Silent Gaps as a Distinctive Feature of Direct Speech

A Corpus Study of Nisvai Narratives, Vanuatu

Jocelyn Aznar
February 11, 2021

Leibniz ZAS
TABLE OF CONTENTS

1. Introduction
   About Nisvai
   Corpus
   Methodology

2. Reported Speech in Nisvai
   Indirect Speech
   Direct Speech

3. The silent gaps of Nisvai Direct Speech
   Silent gaps

4. Conclusion
INTRO
Problematic

Study of silent gaps as part of the Nisvai Direct Speech device

- how silent gaps contribute to the Nisvai’s Direct Speech device,
- developing tools for doing cross-linguistic comparison of silent gaps.
The Nisvai Community and Corpus’ Project

An oral language which has been the subject of only one previous lexical study by Charpentier (1984).

Around 200 natives speakers: the native language of the children.

The language is used in most daily activities, except:

- Relationships with other village communities: Bislama.
- Church: Bislama, a bit of English. Translating to Nisvai.
- School: Teaching language is French, but children actually learn to speak Bislama fluently there, and just few words of French.

Corpus original purposes:

- Ethnolinguistics research on Nisvai narrative practices,
- For the Nisvai speakers: having a written system and language resources for the local school.
ABOUT NISVAI

- Austronesian language > Oceanic Language > Central Vanuatu
- SVO language
- Offers some lexical flexibility (van Lier, 2016), but not omnipredicative as other languages of the area.
- Very few adjectives and adverbs, but many stative verbs (including numbers, color terms):
- Only animate objects can be subject and no passive.
- Tense: non Future/ Future
- Mood: Irrealis/ Realis

1See https://glottolog.org/resource/languoid/id/nisv1234 for a more detailed description.
2Many Austronesian language are described as omnipredicative (Launey, 2004) (François, 2001).
3The name, used by the speakers, means ”what”, as it is very common in the area.
The corpus of Nisvai narratives:

- 165.5 minutes with 32 narratives of between 0.67 to 13.32 minutes,
- 3041 annotations units: 1301w, 1740m,
- 18 speakers, 2 sex (5w, 13m), from 5 to 80 years old,
- practical orthography: no characters outside the ASCII set,
- Nisvai speakers regard their narratives as true stories.
Figure 1: Segmentation into annotation units: interpausal units.
### Figure 2: Content annotation using ELAN

<table>
<thead>
<tr>
<th>No</th>
<th>Type 1: Transcription</th>
<th>Type 2: Textualre</th>
<th>Type 3: Reference</th>
<th>Type 4: Morphology</th>
<th>Type 5: Note</th>
<th>Type 6: Traduction</th>
</tr>
</thead>
</table>
| 93 | Ga-nal kal van : | pépétv_9 | T21.12013.93 | 3SG=dira derrière ASP-A | Il dit, derrière : | Mais non, j’ai déjà ramené les enfants, reviens, allons-y, |}
| 94 | 4daa, ra-pa=ri ku, nahi moq, qa=ri tu mai, danu | pépétv_9 | T21.12013.94 | FPER.NEG.1SG=porter=entrer enfant ASP.F.2SG=entrer.RED.revenir 1NML | pâgulî : karem bâk | |}
| 95 | Be naly-m ga=sab mi cab. | pépétv_9 | T21.12013.95 | COOG=OP entraînes 2SG. 3SG=être en colère aussi FIV | | Ne sois pas en colère. |}
| 96 | Ga-nal danu, dara=la va=van dy vey=roh. dara=la=nali pai=ri moq bul... | pépétv_9 | T21.12013.96 | ZSG=enfer 1NML 1NCL 1NLC=enfer aller C0G aller façon entferner 1NCL 1NRC=aller 1NLC=aller 1NML=aller. 3SG=être en colère aussi FIV | | Mêmes, alors-y, revenons et restions pour ensemble, c’est bien comme ça mais... |}
| 97 | qa=pac=ri annahra moq dry van va=li=ti. Dari dara=va va=nabû=ku hafrî. | pépétv_9 | T21.12013.97 | ZSG=porter=entrer ART.P enfant ASP.F 2GO aller façon déposer-INTR. 1NML 1NCL=aller 1NLC=aller 1NML=aller 1NRC=aller 1NLC=aller 1NML=aller. 3SG=être en colère aussi FIV | shamus bâ : donner le sein. | Rapporte les enfants et vales déposer. Alors les nourrir, |}
| 98 | Bul b=Uwâreme, nanovin syn ga=sab. | pépétv_10 | T21.12013.98 | COOG=OP ART.P=XOM penser.NOMF 1NML 1NCL=aller 1NLC=aller 1NML=aller. 3SG=être en colère aussi FIV | | |}
| 99 | Ga-nal ga=mi=as (ga=roh dry) ni. | pépétv_10 | T21.12013.99 | 2SG=dira 2SG=devir 2SG=rester 2GO MORT 1NML | Il se dit qu’il doit mourir. | |}
| 100 | Ga-câh-ni mitas, ga=dûby avun=Barbaruh. | pépétv_10 | T21.12013.100 | 3SG=peter-INTR à nouveau 3SG=tomber INDEF-INTR.LEU | barbaruh : blaisand / nabanbara. | Il en jette une à nouveau, elle tombe dans un endroit comme Barbâruh. |}
| 101 | Ga-la i dry-sôh. | T21.12013.101 | 3SG=chap=cher COOG | Il chante et vole | | |}
| 102 | e=sa=mnatum anu meh=ni, i=sati= Quartât au anu vog=voq= | champ_5 | T21.12013.102 | PRO.PA.2SG enfant=2SG 3PL=AASP BC être.NO.RED PRO.PA.1SG enfant=1SG 3PL=AASP BC être.MORT=RED | les tânes, tes deux enfants sont morts, les miens, mes deux enfants, sont morts. | |}
| 103 | Ga=sôh dry... | pépétv_10 | T21.12013.103 | 3SG=s’envole COOG | Il s’envole et | |}
| 104 | ga=Kâmitas van bulahe qanag | pépétv_10 | T21.12013.104 | 3SG=chap à nouveau ASP.A MUL attendre comme | chante à nouveau. L’attente commence | |}
| 105 | pac=nâmitas avun=na=ri ga=dûby il=keng. | pépétv_10 | T21.12013.105 | prendre chose DEF à nouveau outer-INTR 3SG=tomber plage | prend une autre larce, la jette et elle tombe sur la plage. | |}
Computer notebook: Jupyter (Kluyver et al., 2016))

- Reproductibility Berez-Kroeker et al. (2018),
- Facilities for doing exploration studies using quantitative and visualisations analysis,
- Algorithmic studies: apply queries on each annotation and have a result,
- Helps for replication as well by providing the algorithms that were used Drummond (2009).

The notebook and the data associated with this presentation will be available soon.
Reported Speech in Nisvai
**GENERAL REMARKS**

Direct speech sequence/frame (Genetti, 2014) = introductory segment + reported speech + conclusive segment

Indirect speech sequence = introductory segment + reported speech

<table>
<thead>
<tr>
<th>Indirect Speech</th>
<th>Direct Speech</th>
</tr>
</thead>
<tbody>
<tr>
<td>Keep the narrative’s language</td>
<td>Language of the reported speech</td>
</tr>
<tr>
<td>Current situation’s mood</td>
<td>Mood of the Reported situation</td>
</tr>
<tr>
<td>Current situation’s deictics</td>
<td>Reported situation’s deictics</td>
</tr>
<tr>
<td>No Enactement</td>
<td>Enactment⁴</td>
</tr>
</tbody>
</table>

⁴different voice timbre, different pitch
Indirect speech is often used to convey the thoughts of the characters.

(1) Interior speech, T21.t2013.98-99
Bul a=Livwume, nanov-in sy-n ga=sab.
COO.OP ART.P=NOM think-NOMI PA-3SG 3SG=bad
‘But the Pigeon is angry.’
Ga=kal ga=mas mac ni.
3SG=say 3SG=must die INT
‘She is telling herself that she has to die\(^5\).’

\(^5\)Little enactment here: lower and angry voice
(2) Indirect speech, T15.2013.10

Ara=lav-lavis van ili, bog a=ckai, tatag
3PL=RED-play ASP.A ASP.R day 3SG=be_one father
sy-n ga=pac-i ga=kal ga=va=vor lhoi.
PA-3SG 3SG=take-INTR 3SG=say 3SG=go=draw_water river

‘They play sometimes, one day, his father carries something and says he is going to draw water in the river.’
(3) Interior speech, T1.2013.125

‘But among the group of the deamon that went to the mud, one says he will hit him, another says he will hit him.’
435 direct speech occurrences in the corpus.

Introductory verbs:

- kal: "to say", not mandatory but... almost.
- tog-i: "to listen-to something",
- qan: "to be like",
- kai: "to sing"

Interjections or Appelative: Marking the beginning of the quotation or turn-taking in a dialogue.

Poetic sequences: The poetic sequences are always introduced as direct speech.
Examples of Direct Speech

(4) “pus”: support verb for direct speech, T44.2015.37

\[
\text{ara=pus} \quad \text{ga=kal}: \quad \text{“Haig, nanag sa-m nmah a?”}
\]

3PL=ask 3SG=say 2SG mother PA-2SG snake INTER

‘They ask, s/he say: «You, your mum is a snake, eh!»’
'«But hers are black, very black.» And he looks at those of the Pigeon and thinks: «Why she says hers have a brown throat like that, hers have a brown throats like that.» he thinks.'
(6) T41.2015.97 T41.2015.101

«E!», ga=kal-i visgy-n napnevur sy-n aru
INTER 3SG=say-INTR OBJ-3SG woman PA-3SG 3PL.DU
ga=kal : «E!
3SG=say INTER

‘«Eh» he says to his two wives: «Eh !’

Mura=ls-i, avys ara=mac-mac pal moq, amru
2PL=see-INTR other 3PL=be_dead ASP.T ASP.F 2PL
ni mycig mra=roh-roh.
INT COO.TMP 2PL=stay-RED

‘You see, the others are dead already, only you two are left.’

Be mra=kaikai nahsy-q,[...]
NEG 2PL=call name-1SG

‘Don’t call my name.[...])’
Silent gap just after the interjection, **T1.2013.189-190**

*Bul*  
a=Tamata  
van  
va=puc  
ari  
kal:  «Eh,  
COO.OP  
ART.P=NOM  
go  
go=lie  
3PL  
say  
INTER  

‘So Tamata goes to lie and says: «Eh,’  

dara=qamqam  
ni.»»  
1INCL=to_speed_up  
INT  

‘let’s go fast.»’

They represent 112 of 435 direct speech occurrences in the corpus.
The silent gaps of Nisvai Direct Speech
Breath gaps included into silent gaps.

<table>
<thead>
<tr>
<th>% of silence $^6$</th>
<th>25%</th>
<th>mean</th>
<th>50%</th>
<th>75%</th>
</tr>
</thead>
<tbody>
<tr>
<td>length</td>
<td>20.6</td>
<td>25.9</td>
<td>26.6</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>430</td>
<td>845</td>
<td>745</td>
<td>1100</td>
</tr>
</tbody>
</table>

Figure 3: Length of silence gaps in the corpus according to the age of the speakers.
What are the factors that influence the length of silent gaps:

- Sex (coef \(+ - 59.5137\), \(P>|t| = 0.000\))
- Age (coef \(1.7553\), \(P>|t| = 0.003\))
- Position in the text (coef \(-2.5702\); \(P>|t| = 0.000\))
- Local Speech Rate\(^7\) (coef \(15.0498\); \(P>|t| = 0.000\))
- Text’s length (coef \(0.3552\); \(P>|t| = 0.000\))

According to the statistical tests (OLS), these variables affects significatively the length of silent gaps, but only very slightly (as the coefficients show.).

\(^7\)Local Speech Rate is calculated by counting the number of phonemes in an annotation units.
A MODEL OF NISVAI’S DIRECT SPEECH SEQUENCE

<table>
<thead>
<tr>
<th>Direct Speech Frequencies</th>
<th>Verb</th>
<th>gap</th>
<th>INTER</th>
<th>gap</th>
<th>RS</th>
<th>gap</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>8</td>
<td>199</td>
<td>356</td>
<td>112</td>
<td>435</td>
<td>417</td>
</tr>
</tbody>
</table>

8I have not counted them yet, but direct reported speech without introductory verbs exists, but are not numerous.

9Only 14 interjections are not used for marking the beginning of direct speech. No use of interjection for indirect speech.
Sample: 435 of Direct Speech sequences over around 3000 silent gaps.

The length of the silent gaps:

- after an interjection (112 occurrences): no significant variation ($P>|t|\ 0.545$),
- before direct discourse (199 occurrences): no significant variation ($P>|t|\ 0.100$),
- after direct reported speech (417 occurrences): significant increase ($P>|t|\ 0.000$), mean length 985 ms (compared to 821 ms).
CONCLUSION
in Nisvai, interjections and silent gaps (before and during) are distinguishing Direct Speech from Indirect Speech.

Silent gaps can be after the introductory verbs, but also after the interjection starting the direct speech.

- Nisvai has a very limited number of introductory verbs
- Direct speech is produced using two terms (usually verbs):
  - A verb or a term of manner,
  - A verb introducing the content of the reported speech.
• Taking into account the role of appelative for direct speech.
• Using this as a framework for cross corpora comparison.
• Sex and age will not be comparable using language documentation corpora as they mostly contain texts performed by few speakers.
REFERENCES


<table>
<thead>
<tr>
<th>sex</th>
<th>age</th>
<th>age_group</th>
<th>speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>5</td>
<td>child</td>
<td>Anici</td>
</tr>
<tr>
<td>F</td>
<td>5</td>
<td>child</td>
<td>Selina</td>
</tr>
<tr>
<td>F</td>
<td>15</td>
<td>child</td>
<td>Kati</td>
</tr>
<tr>
<td>F</td>
<td>60</td>
<td>elder</td>
<td>Limeten</td>
</tr>
<tr>
<td>F</td>
<td>70</td>
<td>elder</td>
<td>Rosie</td>
</tr>
<tr>
<td>M</td>
<td>5</td>
<td>child</td>
<td>Enfant_2</td>
</tr>
<tr>
<td>M</td>
<td>10</td>
<td>child</td>
<td>Epel</td>
</tr>
<tr>
<td>M</td>
<td>10</td>
<td>child</td>
<td>John-Hari</td>
</tr>
<tr>
<td>M</td>
<td>30</td>
<td>adult</td>
<td>Aven</td>
</tr>
<tr>
<td>M</td>
<td>35</td>
<td>adult</td>
<td>Filip</td>
</tr>
<tr>
<td>M</td>
<td>35</td>
<td>adult</td>
<td>Kalbatik</td>
</tr>
<tr>
<td>M</td>
<td>40</td>
<td>adult</td>
<td>Bongman</td>
</tr>
<tr>
<td>M</td>
<td>50</td>
<td>adult</td>
<td>Dale</td>
</tr>
<tr>
<td>M</td>
<td>60</td>
<td>elder</td>
<td>Luan</td>
</tr>
<tr>
<td>M</td>
<td>60</td>
<td>elder</td>
<td>Sumbles</td>
</tr>
<tr>
<td>M</td>
<td>70</td>
<td>elder</td>
<td>Bogalgal</td>
</tr>
<tr>
<td>M</td>
<td>80</td>
<td>elder</td>
<td>Bongmeme</td>
</tr>
<tr>
<td>M</td>
<td>80</td>
<td>elder</td>
<td>Salben</td>
</tr>
</tbody>
</table>
... the relevance of the model:

T16.t2013.96 - 97

- Ara=tog-i van dry taid ili ara=kal: «Tck.»
- Ara=wuc moq-i, ara=lig plan avyn dry kal: «E

Aven then restarts with direct discours.