN 2SG
                               CLF this
                     mango
         two CLF
mango
                                            'your meat'
                    'this mango'
'two mangos'
                                                 (tìn) [khỏon
                                                 CLF GEN
                                                 mii
                                                      năy?
                                                 have where
                                                 'Where is yours?'
```

More on this kind of pattern below

tŭa

2SG

Even if we pick up just NUM CLFs, like WALS, we cannot say, more often than not, simply that CLF is Absent, Optional, or Obligatory.

All the three structures involved in NUM CLF constuctions affect the use of CLFs

NUM CLF N'

CLF Different classifiers may behave differently

Murui (Witotoan; South-western Colombia; Wojtylak 2017)

General classifier

```
(4.30) jiibi-e-na<sub>o</sub> kue<sub>o:ADDRESSEE</sub> ine!<sub>PRED</sub> coca-CLF:G-N.S/A.TOP 1sg give 'Give me coca (general, in any form; usually in powder)!'
```

Specific classifiers

```
(4.29) jiibi na-iai
coca-CLF:TREE-PL
'coca trees'

(4.24) jiibi-koji
coca-CLF:CONTAINER.SMALL.ROUND
'small round coca container'
```

Wojtylak Katarzyna. 2017. A Grammar of Murui (Bue). James Cook University Ph.D. thesis

Now published as *A Grammar of Murui (Bue) A Witotoan Language from Northwest Amazonia* Series: <u>Brill's Studies in the Indigenous Languages of the Americas</u>, Volume: 15.

Murui Num CLFs

- General classifier forms modify a noun
- (4.17) [da je neki-na]_{NP} one-CLF:G chambira-CLF:TREE 'one *chambira* palm tree'
- (4.74) [da je jano-tava-na]_{NP:O} ati-d-e_{PRED} one-CLF:G small-CLF.REP:CHICKEN-N.S/A.TOP bring-LK-3 '(She) brought one small (chicken).'
- Specific classifier forms do not
- (4.33) da ñaiños beno-mo_{LOC} fiebi-t-e_{PRED} one-CLF:PR.F HERE.CLF:SP.PLACE-LOC stay-LK-3 'One (female) stayed here.'
- (3.49) mena kai_O kue-mo_{O:ADDRESSEE} ine!_{PRED} two-CLF:STICK 1sg-LOC give 'Give me two (stick-like forms, here: cigarettes)!'

NUM CLF N'

NUM Different numerals may affect the use of CLFs

In general lower numerals call for CLFs, while higher numerals may not

日本語 A: 昨日のシンポジウム、何人ぐらい来てた?

B': 150(人)ぐらいは、来てたやろ。(CLF optional)

B": 5*(人) ぐらいは、来てたやろ。 (CLF obligatory)

Only "one" and "two" are gender-marked in Portuguese and Rumainian, and only "one" is so marked, as in some Germanic languages, in the other major Romance languages (more on this below).

But in Mi'kmaq (Eastern Algonquian) CLFs are limited to numerals 6 and above:

- (11) a. na'n-ijig ji'nm-ug
 five-AGR man-PL
 'five men'
 b. *na'n te's-ijig ji'nm-ug
 five CL-AGR man-PL
- (12) a. *asugom-ijig ji'nm-ug six-AGR man-PL
 - b. asugom te's-ijig ji'nm-ug six CL-AGR man-PL 'six men'

Bale, Alan, and Jessica Coon. 2014. Classifiers are for numerals, not for nouns: Consequences for the mass/count distinction. *Linguistic Inquiry* 45:695–707.

You cannot simply say NUM CLF is obligatory or absent

• Numeral "one" behaves differently vis-a-via other numerals

Wuming Zhuang (武鸣壮語): "one" does not occur together with a CLF; CLF w/o a numeral is understood as "one CLF". "one CLF X"

tu2 pit7 ko:n3 pit7
CLF duck CLF pen
一只鸭子 支笔
yī zhǐ yāzi yī zhī bǐ 'one duck' 'one pen'

Cf.
sɔŋ1 ta8 nuəŋ4
two CLF y.sibling
两. 个 妹妹
liǎng gè mèimei
'two y. sisters'

sam1 pau4 luk9ŋe2 three CLF child 三 个 孩子 sān gè háizi 'three children'

Bahasa Indonesia

se-*(orang) guru
one-CLF teacher
'a/one teacher'

satu (orang) guruone CLF teacher'a/one teacher'

dua (orang) gurutwo CLF teacher'two teachers'

Ch'ol (Mayan), numerals of Mayan origin require a classifier (13)-(15a), while Spanish-based numerals block it (14)-(15b):

Ch'ol numeral (CLF obligatory)

(13) a. ux-p'ej tyumuty
three-CL egg
'three eggs'
b. *ux tyumuty
three egg

Spnaish loan numeral (No CLF)

(14) a. *nuebe-p'ej tyumuty
nine-CL egg
b. nuebe tyumuty
nine egg
'nine eggs'

- (15) a. Tyi k-mäñä ux-p'ej mansana. Ch'ol numeral (CLF obligatory)

 ASP lerg-buy three-CL apple

 'I bought three apples.'
 - b. Tyi k-mäñä nuebe mansana. Spanish loan numeral (No CLF) ASP lerg-buy nine apple
 'I bought nine apples.' (Bale and Coon 2014:701, 702)

Cf. Japanese

Native: hito-ri 'one (person)', huta-ri 'two (persons)', mit-ari 'three (persons), yot-tari, *it-tari, *mut-tari... Chinese loans: *iti-nin (iti-nin mae), *ni-nin (ni-nin mae/gumi), san-nin, *si-nin, go-ni, roku-nin... yo-nin

NUM CLF N'

N' In some languages CLF marking depends on the head N'.

Vietnamese

```
hai *(cái) chân vịt
two CL propeller (leg duck)
'two propellers'

hai (cái) nhà máy
two CL factory (house machine)
'two factories'

(CLF obligatory)

(CLF optional)

(CLF not possible)
```

two color

'two colors' (Simpson and Ngo 2018)

Hungarian

```
CLFs generally optional with count nouns
```

```
nyolc (szem) alma/gyöngy
eight CL apple/pearl
'eight apples/pearls'
```

But for a small number of nouns CLF-marking is obligatory

```
nyolc *(szem) kávé/bors/mák/rizs/búza/homok
eight CL coffee/pepper/poppy/rice/wheat/sand
'eight coffee beans/peppercorns/poppy seeds/grains of rice/wheat/sand'
```

hét (szál) gyertya/répa seven CL candle/carrot 'seven candles/carrots'

hét *(szál) fű/haj/széna/szalma seven CL grass/hair/hay/straw 'seven blades of grass/hair/hay/straw'

Dékány, Éva. 2022. Classifiers for nouns, classifiers for numerals. Proceedings of NELS 52 (pp.245-258).

Functional typology (機能類型論) reveals other large-scale limitations of form-based, holistic typology

Resarch method based on the form-function correlation patterns

CLF-(/Gender-) marking is not only construction-specific but also use/function-dependent

	NP use/Referring function (as HEAD of NP)	Modification use/Restriction (as MODIFIER of head)	
CLF-marking of DEM (this/that) (4 logical possibilities)) NO	NO	(English)
	YES	YES	(Mandarin)
	YES	NO	(Khorta)
	NO	YES	?

NP use

Modification use

English

This is mine.

This book is mine.

Mandarin

这本是我的。

Zhè běn shì wǒ de.

this CLF COP 1SG NMLZR

'This is mine.'

这 本 书 是 我的。 Zhè běn shū shì wǒ de.

this CLF book COP 1SG NMLZR

'This book is mine.'

Khorta#

i=ta hamnr lagi

this=CLF 1SG.NMLZR COP.1SG

'This is mine.'

i=*ta kitap ham^r lagi this=CLF book 1SG.NMLZR COP.1SG 'This book is mine.'

Cf. i kitap= ta hamAr lagi this book=CLF 1SG.NMLZR COP.1SG 'This book is mine.'

kitap= ta hamAr lagi book=CLF 1SG.NMLZR COP.1SG 'The book is mine.'

#Indo-Aryan spoken mainly in Jharkhand, India; courtey of Petra P. Paudyal

Similar patterns are	seen cross-constructionally	in a	single	language
			<u> </u>	

ern inal Numeral	NP use	Modification use
tǔa khây dâay bà-mûaŋ	(pôən khây dâay)	
2SG want take mango	[sɔ̃ɔŋ kɛ̀n]	bà-mûaŋ [sɔ̆ɔŋ <mark>kɛ̀n</mark>]
kìi kèn?	two CLF	mango two CLF
how.many CLF _{FRUIT}	'(I want) two.CLF'	'two mangos'
(11 1	13	

^{&#}x27;How many mangos do you want?'

Demonstrative

pə̂ən mii bà-mûaŋ	sìi kèn.	[kèn níi] wăan.	bà-mûaŋ	[(<mark>kèn</mark>) níi]
1SG have mango	four CLF	CLF this sweet	mango	CLF this
'I have four mangos.'		'This one is sweet.'	'this man	go'

$\boldsymbol{\mathcal{C}}$		\mathcal{L}
Genitive tìn níi cín khỏoŋ pôon. CLF _{PIECE} this meat GEN 1SG 'This is my meat.'	(tìn) [khɔ̃ɔŋ tŭa] CLF GEN 2SG mii năy? have where	cín *tìn [khɔ̃ɔŋ tŭa] meat CLF GEN 2SG 'your meat'
Again, you cannot simply say	'Where is yours?'	

Again, you cannot simply say CLFs are obligatory, optional, or absent in Northern Thai.

NO/OPT YES is not seen

What is generally believed to be optional marking is revealed to be untrue once usage patterns are considered

```
In Western Armenian, the presence or absence of a classifier is com-
pletely optional, as shown in (8). (For similar examples and observa-
tions, see Donabédian 1993.)
```

(8) yergu (had) dəgha two cl. boy 'two boys'

Bale and Coon (2014:699)

A: How many apples did you buy?

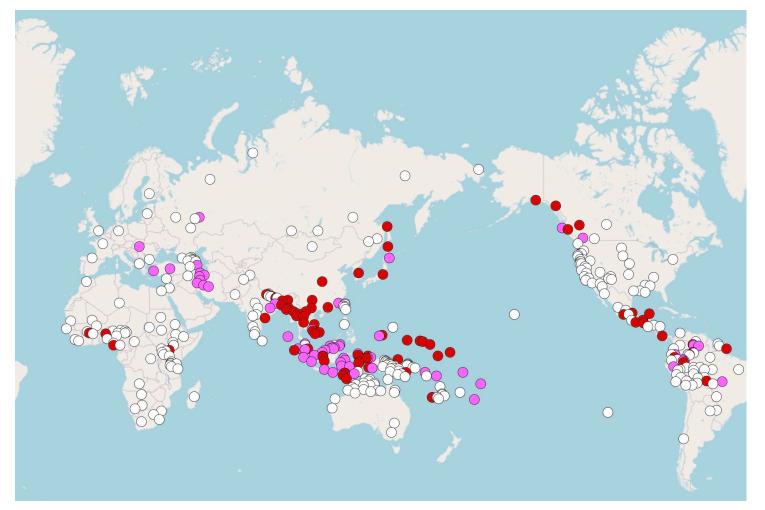
B: Armenian (Luiza Kloyan) 'I bought two (apples).'

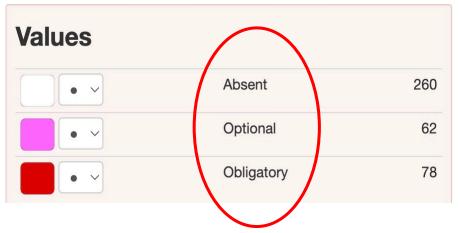
- (1) Erku (hat) xnjor gnec'i. (Optional in Modification use) two CLF apple buy.AOR.1sg
- (2) Erku *(hat) gnec'i. (Obligatory in NP use) two CLF buy.AOR.1SG
- (3) Erku (hat). (Optional in NP use--may be absent in informal speech)

Obligatory CLF-use seen in NP use; similar pattern also seen Turkish (Yu Kuribayashi, Yui Suzuki p.c.)

As seen above, form-based, holistic typology assumed in WALS has numerous problems

Feature 55A: Numeral Classifiers





David Gil. 2013. Numeral Classifiers. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) WALS Online (v2020.3) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.7385533 (Available online at http://wals.info/chapter/55, Accessed on 2023-06-14.)

On the other hand, functional typology, seeking form-function correlations, reveals very fundamental methodological issues associated with the past approaches to classifier-/gender-marking and directs us to a more adequate approach.

Methodological problems in the current treatments of gender and CLFs

Only one of the two major uses/functions of CLF- and gender-marking is seriously considered

Thai numeral	NP use/	Modification use/
khun tôŋkaan mǎa	Referring function	Restricting function
2SG want dog kìi tua?	(chǎn tôŋkaan) [sìi tua] .	mǎa [sìi tua]
how.many CLF _{BODY} 'How many dogs do you want?	four CLF _{BODY} (I want) 'four.CLF.'	dog four CLF 'four dogs'
Portuguese "adjective"		
Qual é o seu carro?	O branco.	carro branco
'Which is your car?'	'The.M white.M.'	'white.M car'
Qual é a sua casa? 'Which is your house?'	A branca. 'The.F white.F.'	casa branca 'white.F house'

NP use/Referential function is treated secondarily as arising from the deletion of a head noun of the modification structure—similar to the treatment of so-called headless relative clauses.

Cf. Greenberg (1974:19) "It is indeed universal in languages with numeral classifier constructions that the head noun may be deleted when it has been either previously mentioned or can be supplied from the non-linguistic context."

Based on this (problematic) methodological perspective, the following kind of characterizations have been made

Aikhenvald (2019: p.9) Genders and classifiers. Oxford: OUP

TABLE 2. Noun categorization devices and their scope

Түре	Scope	What is categorized	
I Gender	Attributive NP or clause	Head noun, A/S or S/O; oblique	
II Numeral classifiers	Numeral/quantifier NP	Head noun	
III Noun classifiers	Noun	Head noun	
IV Possessive classifiers	Possessive NP	Possessed noun	
V Verbal classifiers	Clause	S/O or oblique	
VI Locative classifiers	Adpositional NP	Noun referring to a location	
VII Deictic classifiers	Attributive NP	Head noun ?	

But there are both classifier- and gender-marked forms, ranging from numerals to verbal-based nominalizations, that do/can not modify a head noun.

NP use
(with no head noun)

YES

Tôi muốn[quả này]

1SG want CLF this
'I want this (e.g., apple).'

Modification use
(with a head noun)

NO

táo [*quả này]

apple CLF this
'this apple'

Cf. Vietnamese,

Summary of the problems in the current methodology

Distribution patterns of NUM+CLF/Gender

	NP use	Modification use
	(i) NO	NO (English)
	→ (ii) OPT	NO (Marathi)
	(iii) OPT	OPT (Jingpo)
Data?	→ (iv) YES	NO (Sikuani)
	\longrightarrow (v) YES	OPT (Halkomelem, Tariana, Khorta)
	(vi) YES ←	YES (Mandarin, Thai, Newar, Korean, Japanese; Gender in Portuguese and other IE lgs)

- This type of diverse marking patterns has not been considered.
- Wrong perspective; modification pattern considered to to the starting point of analysis
- If the function of CLFs/Genders were to categorize (the referent) of a head noun, (ii) and (iv) would be hard to explain; (v) would also be problematic.
- Missing the primacy of gender/CLF marking in NP use, as in the case of ordinary nominalizations.

Classifier-marked forms that cannot modify

Numerals

Marathi classifier marking

```
ek-jaN aal-aa one-CLF came.M 'One (HUMAN) came.'
```

```
a'. ek *-jaN mulgaa
one CLF boy
'one boy'
```

Demonstratives

Vietnamese classifier marking on demonstratives

```
a. tôi thích [con này]. NP use
I like CLF this
'I like this (e.g. a cat).'

a'. tôi thích con mèo [*con này]. Modification use
I like CLF cat CLF this
'I like this cat.'
```

Genitive

Northern Thai

NP use

[(tin) [khɔɔŋ pə̂ən]]
CLF GEN 1SG
mii nay?
have where
'Where is mine?'

Modification use

cín [*tin [khɔɔŋ pə̂ən]
meat CLF GEN 1SG
mii nǎy?
have where
'Where is my meat?'

Adjectives, V-based nominalizations

'a small dog'

Zauzou若柔語(Tibeto-Burman, Loloish; southern China; (courtesy of Tetsuya Miyagishi)

```
Stative ("Adjectival")

a. [na<sup>55</sup> 7õ<sup>33</sup>] nɛ<sup>31</sup> nu<sup>33</sup> khyi<sup>31</sup> 7õ<sup>33</sup> NP use

small CLF TOP 1SG dog CLF

'The small one is my dog.'

a'. [na<sup>55</sup> *7õ<sup>33</sup>] khyi<sup>31</sup> Modification use small CLF dog
```

Eventive

- b. [ʔa³³ɲi³³ κi⁵⁵ ʔõ³³] nε³¹ ŋu³³ (zε³³) ɣou⁵³phõ³⁵
 NP use yesterday die CLF TOP 1SG GEN cock
 'The one that died yesterday is my cock.'
- b.' [?a³³ni³³ si⁵⁵ *?õ³³] you⁵³phõ³⁵ Modification use yesterday die CLF cock 'the cock that died yesterday'

Gender-marked forms that cannot modify

Numerals

Telugu (Dravidian; courtesy of Niranjan Uppoor)

```
NP use
                                           Modification use
                                        a'. okka/*okka<mark>du</mark>
a. okkadu
                                                           abba:yi
              vacca:du
                                            one/one.M
                                                           boy
  one.M
              came.3M.SG
                                            'one boy'
  'One.M came'.
b. okkarti vaccindi
                                        b'. okka/*okkarti
                                                           amma:yi
   one.F came.3F.SG
                                            one/ one.F
                                                           girl
   'One.F came.'
                                            'one girl'
```

c. okkaţi vaccindi one.N came.3F.SG 'One.N (e.g. a dog) came.'

c'. okka/*okkaţi kukkalu one/one.N dog' 'one dog'

Demonstratives

Kannada (Dravidian) gender marking (courtesy of Niranjan Uppoor)

- a. a-vanu nanna geleya that-M my friend Lit. 'That masculine near the third person or someone not present is my friend.' 'He who is away from me is my friend.'
- a'. aa geleya *a-vanu geleya that friend that-M friend 'that friend'
- b. i-valu nanna gelati this-F my friend Lit. 'This feminine one near me is my friend.' 'She who is near me is my friend.'
- b'. ii huḍugi *i-vaļu huḍugi this girl this-F girl 'this girl'

- c. a-du nanna mane that-N my house Lit. 'That neuter one is my house.' 'That is my house.'
- c'. aa mane *a-du mane that house that-N house

Genitives

Kannada (Dravidian) gender marking (courtesy of S.N. Sridhar)

- a. Aa seevaka(nu) [nanna tande-ya-vanu]. that servant (M) my father-GEN-M 'That manservant is my father's.'
- a'. [nanna tande-ya(*-vanu)] seevaka(nu) my father-NMLZR-M servant(M) 'my father's manservant'
- b. Aa seevaki [nanna tande-ya-vaļu]. that maid my father-GEN-F 'That maid is my father's.'
- b'. [nanna tande-ya(*-valu)] seevaki my father-NMLZR-F maid 'my father's maid'

Adjectives, V-based nominalizations

Kannada gender marking (courtesy of Niranjan Uppoor)

Stative ("Adjectival")

- a. [[Ettarad-a]-vanu] nanna maga.
 tall-NMLZR-M my son
 'The tall one is my son.'
- a.' [[[ettarad-a](*-vanu)] huḍuga] tall-NMLZR-M boy 'tall boy'
- b. [[Ettarad-a]-vaļu] nanna magaļu.
 tall-NMLZR-F my daughter
 'The tall one is my daughter.'
- b'. [[[[ettarad-a](*-vaļu)] huḍugi] tall-NMLZR-F girl 'tall girl'

Eventive

- c. [Band-a-vanu] nanna geleya. came-NMLZR-M my friend 'The one.M who came is my friend.'
- c'. [[Band-a (*-vanu)] huḍuga] nanna geḷeya. came-NMLZR(*-M) boy my friend 'The boy who came is my friend'
- d. [Band-a-vaļu] nanna geļati. came-NMLZR-F my friend 'The one.F who came is my friend'
- d'. [[Band-a (*-valu)] huḍugi] manna gelati. came-NMLZR(*-F) girl my friend 'The girl who came is my friend.'

Nominalization analysis of CLF/Gender marking

Must analyze the forms that cannot/do not modify first

```
Marathi classifier marking

NP use/Referring function

a. don-jaN

two-CLF<sub>HUMAN</sub>

'Two (HUMAN(.Epicene)) came.'

Modification use/Restricting function

a'. don *-jaN mule

two CLF<sub>HUMAN</sub>

'two boys'
```

CLF -jaN nominalizes the numeral "don" deriving an N-based NMLZ that denotes "thing" entities embodying the concept of the quantitative unit "two", and which fall in the HUMAN category (e.g., two boys, two men, two teachers).

Contrary to the conventional wisdom, (numeral) CLFs do not classify the referent of a head noun, which may not exist. Instead, they operate on the numeral (and others) (with) which they mark/form constituency.

Same applies to Gender marking

Must analyze the forms that cannot/do not modify first

Telugu gender marking

NP use/Referring function

a. okka-rti vaccindi one-F came.3F.SG 'One.F came.' Modification use/Restricting function

a'. okka *-rti amma:yi one -F woman 'one woman'

-rti nominalizes the numeral "okka" deriving an N-based nominalization that denotes "thing" entities embodying the concept of the quantitative unit "one", and which fall in the FEMININE category (e.g. one girl, one woman, one female teacher)

Contrary to Corbett's (1991, others) analysis, gender marking is not triggered by an agreement controlling head noun, which may not exist. Instead, genders, just like CLFs, operate on numerals (and others) to/with which they mark/form constituency.

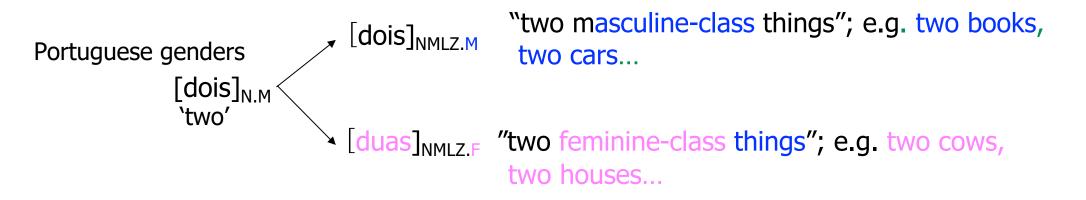
Same analysis holds for those gender/CLF forms that may also modify, as in Vietnamese numeral CLFs

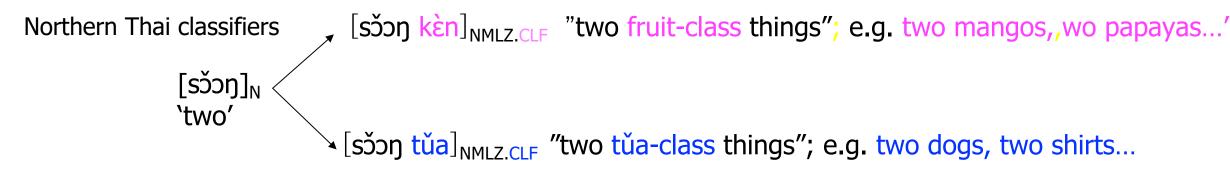
Tôi muốn [ba quả] [[ba quả] táo] (simply a use of N-based NMLZ) 'I want three (e.g., apples).'

quả nominalizes the numeral "three" deriving an N-based nominalization that denotes "thing" entities embodying the concept of the quantitative unit "three", and which fall in the quả/fruit category (e.g., apples, mangos, oranges)

Summary of the nominalizaiton analysis of gender-/classifier-marking

Genders and classifiers are classifying nominalizers that yield/derive lexical/grammatical nominalizations that denote "things" and thing-ike entities that are in metonymic relation to the concept of the base structure and that are classified according to the nominal classes of the language.





Context determines what these nominalized structures actually denote when used in discourse

Northern Thai classifiers Portuguese gender Context 1 Context 1 tǔa khây dâay Qual é o seu carro? 2SG want take 'Which is your car?' măa kìi tŭa? branco. dog how.many CLF [sɔɔn tŭa]. the.M white.M 'How many dogs two CLF 'The white (one).' (car) do you want? 'Two.' (dogs Context 2 Context 2 tŭa khây dâay sîn Qual é o seu chapéo ? 2SG want take sarong 'Which is s your hat?' kìi tŭa? branco. [sɔɔn tua]. how.many CLF the.M white.M two CLF 'How many sarongs 'The white (one).' (hat) 'Two.' (sarongs) do you want?'

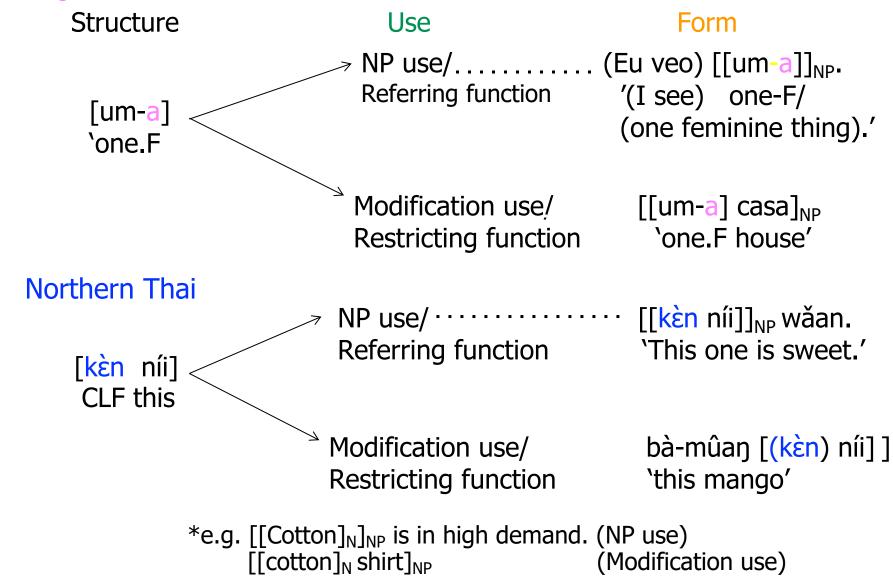
Cf. Ordinary metonymy:

Speaker A: I heard the health benefit of red wine/orange juice...

Speaker B: Oh, we should then drink a glass a day. (glass > red wine/orange juice)...

Just like ordinary nouns* grammatical NMLZs have two uses

Portuguese



Evidence supporting our analysis: The nominalization function of gender-/classifier-marking

Portuguese

Lexical (noun-forming) nominalization by genders

```
'kill/assasinate' kill-er -er (non-classifying nominalizer)

assessin-o
'(male) killer'

assessin-a
'(female) killer'
```

amigo/amiga 'friend.M/friend.F', filho/filha 'son/daughter', tio/tia 'uncle/aunt'

```
japonês > japonesa 'female Japanese' professor > professora 'female teacher', freguês > freguesa 'female customer', camponês > camponesa 'female farmer', marquês > marquesa 'marquise'
```

```
doutor > doutora 'female doctor', matador > matadora 'female killer', cantor cantora 'female singer'
```

Grammatical (not nouns) nominalization by genders

```
O [sentado na cadeira] é meu pai
ART.M sitting.M in.the chair is my father
'The one sitting in the chair is my father.'
```

```
A [assessin-a] é minha mãe

A [sentada na cadeira] é minha mãe.

ART.F sitting.F in.the chair is my mother

'The one sitting in the chair is my mother.'
```

Cf. Non-classifying nominalization English

[Sitting in the chair] is none other than my old man/woman.

Lexical: (I read it in one) [sitting]

Northern Thai

Lexical (noun-forming) nominalizations by CLFs

V-based N-based

khon khàp lót 'driver' khon nua 'northern (Thai) people'

CLF drive(v) car CLF north

hoon hian 'school' hoon yaa 'hospital'

CLF study(v) CLF medicine

kan thảy 'plow' kèn tảa 'eyeball'

CLF plow (v) CLF eye

láan káa 'shop' nîw tĭin 'toe'

CLF sell (v) CLF foot

Cf. V-based and N-based English nominalizations

sing--er London-er

speak-er village-r

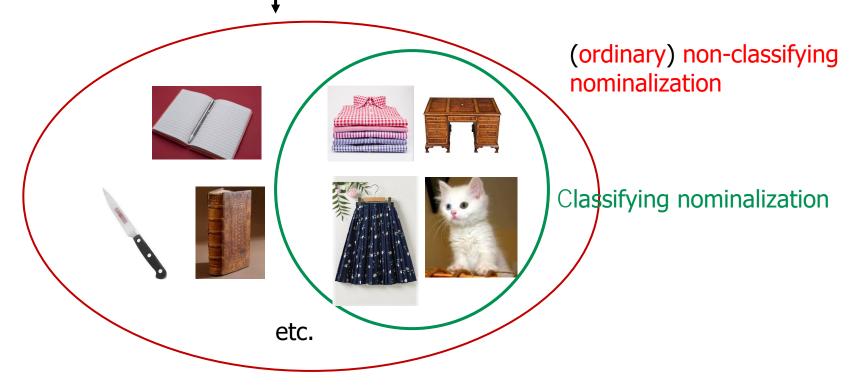
wash-er YouTube-r

Grammatical nominalizations

Northern Thai

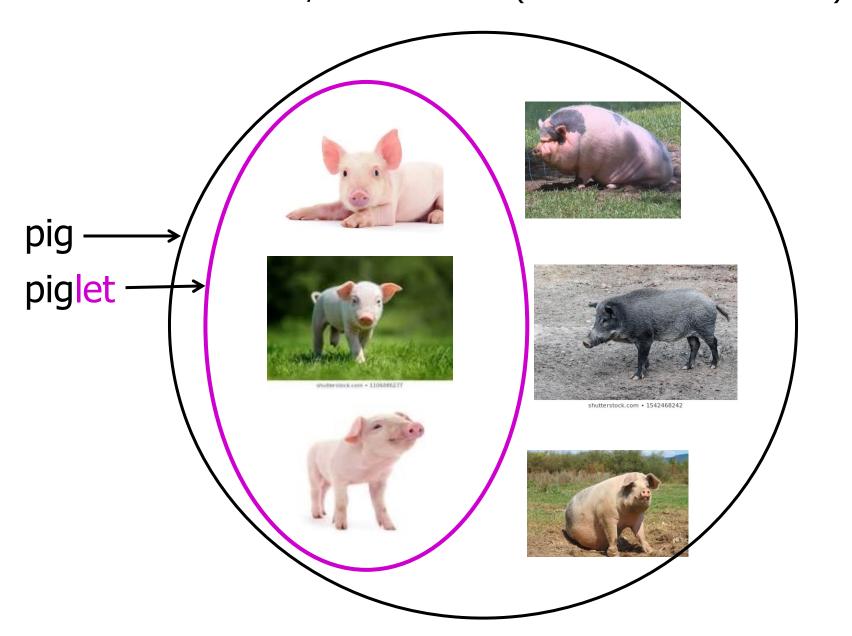
tửa [tîi pôn sûu tàwaa] péεŋ khanàat NMLZR 1SG buy yesterday very expensive '[What I bought yesterday] was very expensive.'

NMLZR=nominalizer



CLF-marking derives a new nominal structure with a new set of denotations, just like a diminutive derivation such as,

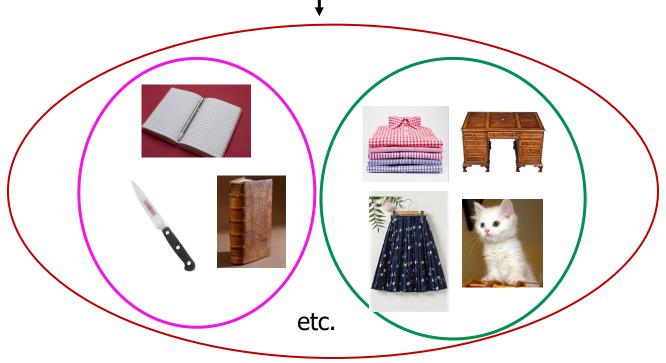
Diminutive derivation/nominalization (=derivation of nominals)



Grammatical nominalizations

```
Northern Thai

| tŭa | ttii | pôn sûu tàwaa | pέεη khanàat |
| NMLZR 1SG buy yesterday very expensive what I bought yesterday was very expensive | the control of the cont
```



Another new struture with a new set of denotations

What is known as a relative clause is no more than a modification use of V-based grammatical nominalizations

```
NP use
                    [(tǔa) [[tîi pôn sûu tàwaa] péɛŋ khanàat]<sub>NMLZ</sub>]<sub>NMLZ</sub>
                     CLF NMZLR 1SG buy yesterday very expensive
                   'What/The one I bought yesterday was very expensive.'
Structure: V-based grammatical nominalizations
             pôn s<del>ûu</del> tàwaa] péɛŋ khanàat]<sub>NMLZ</sub>]<sub>NMLZ</sub>
[(tǔa) [tîi
  CLF NMZLR 1SG buy yesterday very expensive
                   Modification use
                  ≯sîn [(tǔa) [tîi
                                   pôn s<del>ûu</del> tàwaa] péen khanàat <sub>NMLZ</sub> <sub>NMLZ</sub>
                   skirt CLF NMZLR 1SG buy yesterday very expensive
                    'The skirt that I bought yesterday was very expensive.'
```

There is nothing like a relative clause as an independent structure apart from the use of a V-based grammatical nominalization as a modifier.

Dynamicization of functional typology

Accounting for crosslinguistic marking patterns

Telugu (Dravidian; courtesy of Niranjan Uppoor)

```
a'. okka/*okkadu
                                                         abba:yi
a. okkadu
             vacca:du
                                          one/one.M
                                                         boy
  one.M
              came.3M.SG
                                          'one boy'
  'One.M came'.
b. okkarti
          vaccindi
                                      b'. okka/*okkarti
                                                         amma:yi
   one.F came.3F.SG
                                          one/ one.F
                                                         girl
   'One.F came.'
                                          'one girl'
c. okkati
           vaccindi
           came.3F.SG
   one.N
                                      c'. okka/*okkaţi
                                                        kukkalu
  'One.N (e.g. a dog) came.'
                                          one/one.N
                                                         doq'
                                          'one dog'
```

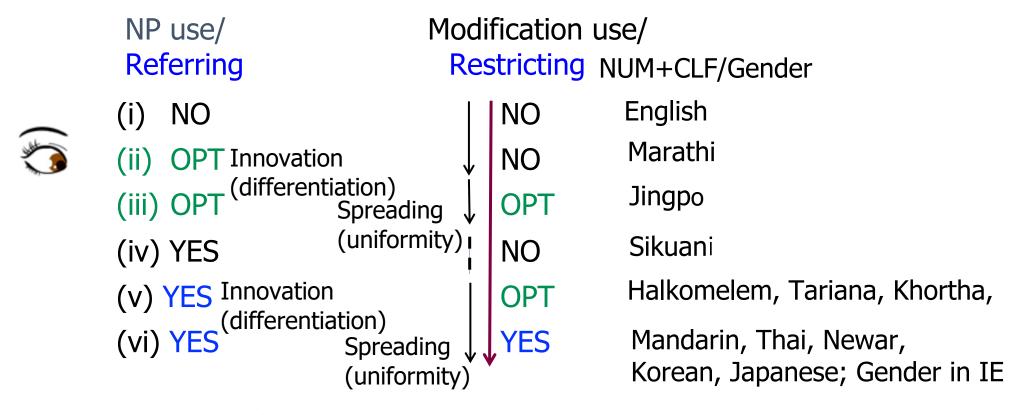
There is no logical reason that gender-marked numerals cannot modify a head noun, as in Telugu, in view of

Same can be said about restirctions in CLF-marking

Vietnamese	NP use		Modification use	Early stage
	Tôi muốn [quả này] 1SG want CLF this 'I want this.'	\	táo [*quả này] apple CLF this `this apple'	NO
Northern Thai				
	[kɛ̀n níi] wǎan. CLF this sweet 'This one is sweet.'	-	bà-mûaŋ [(kɛ̀n) ní] mango CLF this 'this mango'	OPT
Zauzou (Loloish)				
this	a ³³ <mark>ʔõ³¹</mark>] nɛ³¹ ŋu33 pe33 ze33 khyui³¹ s CLF TOP 1.PL.EXCL GEN dog nis is our dog.'		khyui ³¹ [ʔa ³³ <mark>ʔõ³¹</mark>] dog this CLF 'this dog'	YES Late stage of change

Plausible explanation is a diachronic one, namely gender-/CLF-marking spreads from NP-use domain to the modification context across time, but the timing of the spread varies across languages.

Dynamic functional-typological perspective



Starting point: the NP-use forms must be the starting point of analysis.

The other crosslinguistic marking patterns represent the developmental stages of the CLF/Gender marking. Like,

What motivates this kind of alternating pattern of change of innovation and spreading/leveling?

Pioneer functionalists

The more economical or more abundant use of linguistic means of expressing a thought is determined by the need... Everywhere we find modes of expression forced into existence which contain only just so much as is requisite to their being understood.

> Paul, Hermann 1880/1889. Prinzipien der Sprachgeschichte (Halle: Max Niemeyer; Translation by H.A. Strong 1881/1891 and by Hans Hock 1991)

In order to understand how and why a language changes, the linguist must keep in mind two ever-present and antinomic factors: first, the requirements of communication, "Hearer's the need for the speaker to convey his message, and second, the principle of least effort, which makes him restrict his output of energy, both mental and physical, to the minimum compatible with achieving his ends.

economy" "Speaker's economy"

Martinet, André 1962. A Functional View of Language.

Oxford: Clarendon Press. P.139

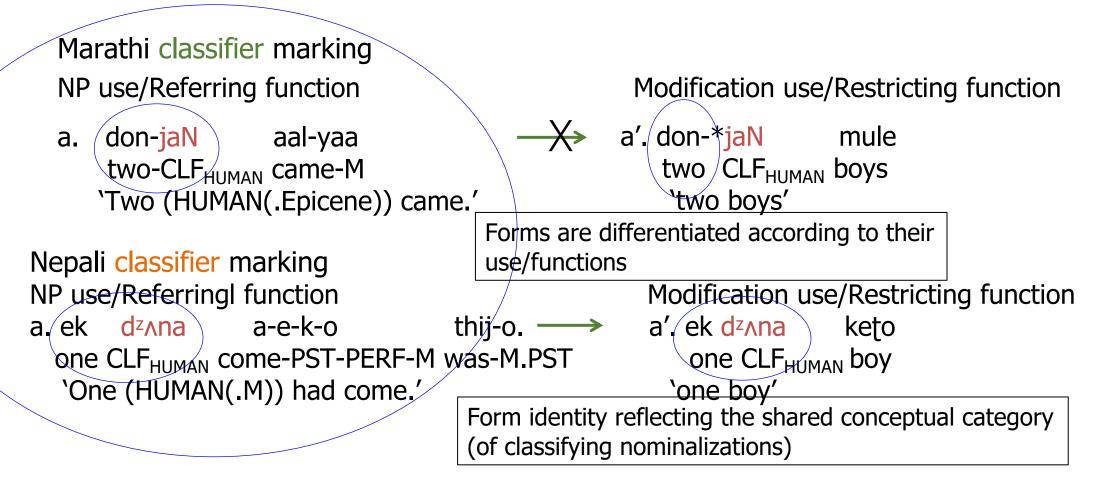
Speaker's economy > Simplification/form uniformity (S) > Reflecting the unity/oneness of the conceptual category involved

Hearer's economy > Diversification/formal variation (D) >

Reflecting differences in communicative function (e.g., referring or restricting the denotation?)

Crosslinguistic difference in the marking pattern

Starting point is the forms that cannot/do not modify



Nepali is a step ahead of Marathi in the spreading of the CLF marking to the modification context. The former has achieved the speaker's economy, while the latter maintains the hearer's economy.

Same pattern of development seen in gender marking

Telugu gender marking (as in Tamil and Malayalam)

```
NP use/Referringl function
a. okka-rti vaccindi one-F came.3F.SG  
'One.F came.'

Modification use/Restricting function a'. okka *-rti amma:yi one -F woman one woman'
```

Kannada gender marking

NP use/Referring function

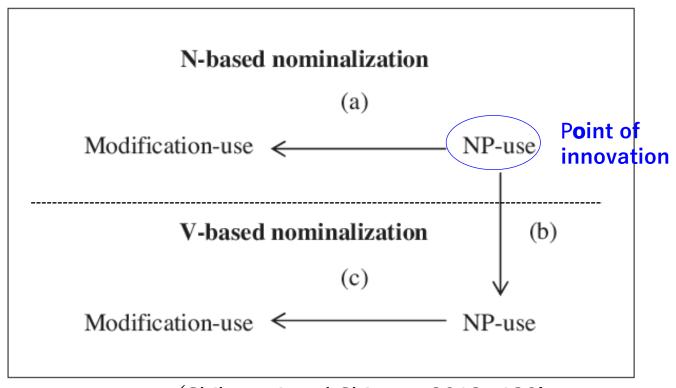
a. obba-lu bandalu
one-F came.3FSG
'One.F came.'

Modification use/Restricting
a'. obba-lu huḍugi
one-F girl
'one girl'

Kannada is a step ahead of Telugu in the spreading of gender marking to the modification context to achieve the speaker's economy.

Manner of spreading of nominalization markers

Figure 1: Directions of spread of nominalization markers

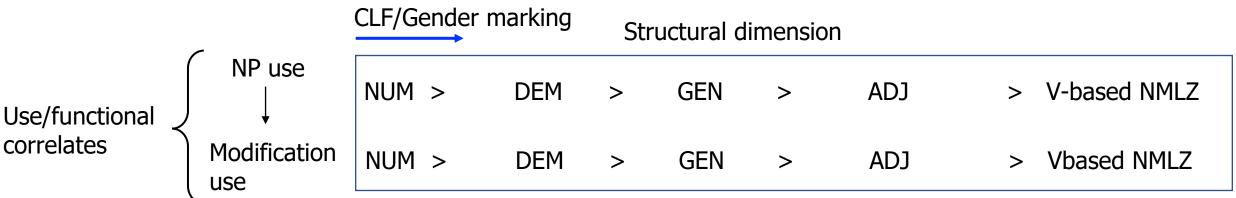


(Shibatani and Shigeno 2013: 120)

Shibatani, Masayoshi and Hiromi Shigeno. 2013. Amami nominalizations. *International Journal of Okinawan Studies*, Vol. 7. 107-139.

Patterns of spread of nominalization markers





Constraints on synchronic distributions of nominalizaton markers, incl CLFs/Genders

- Cross-dialectal/linguistic variations in CLF-/Gender-marking are expected due to communicative dynamism
- (Overt) CLF-/Gender-marking is likely found (in the NP-use context) at the left edge of the hierarchy
- Universal constraints can drawn along the two dimensions, similar to the Keenan-Comrie attempt on those on relative clause constructions

Predications on diachronic developments of nominalization markers, incl CLFs/Genders

- A historically older form of a language shows a less developed pattern of CLF-/Gender-marking
- Cross-dialectal/linguistic variations are expected, reflecting the historical pattern above

N-based V-based (less developed)

NP use: NUM > DEM > GEN > ADJ > V-based nominalization

Mod use: NUM > DEM > GEN > ADJ > V-based nominalization

(less developed)

Assamese and Assamese Bengali (Case study by Gitanjali Bez, Gauhati University)

Old Assamese (less developed than both Nalbariya dialect and Modern Standard Assamese)

NP use: OPTNUM > OPTDEM > *GEN > *ADJ > *V-based nominalization

Mod use: OPTNUM > OPTDEM > *GEN > *ADJ > *V-based nominalization

Nalbariya dialect (reflects a less developed pattern than Modern Standard Assamese)

NP use: NUM > DEM >*OPTGEN >*OPTADJ >*OPTV-based nominalization

Mod use: NUM > OPTDEM > *GEN > *ADJ > *V-based nominalization

Modern Standard Assamese

NP use: NUM > DEM > *OPTGEN > *OPTADJ >*OPTV-based nominalization

Mod use: NUM > OPTDEM > OPTGEN > *ADJ > *V-based nominalization

Cross-linguiscally

Bengali (similar to Nalbar va Assamese

NP use: NUM > DEM > *OPTGEN > *OPTADJ > *V-based nominalization

Mod use: NUM > *DEM > *GEN > *ADJ > *V-based nominalization

Development of gender marking in major Dravidian languages

```
Telugu, Tamil, Malayalam (K V Subbarao, Niranjan Uppoor, Rajendran S)
```

```
NP use: NUM > DEM > GEN > ADJ > V-based (participle)
```

Mod use: *NUM > *DEM > *GEN > *ADJ > *V-based (participle)

Kannada (Niranjan Uppoor)

```
NP use: NUM > DEM > GEN > ADJ > V-based (participle)
```

Mod use: NUM > *DEM > *GEN > *ADJ > *V-based (participle)

No mark=CLF-marking obligatory, OPT=optional marking, *OPT="optional"; marking when the referent is definite, *X=markingof X not possible

Implications for non-gender/non-CLF languages

Describing "adnominals" (noun modifiers) in English?

Numeral three dogs

Demonstrative this dog, that dog

Genitive/Possessive my dog, my neighbor's dog

Adjective white dog,

Relative clause The dog [which is barking] is yours.

Noun cotton shirt, London parks

Why do these forms belonging to different parts of speech (word classes)/constructions all modify a noun? Do they have anything in common, which unifies them and allows them to function as a noun modifier?

What CLF/Gender languages suggest

These modifiers
are all
grammatical <
nominalizations
in modification
use

Numeral three dogs

Demonstrative this dog, that dog

Genitive/Possessive my dog, my neighbor's dog

Adjective white dog,

Relative clause The dog [which is barking] is yours.

Noun cotton shirt, London parks

We have seen that these "adnominals" from numerals to relative clauses are a unified phenomenon as indicated by gender-/classifier-marking; they are all nominalized structures. What modifies a noun isn't a NUM, DEM, etc. but a derived nominal structure, similar to nouns (e.g., cotton, London).

Some internal evidence for the nominalization analysis of English adnominals

Basic meanings of numerals

(1) quantitative units

(Talking about numbers denoting quantitative units)

Two is larger than one. (quantitative unit > a singular verb form)

Dois é maior que um. (Portuguese)

(2) numeral letters



(Looking at the above numeral letters)

One is blue and two is red. (a letter "2" > a singular verb form) Um é azul e dois é vermelho.

But when numerals nominalize, they denote "things" (metonymic to the basic quantitative units; "things" embodying the basic numeral concepts)

```
Speaker A: Are there cows out there?
            Há vacas ali? (Portuguese; feminine denotation)
Speaker B: Yes, I see three.
                                                 two "things" > a plural verb form
           Sim, eu vejo três.
           One is eating grass, and two are drinking water.
           Uma está comendo capim e duas estão bebendo água.
                                           two feminine "things" > a F plural
Speaker A: Are there cars out there?
                                                                verb form
            Há carros ali? (masculine denotation)
 Speaker B: Yes, I see three.
          B: Sim, eu vejo três,
                                              two "things"> a plural verb form
             One is blue and two are red.
```

dois são vermelhos.

two masculine "things" > a M plural verb form

Um é azul e

Traditional analysis of genitives; as noun declension/case inflection

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. A Comprehensive Grammar of the English Language. Essex: Longman.

P. 336

SUBJECTIVE	I	we	he	she	distant	they	who	
OBJECTIVE	ine u	us	you	him	her	it	them	who(m)
GENITIVE determinative independent	Modif my mine	cation our ours	use for your yours	ms his	her hers	its	their theirs	whose

One immediate probrem with this: what about *John's*, *my neighboor's*, which are not pronouns?

Many languages treat these as a unified phenomenon:

mǎa [khɔ̃ɔŋ tǔa] 'your dog', mǎa [khɔ̃ɔŋ Kultida] 'Kultida's dog' (Northern Thai)

Major problems with the traditional analysis of genitives/possessives

SUBJECTIVE	I	we	you	he	she	it	they	who
OBJECTIVE	me	us		him	her		them	who(m)
GENITIVE determinative independent	my mine	our ours	your yours	his	her hers	its	their theirs	whose

These two groups of forms differ significantly in both semantic and syntactic properties, warranting two separate treatments

Semantically, Subective (or Nominative) and Objective (or Accusative) forms (I/me, you/you, he/him, etc.), on the one hand, and Possessive (or Genitive) forms (my/mine, your/yours, his, etc.), on the other, denote/refer to different entities

I (pôən) like you. Look at me (pôən)! Refer to the speaker



Look at mine (khɔ̃ɔŋ pə̂ən)!

Refers to what is metonymically/ intimately connected to the speaker







Etc.

Syntactically, Subjective/Objective forms and Genitive forms are distributed differently

Subject position
He/*him/his/John's is big.

Subjective forms occur only in subject position.

Object forms cannot; but Genitive/possessive forms can.

Object position
Look at *he/him/his/Johns.

Object forms occur only in object position
Subject forms cannot; but Genitive/possesive
forms can.

Modifier position
*he/*him/his/John's book

- Inflected forms occur only in specific syntactic positions
- Genitives/Possesives, on the other hand, denote things, like ordinary nouns—they are grammatical nominalizations marked by inflectional morphology or cliticization of /z/ (written as 's). Syntactically, they occur anywhere like ordinary nouns.

We see this clearly in CLF languages because things denoted by genitives may be classified;

```
Northern Thai

[khɔ̃ɔŋ pə̂ən] pɛ́εŋ khanàat

kèn

Northern Thai

[khɔ̃ɔŋ pə̂ən] pɛ́εŋ khanàat

'Mine is very expensive.'
```

Implications for classifier/gender languages

Nominalization does not necessarily exhibit overt marking: e.g. [three]_N 'three' > [[three]_N]_{NMI 7} 'three (things)'



'three houses'

There can be classiffying nominalizations that do not exhibit overt marking.

Of the major Romance languages Portuguese and Romanian both morphologically distinguish gender for "one" and "two" but not for higher numbers; the others do so only for "one".

```
overtly gendered/classifying
Esses dois são meus.

'These two.M are mine.M..'

Essas duas são minhas.

'These two.F are mine.F.'
                                                                                   dois carros
                                                                                  'two cars'
                                                                                  duas casas
                                                                                   'two houses'
                              covertly gendered/classifying
             três são meus.
                                                                                    três carros
'These.M three are mine.M.'

Essas três são minhas.

'These.F three are mine.F.'
                                                                                    'three cars'
                                                                                    três casas
```

One should not assume that a given form is not classifying simply because morphological marking is absent.