Symposium on Amazonian Languages II - April 9th, 2017

The expression of topic in Kakataibo

Daniel Valle
University of Texas at Austin

1 Introduction

- This paper explores the morpho-syntactic expression of topic in Kakataibo.

- In doing so, this research aims to shed some light into the information structural factors that condition the use of certain alternating structures conveying the same propositional content.

(1) a. uni=n=ka=a  ‘inu  ‘a-i-a.
   Kakataibo.person=A/S=VAL=3A/S  jaguar  do-IPFV-N.PROX

b. =ka=a  ‘uni=n  ‘inu  ‘a-i-a.
   =VAL=3A/S  Kakataibo.person  jaguar  do-IPFV-N.PROX

c. =ka=a  ‘a-i-a
   =VAL=3A/S  do-IPFV-N.PROX
   ‘A/the man kills/is killing a/the jaguar.’

Topic

- What the sentence is about (file card metaphor, Gundel 1988; Lambrecht 1994; Reinhart 1981, among others)
- Psychological subject (Chafe 1976)
- Old (given) information (Firbas 1966)
- Topic may even constitute a family-term for associated meanings such as information separation, addressation, semantic subject and frame setters (Jacobs 2001).

- In this research, topic is viewed as what the sentence is about.
  -Tests to identify topichood (Gundel 1988; Reinhart 1981):
    - Tell me more about x
    - What about x

- The information expressed by the topic may have different relations to the information already stored in the common ground (Givón 1983; see Stalnaker 1974 for common ground):
  - New topic: when a referent achieves topical for the first time in discourse.
  - Change of topic: a referent that had topical status in the previous discourse but is different from that of the previous sentence.
  - Continuing topic: topic already introduced into the common ground and is similar to the previous sentence’s topic.
- Cross-linguistically, topics are expressed by dedicated morphemes (e.g. -wa in Japanese Kuno 1973), a specialized position in the clause (e.g. initial position in Turkish Erguvanlı 1984); or phonological clues (e.g. f0 variations in Italian (Frascarelli and Hinterhörlz 2007).

**Kakataibo language**

- Kakataibo (ISO 639.3 code ‘cbr’) is a Panoan language spoken by approximately 1500 people (Frank 1994) along the Aguaytía, San Alejandro and Sungaroyacu rivers in the Peruvian Amazon regions of Ucayali and Huánuco (see Map 1).

**Map 1. Kakataibo native communities territory (Peruvian Ministry of Culture)**

- Pro-drop is highly frequent; subject arguments are cross-referenced on the second-position clitic complex (2CL) and on the verb.

- Split-alignment: ergative for nouns (A marked by =n; S/O marked by Ø); accusative for pronouns (A/S =n and O Ø; ; cf. Zariquiey 2011). Overall, accusative profile throughout the grammar.

- Complex system of switch reference, and pervasive use of clausal nominalizations.

- Relatively free constituent order restricted by two syntactic landmarks and dependent on information structure, e.g. focus and topic.

(1) _____ =2CL _____ V

---

1 This map is available at: [http://bdpi.cultura.gob.pe/mapa/407](http://bdpi.cultura.gob.pe/mapa/407)
The expression of topic in Kakataibo

(2) a. ‘inu=n=ka=a
jaguar=A/S=VAL=3A/S
nis-i-a.\(^2\)
stand.up-IPFV-N.PROX
‘A jaguar is standing up.’

b. kwan-i-a.
‘(The jaguar) goes.’

c. kwan-kin=ka=a
chaxu
mëra-i-i. (Continuing topic)
go-A/S>A:SE=VAL=3A/S
deer
find-IPFV-PROX
‘(The jaguar) finds a deer while going.’

d. mëra-kin=ka=a
chaxu
kwan-i-a. (New of topic)
find-A/S>A:SE=VAL=3A/S
deer
go-IPFV-PROX
‘The deer walks away while (the jaguar) finds (it).’

e. a=n nui-i ‘i-kë
tan
kwan-i-a
baka=nu. (CT)
3=A/S follow-A/S>S:SE
be-NFUT.NMLZ
go-IPFV-N.PROX
water=LOC
‘What the jaguar follows goes to the water’.

d. baka=nu
water=LOC
go-PFV
‘(The deer) went.’

e. baka=nu
water=LOC
asin-i-a. […]
go-in-IPFV-N.PROX
‘(The deer) goes into the water’.

f. kwan-tan-mainun=ka=a
nükën
chichi
baka=nu
go-GO.IMP-A/S≠A/S:S=VAL=3A/S
1PL.O/POSS
grandmother
water=LOC
nain-i-i. (New topic)
shower-IPFV-PROX
‘While (the deer) is going away, the water’s mother is taking a shower on the river.’
(Nükën_chichi-ET-2012-11-05-CBR-1.1-1.7, 1.11)

(3) Morpho-syntactic strategies to express topic
- Grammatical encoding
- Type of NP
- Position in the clause
- Grammatical relation
- Type of clause
- Subordinating morphology

2 Methodology

- The sample comes from 22 naturalistic texts (n=546, min = 100 approx) produced by fourteen native speakers (2 female) whose ages range from fourteen to more than eighty years.

- This sample is a subset of the corpus created under the a documentation project of the San Alejandro dialect of Kakataibo, sponsored by ELDP (IGS #0165), which contains more than 22 hours of naturalistic speech transcribed, translated and analyzed to the morpheme level.

- The selected corpus was chosen mainly because:
  - The referents occurring in them were unambiguously identified.
  - Average syntactic complexity in the texts (number of clauses per sentence).

- Texts in the sample include both conversations and monologues.

- Two native speakers were asked to judge each sentence and decide what referent of those present in the proposition the sentence was about, e.g. what about x test, tell me more about x.

- A Kappa test, which measures the interrater (=speaker) agreement for categorical data, shows that overall they had a good level of (Kappa = 0.635) when identifying the topic.

(4) Morpho-syntactic strategies to express topic

- Grammatical encoding
- Type of NP
- Position in the clause
- Grammatical relation
- Type of clause
- Subordinating morphology

**Grammatical encoding**

(5) Less coding

<table>
<thead>
<tr>
<th>Grammatical encoding</th>
<th>More coding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zero</td>
<td>Nominalizing suffix plus a pronoun (NMLZ pro)</td>
</tr>
<tr>
<td>Bound second-position enclitic subject pronoun (2CL)</td>
<td></td>
</tr>
<tr>
<td>Switch-reference suffix (SR)</td>
<td></td>
</tr>
<tr>
<td>Pronoun</td>
<td></td>
</tr>
<tr>
<td>Switch-reference suffix plus a pronoun (SR pro)</td>
<td></td>
</tr>
<tr>
<td>Full NP</td>
<td></td>
</tr>
</tbody>
</table>

Type of NP

(Bare) noun > Complex NP > Noun plus a third person pronoun (Noun 3p)
The expression of topic in Kakataibo

Position in the clause

<table>
<thead>
<tr>
<th>Position</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial position:</td>
<td>Topic =2CL _____ V</td>
</tr>
<tr>
<td>Middle position:</td>
<td>____ =2CL Topic XP V</td>
</tr>
<tr>
<td>Preceding verb:</td>
<td>____ =2CL _____ Topic V</td>
</tr>
</tbody>
</table>

Grammatical relation

A: subject of transitive verb
S: subject of intransitive verb
O: object of transitive verb

Type of clause

Main clause
Subordinate clause

Subordinating morphology

Same subject (SS)
Argument to argument (ARG to ARG)
Different subjects (DS)
Nominalizer (NMLZ)

3 Results

3.1 New topic

Grammatical encoding

- New topics are significantly preferred to occur overtly in the form of a Full NP, $X^2 (6, N=56) = 25.38, p<.001$.
- A post hoc test shows that new topics are preferred to be grammatically encoded as an NP over 2CL ($p<.001$), switch-reference (SR) suffixes ($p=.004$), SR plus a pronoun ($p=.009$) and pronoun ($p=.003$).
- Other pairwise comparisons with NP as one of the elements did not show a significant difference.
(6) a. ‘aiká  uda  ‘ióxa.
   ‘ai=ka=a  u=da  ‘i-on-x-a
   then=VAL=3A/S  3=COMP  be-HST-3-N.PROX
   ‘Then, it was like that.’

   b. ‘ai=ka=a  bërí  unikama  desfilá  ‘ibaitia
   bërí  uni=kama  desfilá  ‘i-bait-i-a
   then=VAL=3A/S  now  K.person=PL  parade  be-DUR-IPFV-N.PROX
   ‘Then, the people were parading.’ (Sobre_aniversario_de_la_comunidad-LIG-20-06-14-CBR.1.21-1.22)

Type of NP

- A noun without any modifier, is significantly preferred over other kinds of NPs, $\chi^2 (2, N=33) = 13.09, p<.001$.
- The simple NP strategy to express new topics was preferred over the use of a complex NP ($p<.001$), one containing a noun modifier, and over the NP plus a co-referential pronoun ($p<.002$) strategies.

(7) ėnë a ě xëki a mëdi ‘ati chakatamánukáidá
   ėnë a ě xëki a mëdi ‘ati chaka-tan-mainun=ka=id=a
   this 3 mmm corn 3 tamal do-FUT.NMLZ cut-GO.IMP-A/S≠A/S:SE=VAL=EVID=3A/S

este  ėn   kukuídá       këkia
este  ě=n   kuku=id=a    kë-ki-ia
hmm  1=POSS father.in.law=EVID=3A/S    scream-INTR-A/S/O>O:SE

kamáno  kuaëxaxa.
kamáno  kuat-ëxan-x-a
non.contacted.Kakataibo group hear-REC.PST-3-N.PROX
‘My uncle heard kamáno people while (people) were cutting down (plants) in order to
make tamales.’ (Encontró_hacha_de_camano-SR-2012-11-09-CBR.1.2)
**Position**
- No significant preference was found, $X^2 (2, N=36) = 2.09, p<.35$.

**Grammatical relation**
- New topics are preferred to occur as the grammatical subject (conflation of A and S), $X^2 (4, N=60) = 24.11, p<.001$.
- A following pairwise *post hoc* test shows that both the A and S grammatical relations are preferred over the O grammatical relation ($p<.05$ for A and $p<.009$ for S) and over non-core arguments ($p<.005$ for A and $p<.001$ for S).
- However, there is no significant preference for new topics as either A instead or S or vice versa.

**Type of clause**
- There is not a significant preference for new topics to occur in either type of clause, ($W = 58, p = 0.8905$), (n=11).
Subordinating morphology

- Any of the subordinating morphology may occur in the subordinate verb; there is not a significant preference for new topics to be introduced by any of the subordinating morphology considered here $X^2 (3, N=44) = 5.45, p<.142$.

(8) […] xanu nētēmainuidá ain tita xēni
xanu nētē-mainun=īd=a ain tita xēni
woman disappear-A/S≠A/S:SE=EVID=3A/S 3.POSS mother old
mēräkëxa.
mēra-t-akē-x-a
find-REFL-REM.PST-3-N.PROX
‘[…] while (the two) women disappeared, their old mother showed up.’
(Historia_de_murciélago-ET-2012-11-16-CBR.1.16)
3.2 Change of topic

**Grammatical encoding**

- A significant preference for using a Full NP to express the change of topic was found, $\chi^2 (7, N=88) = 16.28, p<.024$.
- Nonetheless, this significant preference was shown to be only applicable to the pairwise comparison between a Full NP and a nominalization plus pronoun (NMLZ pro, $p<.026$).
- Other pairwise comparisons involving the Full NP were not found to be significant.

![Change of topic grammatical encoding](image)

(9) a. [...] ñapó nê ëo ‘amainunká kuëuni
    ñapon nê ëo ‘a-mainun=ka=a kuëun-i
    sábalo mm different do-A/S≠A/S;SE=VAL=3A/S stab.on.a.stick-A/S>S;SE

    kwanti kixuidá ka kakëxa.
    kwan-ti ki-xun=id=a =ka=a ka-akë-x-a
    go-FUT.NMLZ say-A/S>A;PE=EVID=3A/S =VAL=3A/S tell-REM.PST-3-N.PROX
    ‘Then, (the man) told his woman “let’s go stab (some fish) while I fish sábalo”’.

b. xanu kwankëxa.
    xanu kwan-akë-x-a
    woman go-REM.PST-3-N.PROX
    ‘(Then, his) wife went (fishing).’ (Abuelo_que_picó_a_su_nieto-ET-2012-11-11-CBR.1.6-1.7)

**Type of NP**

- The use of a bare noun is significantly preferred over the other two options, $\chi^2 (2, N=33) = 19.58, p<.001$.
- A post hoc analysis shows that the use of the bare noun is significantly preferred over both complex NP ($p<.001$) and over nouns plus third person pronoun ($p<.001$).
Position
- Change of topic constituents are preferred to occur in the middle position of the clause, $X^2(2, N=27) = 12.07, p=.002$.
- Nonetheless, this preference only holds when analyzing the pairwise comparison between the initial and middle position of the clause as suggested by a post hoc test ($p<.001$).

\[
\begin{align*}
\text{a.} & \quad \text{chuna inmù nipakë sotia.} \\
\text{chuna} & \quad \text{inmù nis-pat-kë so-t-i-a} \\
\text{spider.monkey} & \quad \text{below walk-DOWN-NFUT.NMLZ sit-REFL-IPFV-N.PROX}
\end{align*}
\]

\[
\begin{align*}
\text{b.} & \quad \text{sotamáinuka} \\
\text{so-t-tan-mainun}=ka=a & \quad \text{'inu maná 'ikë.} \\
\text{sit.down-REFL-GO.IMP-A/S≠A/S:SE=VAL=3A/S jaguar above be.IPFV.3} \\
\text{‘[...] the spider monkey had gone down (the tree) and sits (there). When (the spider monkey) was sitting down, the jaguar is up (there).’} & \quad \text{(Historia_de_chuna-MB-2012-11-17-CBR.1.14-15)}
\end{align*}
\]

Grammatical relation
- The change of topic constituent significantly occurs as the A or S argument of the clause, $X^2(4, N=44) = 27.20, p<.001$.
- A post hoc test shows that both the A and S arguments are preferred to express change of topic over both the copula subject and O arguments (A vs. O, $p=.03$; A vs. COP SUB, $p=.02$; S vs. O, $p=.02$; S vs. COP SUB, $p<.001$).
Type of clause

- No significant preference was found with regard to placing the topical constituent in either the main or subordinate clause, $W = 55, p=0.2097$ ($n = 18$).

Subordinating morphology

- None of the morphological markers reached significance, $\chi^2 (3, N=36) = 0.96, p=.81$. 
3.3 Continuing topic

Grammatical encoding
- A statistically significant preference for certain strategies marking a continuing topic over the others, $X^2 (7, N=104) = 54.38, p < .001$.
- A post hoc analysis shows that the strategies that involve the least morpho-syntactic coding, zero marking and second-position enclitics, are favored to express this kind of topic.
- Continuing topics with zero marking are preferred over the NMLZ ($p < .001$) and SR ($p < .001$) strategies.
- The use of second-position enclitics to signal continuing topics has a wider preference over other grammatical devices such as NMLZ ($p < .001$), NMLZ pro ($p < .001$) and SR pro ($p < .001$).
- However, the use of a Full NP was also found statistically significant in marking continuing topics over the NMLZ ($p < .001$), NMLZ pro ($p < .001$) and SR pro ($p = .013$) strategies.

(11) a. anúbi kwankëxa.
   a=nu=bi kwan-akë-x-a
   3=LOC=CONT go-REM.PST-3-N.PROX
   ‘(They) kept going there.’

b. ‘aidá kwankëxa.
   ‘ai=id=a kwan-akë-x-a
   then=EVID=3A/S go-REM.PST-3-N.PROX
   ‘Then, (they) went.’

c. como trescientos metroín dapiidá kwankëxa.
   como trescientos metro=in =dapi=id=a kwan-akë-x-a
   like three.hundred meter=EXCL =DUBT=EVID=3A/S go-REM.PST-3-N.PROX
   ‘(They) walked only for three hundred meters.’

d. maënu kwankë ‘aiibiidá
   maë=nu kwan-kë ‘ai=bi=id=a
   abandoned.farm=LOC go-NFUT.NMLZ then=CONT=EVID=3A/S
nukuákëxa.
nuku-t-akê-x-a
reach-REFL-REM.PST-3-N.PROX
‘When (they) had gone to the purma, but (they) arrived (there).’
(Historia_de_murciélagos-ET-2012-11-16-CBR.7.10)

**Type of NP**
- A preference for the type of NP expressing continuing topic was found, $X^2(2, N=33) = 19.44, p<.001$.
- A following *post hoc* test showed that the bare NP strategy was preferred over the other two strategies, complex NP ($p<.001$) and Noun 3p ($p<.001$).

**Position**
- It was found that no position was favored to express this kind of topic, $X^2(2, N=39) = 0.474, p=.079$.

**Grammatical relation**
- Continuing topics are preferred statistically to occur as the A or S argument, the subject grammatical relations, $X^2(4, N=60) = 22.2, p<.001$,.
- A post hoc test shows that both A and S are significantly favored over all the other grammatical relations considered in this variable.

Type of clause
- No statistically significant preference was found for any of these two clause types, $W = 62$, $p = .062$, $n = 18$.

Subordinating morphology
- Only the same subject (SS) switch reference strategy is statistically favored to be used with continuing topics, $X^2 (3, N=44) = 21.93$, $p < .001$.
- A post hoc test evidences that the SS strategy was preferred over the three other alternatives: same subject over nominalizer ($p < .001$), same subject over argument to argument ($p < .001$) and same subject over different subjects ($p < .001$).
The expression of topic in Kakataibo

(12) a. mëranakinka chaxu nuiia
    mëra-anan-kin=ka=a chaxu nui-i-a.
    find-REC-A/S>S:SE=VAL=3A/S deer follow-IPFV-N.PROX
    ‘Finding each other, (the jaguar) follows the deer.’

b. nuixuka chaxu bisia
    nui-xun=ka=a chaxu bis-i-a.
    follow-A/S>A:PE=VAL=3A/S deer grab-IPFV-N.PROX
    ‘After following (it), (the jaguar) grabs the deer.’

c. bixunka chaxu rëtëia
    bis-xun=ka=a chaxu rëtë-i-a.
    grab-A/S>A:PE=VAL=A/S deer kill-IPFV-N.PROX
    ‘After grabbing (it), (the jaguar) kills the deer.’

4 Discussion and conclusions

- The main properties of topics in Kakataibo are summarized in Figure 4-1.
- There is a clear morpho-syntactic divide between new and change of topics, and continuing topics (or a more general less – more grammatical encoding continuum, Givón 1983).
- Overt vs. non-overt presence of NP plays a central roll in the expression of topic in Kakataibo (see also Valle n.d. for the expression of focus).
- Topics, regardless of the type, tend to occurs as the A/S argument (Chafe 1976; Givón 1975; Kiss 1995, among others).
- These results also show that topics in Kakataibo are not expressed by a dedicated morphological marker nor a dedicated topic position.
- Methodologically, this research shows that information structural categories, e.g. topic, follow tendencies rather than categorical distinctions.

Figure 4-1

<table>
<thead>
<tr>
<th>Grammatical encoding</th>
<th>New topic</th>
<th>Change of topic</th>
<th>Continuing topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full NP***</td>
<td>Full NP***</td>
<td>Zero***, 2CL***, Full NP***</td>
<td></td>
</tr>
<tr>
<td>Simple NP***</td>
<td>Simple NP***</td>
<td>Simple NP***, Complex NP***</td>
<td></td>
</tr>
<tr>
<td>No preference</td>
<td>Middle***</td>
<td>No preference</td>
<td></td>
</tr>
<tr>
<td>A/S***</td>
<td>A/S***</td>
<td>A/S***</td>
<td></td>
</tr>
<tr>
<td>No preference</td>
<td>No preference</td>
<td>No preference</td>
<td></td>
</tr>
<tr>
<td>No preference</td>
<td>No preference</td>
<td>SS</td>
<td></td>
</tr>
<tr>
<td>No preference</td>
<td>No preference</td>
<td>No preference</td>
<td></td>
</tr>
<tr>
<td>SS</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
References


———. “Grammar and Information Structure of Kakataibo.” University of Texas at Austin.