The coding of (in)definiteness in northern Vanuatu
Anaphora, specificity, topicality

Alexandre FRANÇOIS
CNRS–LACITO ¹ – A.N.U.
alex.francois.cnrs@gmail.com

1. Indefinites in northern Vanuatu: the question

1.1 The languages of the Banks & Torres Islands

Banks and Torres Islands, northern Vanuatu: 17 languages, all Oceanic. Their grammatical and semantic structures are generally parallel or “isomorphic” (François 2011)... And yet the organisation of (in)definiteness is quite diverse amongst them.

My corpus: (a) grammatical, comparative questionnaire
(b) recordings of spontaneous speech: 104 h, incl. {50 h = 389 narratives} in 21 languages.
→ Focus on one language: Hiw.

¹ This work was first presented in LACITO’s research group For a linguistic typology of (in)definiteness, Oct 2015.
1.2 When definiteness is underspecified

The semantic feature \(\pm\)definite] is **encoded** systematically in some languages (Romance, Germanic, Greek, Arabic…) but is left **underspecified** in others (Russian, Mandarin, Japanese…).

**DEFINITE** expression:
expression construing a referent X with the specific instruction, given to the addressee, to retrieve the identity of that X among the already known (or identifiable) referents of their representational world.
- e.g. *She* fell on the road; *the* children are quiet; *I know the* harpist

**INDEFINITE** expression:
expression construing a referent X with the specific instruction, given to the addressee, to create a new X, without trying to equate it with an already known (or identifiable) referent.
- e.g. *Someone* fell on the road; *some* children are quiet; *I know a* harpist


What about the Oceanic languages of northern Vanuatu?
At first glance, the contrast **definite** – **indefinite** is left underspecified:

1. **Hiw**
   - *Nine yō me:* see ART eel crawl out hither
     - ‘Suddenly he saw **AN** EEL crawling out to him.’ [Eel_10]

2. **Tom**
   - *Ne meřempē pe noke meñmeñon ti,*
     - ‘He said: “**THE** EEL I’ve been feeding, THAT EEL has now become a boy!”’ [Eel_47]

Languages of northern Vanuatu have a noun article (usually /\(n\)/ < POc *na*), which is non-specific with respect to definiteness. Its function is that of a determiner, which allows common nouns to form a valid referential phrase. \[\Rightarrow\] article = the D in a ‘DP’ [formal syntax]

The articles of Germanic & Romance languages are really **portmanteau** forms stacking up several functions { DET, \(\pm\)DEF, gender, number }... The common noun article of N. Vanuatu lgs has essentially one function, namely DET.

NB: /\(n\)/ is only found with **COMMON** N (all non-human N + some human N) \(\neq\) “personal nouns” (=individuated human) take *i* or Ø [François 2007]

The /\(n\)/ article may read as **\(\pm\)**definite, **\(\pm\)**specific, **\(\pm\)**generic:

3. **Hiw**
   - *NE TEMET tati yō vegyaye tom NE TAYO giy NE TĒN ti.*
     - ‘The ghost didn’t realise that *someone* had been digging the **ground**.’ [Brothers_09]

4. **Hiw**
   - *Tuwtōw, tekīwa tati gengon NE TAYO ti.*
     - ‘In the olden days, there was no cannibalism.’ [Stories.001]
In sum, NV languages do not encode definiteness on their articles... But do they encode it at all?

2 Definite and indefinite in Hiw

Nevertheless, some morphemes do exist, that encode such values as [+DEF] or [+SPEC].

2.1 The anaphoric in

(5) Hiw Se toge vën vën, ne qin tamesö NOT mét. 3pl stay:PL DUR DUR ART person old INDF die:NPL

Ne qin tamesö IN mét : teñwafe ve toge ART person old ANAPH die:NPL HUM:M:PL IPFV stay:PL

ve wane řog ne mëse. IPFV drink.kava PREP ART death

‘They were living like that, when one day AN old man passed away. As THAT man had died, the men of the village came together to drink kava at his wake.’ [Hades.08]

in ANAPHORIC (therefore +DEF): always points to a referent previously mentioned in the context, typically in the preceding clause. ‘that X in question, the aforementioned X’

Origin of in = associative noun linker i + suffix -n ‘3sg:ANAPH’ ⇒ /in/ = ‘of it, its’

(6) ne vegevag’ i meřempē → ne vegevag’ i-n ART story ASSOC eel ART story ASSOC-3sg

‘the story of the eel’ → ‘its story’

⇒ i-n grammaticalised as an anaphoric particle in:

(7) Ne vegevag’ in ppa pe ne. ART story ANAPH finish FOC DX1

‘So that’s how THE STORY ends.’ [Eel_86]

Yet, in cannot be used for non-anaphoric definite (‘familiarity’, ‘recognition’, ‘uniqueness’):

(8) Noke peon tō ŋwuye yö vonyō (’in). 1sg FUT go:NPL return LOC village (’ANAPH)

‘Let me go back to THE VILLAGE.’ (’... in question)

⇒ anaphoric definites > non-anaphoric definites > specific indefinites > non-specific indefinites

Hiw ne X ne X ne X ne X

ne X in

cf. referentiality scale (Dryer 2014):

<table>
<thead>
<tr>
<th>anaphoric definites</th>
<th>non-anaphoric definites</th>
<th>specific indefinites</th>
<th>non-specific indefinites</th>
</tr>
</thead>
<tbody>
<tr>
<td>the X</td>
<td>the X</td>
<td>a X</td>
<td>a X</td>
</tr>
<tr>
<td>ne X</td>
<td>ne X</td>
<td>ne X</td>
<td>ne X</td>
</tr>
</tbody>
</table>

Eng.
2.2 Indefiniteness and specificity

Three different indefinite articles in Hiw! → What differences?

‘an N’: ① sê N — ② ne N sê — ③ ne N ŋot

2.2.1 Non-specific indefinite

The specific indefinite construes a unique referent, unknown to the addressee (→ [-DEFINITE]), yet endowed with individual existence. ⇒ EXTENSIONAL reading → [+specific]:

She wants to marry A FIREMAN... He’s called Jack and he’s from Dublin.

The non-specific indefinite construes a type of referent based on a qualitative property, without entailing the existence of an individual ⇒ INTENSIONAL reading → [-specific]:

She wants to marry A FIREMAN... but she hasn’t found any to her taste.


Hiw /sê N/ encodes exclusively non-specific indefinites:

(9) Pavën ike yēaŋ sê ořyē on kōge sise on tōgō.
then 2sg seek INDF:NSPEC rope SUBJ 3pl SUBJ hard

‘Then you look for A STRING so as to tie them firmly.’ [q.d07.Kenu:11]

Typical of [-SPEC] indefinites (cf. François 2002: 60 sqq.), { sê N } is incompatible with realis declarative clauses (which entail an actual event, and therefore the existence of its participants):

(10) *Noke tēwōri sê ořyē piti.
1sg find INDF:NSPEC rope CPLT

*I’ve found A[+spec] STRING.

If the sentence is realis declarative, an indefinite is normally [+SPEC]. Instead of { sê N }, the only grammatical construction is { ne N sê } [-DEF +SPEC]:

(11) Noke tēwōri n’ ořyē sê piti.
1sg find ART rope INDF:+SPEC CPLT

‘I’ve found A[+spec] STRING.’

A realis declarative is only compatible with { sê N } with verbs that are intrinsically INTENSIONAL (e.g. ‘want’, ‘look for’) as they don’t entail the existence of X (cf. Moltmann 1997). → (9) ‘seek’.

The typical context for { sê N } are predicates which are made semantically intensional through their MODAL specifications: conditional clauses, irrealis or habitual predicates:

(12) Tomñwē sê tayō ve yuŋ ton inine, nine tõ vën
if INDF:NSPEC person IPFV ask ABL 3sg 3sg go:NPL up
yōñwōwōn, tēwāri sê řēřē pe nē meřawë, tärë surř.
in:bush find INDF:NSPEC tree REL STAT perfect cut down

Non-specific { sê N } typically shows up in negative sentences ...

(13) Noke tatī yoš sê gē ti ŋwute pene.
1sg NEGːR see INDF:NSPEC thing PAST place DX1

‘I didn’t see anything.’ — LIT. I didn’t see A[+spec] thing here. [q.d5.Naef:08]
(13’)
*Noke yô sê gê ti ñwute pene.
1sg see IND:SPEC thing PAST place DX1
*I saw *{sê} thing here.

→ ENG. I didn’t see anything. ≠ *I saw anything.

... or as the subject of a verb of absence (intrinsically negative):

(14) Tego, sê votwu ve yeër ñwute pene.
no IND:SPEC knife IPFV lack place DX1
LIT. *{sê} knife is lacking here. = ‘There is no knife here.’ [q.d5.Naef.06]

⇒ see this pair of sentences in a text:

(15) Ne tayô sê o ne têg ti nêne!
ART person IND:SPEC open ART door PAST DX2
‘SOMEBODY opened this door!’ [Grouper_32]
Declarative realis ⇒ EXTENSIONAL reading ⇒ [+specific]

(16) Sê tayô tati me tô me!
INDF:SPEC person NEG:R INTSF go:NPL hither
‘NOBODY came here!’ [Grouper_34]
Negative realis ⇒ INTENSIONAL reading ⇒ [-specific]

Cf. semantic map of indefinites (Haskelma 1997:249) ⇒ Hiw \{ sê N \}

(1) specific known (2) specific unknown
(3) question (4) indirect negation (5) direct negation
(6) direct negation (7) conditional (8) comparative (9) free choice

It is always possible to underspecify the definiteness status of the NP: \{sê N = ne N\}

(17) Ne tayô tati me tô me!
ART person NEG:R INTSF go:NPL hither
‘NOBODY came here!’ [Grouper_36]

2.2.2 Specific indefinite

If the NP is indefinite [+SPEC], it will normally be encoded by \{ ne N sê \} – see ex.(11), (15)...

(18) Noke nsis tom noke vati-k’i ne yekeyake sê.
1sg want COMP 1sg show-2sg OBL ART dance IND:SPEC
‘I’d like to teach you A_{[SPEC]} DANCE.’ [Music.43]
(‘there is a specific dance I want to teach you’)

(19) Sise tog’i ne metëvônëyô sê. Sise to—ge toge:

teñwên sê nine tati yog.
man IND:SPEC 3sg NEG:R married
‘They lived in A_{[SPEC]} VILLAGE. They lived on and on...
But (there was) A_{[SPEC]} MAN (who) wasn’t married yet.’ [Grouper_03]
6 — Coding of (in)definiteness in northern Vanuatu

(20) Könh së, Meřavitit gatet ti tekña ququy ena...
day INDF:SPEC M. say DAT HUM:MX:PL friend POSS:3sg
‘ONE[+SPEC] DAY, Megrvavit said to his friends...’ [Tyingevuv.007]

⇒

<table>
<thead>
<tr>
<th>Hiw</th>
<th>anaphoric</th>
<th>non-anaphoric</th>
<th>specific</th>
<th>non-specific</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ne X</td>
<td>ne X</td>
<td>ne X</td>
<td>ne X</td>
</tr>
<tr>
<td></td>
<td>ne X in</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ne X së</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>së X</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


2.3 The pragmatic parameter

2.3.1 Discourse topicality

The contrast [+SPECIFIC] belongs to logical semantics:

<table>
<thead>
<tr>
<th>[+DEF]</th>
<th>x is presented as ±familiar to the addressee</th>
</tr>
</thead>
<tbody>
<tr>
<td>[+SPEC]</td>
<td>x is presented as ±endowed with extensionality (existence)</td>
</tr>
</tbody>
</table>

Yet some languages combine these semantic dimensions with PRAGMATIC parameters – particularly, discourse saliency or topicality:

1. An indefinite may serve to construe a new referent with LOWER TOPICALITY, only once, with no further mention in the subsequent text
2. An indefinite may serve to construe a new referent with HIGHER TOPICALITY, showing more cognitive and discursive persistence in the subsequent text

Compare:

1. Anna seized A SUITCASE, a coat, a hat, and ran out to the train station. ['IT was heavy...]
2. Anna seized A SUITCASE that had been left there in the corner of the attic. It was an old leather case covered in dust, and surprisingly heavy. She tried to see what was inside, but it was locked. How was she going to open it?

Givón (1992): “GRAMMAR OF REFERENTIAL COHERENCE” – TOPICALITY

| #1 | unimportant indefinite ≠ #2 = important indefinite |

Dryer (2014): #1 = pragmatically non-specific (but semantically specific) indefinite [sic]
#2 = pragmatically specific (and semantically specific) indefinite

2.3.2 The presentative indefinite

Hiw really has two morphemes coding for INDF:SPEC : {NE} N së, but also {NE} N ṇot.

At first sight, së and ṇot are equivalent... Yet my corpus suggests they differ in TOPICALITY.

1. /së/ usually goes with indefinite referents with LOW TOPICALITY
   { ID:SP:BKG } = indefinite, specific, BACKGROUND
2. /ṇot/ usually goes with indefinite referents with HIGH TOPICALITY
   { ID:SP:TOP } = indefinite, specific, TOPICAL

E.g. ṇot flags the first mention of a new character, which is later central in the text:

(21) Könh së, NE TAMESŌ ṇot nine tō nine ōnwe nōna yöte ti.
day ID:SP:BKG ART old ID:SP:TOP 3sg go:NPL 3sg weed his garden PAST
‘ONE day, AN OLD MAN went to work at his garden.’ [Yams_02]
Kōn sē, NE YEQĒN TAMESŌ ǸOT, mi ne megoye na viŋō. lit. 'One day, AN OLD WOMAN with her two children.' [Brothers_03]

possible gloss for ǸOT: indefinite presentative (INDF:PRSTV)

Tameřën nine putput vage-ṱō-on, NE YEQĒN TAMESŌ ǸOT
moment 3sg sing times-two-ORD ART woman old ID:SP:TOP

And as he sang for the second time, he was heard by AN OLD WOMAN: “Hey, friends!” [She said]…’ [Eel_74]

Sometimes we get a cascade of ǸOT-marked NPs:

Kōn sē, nine tō ti űrow eyo. Tō űrow me,
day ID:SP:BKG 3sg go:NPL PAST out shore go:NPL out hither
‘One day, he walked down to the shore. As he got on the shore,

yō NE QEṉŌN PE ǸOT ve tu.
see ART ditch water ID:SP:TOP

he discovered A CREEK that was flowing there.

Pataŋ i ne qeṉōn pë in, NE TŌT VOT EṉOT ve sag.
close to ART ditch water ANAPH ART CLF:VERTIC stone ID:SP:TOP IPFV sit
Next to THAT creek, A ROCK was standing.

Nine sag, ne sag p’ënē, ne sag ne gengon; (...) 3sg sit:NPL 3sg sit:NPL FOC DX1 3sg sit:NPL 3sg eat~INTR
So he sat there; and as he was sitting, he began to eat [his yam];

ne trog ne gengon ena vën yō qeṉōn pē, ne yō: 3sg throw:PL ART food his thiser LOC ditch water 3sg see
as he was throwing crumbs into the river, he saw:

⇒ NE MEṆEMPË ēy o me.
ART eel crawl out hither
he saw (AN) EEL crawling out to him. [→ex.(1) p.2]

⇒ NE MEṆEMPË, pa nē~~~ kkē! Nē kkē, nē kkē ņēt!
ART eel but STAT small STAT small STAT small INTSF
(THAT) EEL was so~~~ small!! It was small, so very small!’ [Eel_09]

Lack of ǸOT (cf. ⇒) is surprising, in a story about an eel... Shows that even the category “INDEFINITE, SPECIFIC, TOPICAL” may be left unspecified, i.e. only encoded with article ne ‘DET’.
2.4 Synthesis: definiteness marking in Hiw

Definiteness is sometimes underspecified (ne), but also sometimes “over-specified”:

<table>
<thead>
<tr>
<th>Hiw</th>
<th>anaphoric definite</th>
<th>non-anaphoric definite</th>
<th>topical, specific indefinite</th>
<th>non-topical specific indefinite</th>
<th>non-specific indefinite</th>
</tr>
</thead>
<tbody>
<tr>
<td>ne X in</td>
<td>ne X</td>
<td>ne X</td>
<td>ne X ŋot</td>
<td>ne X së</td>
<td>së X</td>
</tr>
</tbody>
</table>

3 An areal typology

The morphosyntactic categories of Hiw are so specific that they constitute a perfect ‘etic grid’ for analysing an areal typology of definiteness marking in N. Vanuatu languages.

→ see next page

4 Conclusion

The languages of northern Vanuatu can thus contribute in a broader endeavour, namely the TYPOLOGY OF DEFINITE AND INDEFINITE ARTICLES (Dryer 2014: e238):

<table>
<thead>
<tr>
<th>AD</th>
<th>TYPE OF NOUN PHRASE</th>
<th>anaphoric definite (AD)</th>
<th>nonanaphoric definite (ND)</th>
<th>pragmatically specific indefinite (PSI)</th>
<th>pragmatically nonspecific but semantically specific indefinite (PNI)</th>
<th>semantically nonspecific indefinite (SNI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ND</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PNI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AD + ND</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ND + PSI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSI + PNI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PNI + SNI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AD + ND + PSI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ND + PSI + PNI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSI + PNI + SNI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AD + ND + PSI + PNI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ND + PSI + PNI + SNI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>all five types</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TABLE 1. A preliminary typology of articles.
## An areal typology of definiteness marking

Languages of Torres & Banks (cf. François 2007, 2011) + Araki (François 2002)

<table>
<thead>
<tr>
<th></th>
<th>ANAPHORIC DEFINITE</th>
<th>NON-ANAPH. DEFINITE</th>
<th>TOPICAL, SPECIFIC INDEFINITE</th>
<th>NON-TOPOCAL, SPECIFIC INDEFINITE</th>
<th>NON-SPECIFIC INDEFINITE</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIW</td>
<td>ne X in</td>
<td>ne X ūnot</td>
<td>ne X sē</td>
<td>sē X</td>
<td></td>
</tr>
<tr>
<td>LTG</td>
<td>ne X in</td>
<td>ne X sise</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LHI</td>
<td>n-X</td>
<td>n-X e(n)</td>
<td>n-X tā</td>
<td>n-X vētwa ['1']</td>
<td></td>
</tr>
<tr>
<td>MTP</td>
<td>nA-X e(n)</td>
<td>nA-X vētwag ['1']</td>
<td>nA-X nan</td>
<td>nA-X vōwal ['1']</td>
<td>te X</td>
</tr>
<tr>
<td>LMG</td>
<td>n-X</td>
<td>n-X e</td>
<td>n-X vōwal ['1']</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VRA</td>
<td>(ē)n X</td>
<td>(ē)n X ne vōwal ['1']</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VRS</td>
<td>o X</td>
<td>o X ni-tiwial ['1']</td>
<td></td>
<td></td>
<td>?</td>
</tr>
<tr>
<td>MSN</td>
<td>o X</td>
<td>o X ni-tawal ['1']</td>
<td></td>
<td></td>
<td>?</td>
</tr>
<tr>
<td>DRG</td>
<td>o X neñ</td>
<td>tuar (o) X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OLR</td>
<td>X</td>
<td>tay X ['other']</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LKN</td>
<td>X</td>
<td>X</td>
<td>too X ['other']</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRL</td>
<td>nV- X</td>
<td></td>
<td>nV- X tuwel ['1']</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARK</td>
<td>X</td>
<td></td>
<td>X mo hese ['1']</td>
<td>ūre X</td>
<td></td>
</tr>
</tbody>
</table>
10 — Coding of (in)definiteness in northern Vanuatu

**Abbreviations**

<table>
<thead>
<tr>
<th>ABL</th>
<th>ablative</th>
<th>MX</th>
<th>mixed gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPLT</td>
<td>complete</td>
<td>NEG</td>
<td>negation Reals</td>
</tr>
<tr>
<td>ANAPH</td>
<td>anaphoric</td>
<td>NPL</td>
<td>non-plural</td>
</tr>
<tr>
<td>ART</td>
<td>article</td>
<td>NSPEC</td>
<td>non-specific indefinite</td>
</tr>
<tr>
<td>ASSO</td>
<td>associative</td>
<td>OBL</td>
<td>oblique</td>
</tr>
<tr>
<td>CONT</td>
<td>continuous aspect</td>
<td>ORD</td>
<td>ordinal numeral</td>
</tr>
<tr>
<td>DUR</td>
<td>durative</td>
<td>PFT</td>
<td>perfect</td>
</tr>
<tr>
<td>DX</td>
<td>deictic (1st, 2nd, 3rd degree)</td>
<td>POT</td>
<td>potential</td>
</tr>
<tr>
<td>HUM</td>
<td>number classifier for humans</td>
<td>PRSTV</td>
<td>presentative</td>
</tr>
<tr>
<td>INDF</td>
<td>indefinite</td>
<td>QUOT</td>
<td>quotative</td>
</tr>
<tr>
<td>INTSF</td>
<td>intensifier</td>
<td>REL</td>
<td>relatiser</td>
</tr>
<tr>
<td>IPFF</td>
<td>imperfective</td>
<td>SPEC</td>
<td>specific indefinite</td>
</tr>
<tr>
<td>IRR</td>
<td>irrealis</td>
<td>SUBJ</td>
<td>subjunctive</td>
</tr>
<tr>
<td>LOC</td>
<td>locative</td>
<td>STAT</td>
<td>stative aspect</td>
</tr>
</tbody>
</table>

**Hiw orthography**

| orthogr. | a e ē ē̂ g i k m n ŋ ňw o ō ṑ p q ō̃ r s t u v w y |
| iPA     | a ə e ɪ y i k m n ŋ ŋʷ ɔ ɵ o p kʷ ɡ l s t ɵ β w j |

**5 References**


