

Relative Clauses in Gwadeloupéen

Emmanuel Schang (LLL, CNRS & Univ. Orléans, FRANCE)

@ FACS-5, *Lexington*

Back in the old days, linguists were writing complete grammars:

- Bernabé (1983) for Guadeloupean Creole
- Ferraz (1978) for Santomense
- Maurer (1995) for Angolar
- Kouwenberg (1994) for Berbice Dutch
- Baptista (2003) for Cape-Verdean
- Kihm (1994) for Kriyol
- and (not that) many others. . .

Solid basis for the development of corpus-based studies.

Background (2)

The problem:

- many (creole) languages remain only partially described

Background (2)

The problem:

- many (creole) languages remain only partially described
- partial formalizations are hard to evaluate (how accurate are the rules and what is the coverage of the grammar?)

Background (2)

The problem:

- many (creole) languages remain only partially described
- partial formalizations are hard to evaluate (how accurate are the rules and what is the coverage of the grammar?)
- difficult to develop cumulative works

Background (2)

The problem:

- many (creole) languages remain only partially described
- partial formalizations are hard to evaluate (how accurate are the rules and what is the coverage of the grammar?)
- difficult to develop cumulative works

It is unfortunate since creole speakers have high expectations for their languages:

- standardization
- tools (spellcheckers, parsers, etc.)

Background (2)

The problem:

- many (creole) languages remain only partially described
- partial formalizations are hard to evaluate (how accurate are the rules and what is the coverage of the grammar?)
- difficult to develop cumulative works

It is unfortunate since creole speakers have high expectations for their languages:

- standardization
- tools (spellcheckers, parsers, etc.)

▮▮▮ **Creole languages are still 'little-equipped languages' in the 21st century.**

Background (3)

Developing electronic resources for languages is usually expensive (many man.year for core grammars).

Background (3)

Developing electronic resources for languages is usually expensive (many man.year for core grammars). We need **efficient description languages** (such as XMG: Petitjean, Duchier, and Parmentier (2016)) to allow for **semi-automatic grammar building**.

A TAG grammar of Gwadeloupéyen

A forest of syntactic trees (\approx rules)

A TAG grammar of Gwadeloupéyen

A forest of syntactic trees (\approx rules)

- freely available on GitHub:

https://github.com/eschang/xmg_GC_metagramar

A TAG grammar of Gwadeloupéyen

A forest of syntactic trees (\approx rules)

- freely available on GitHub:
`https://github.com/eschang/xmg_GC_metagramar`
- free access to the documentation (book in preparation ©)

A TAG grammar of Gwadeloupéyen

A forest of syntactic trees (\approx rules)

- freely available on GitHub:
`https://github.com/eschang/xmg_GC_metagramar`
- free access to the documentation (book in preparation ©)
- compiler freely available:
`http://xmg.phil.hhu.de/index.php/upload/compile_grammar`

A TAG grammar of Gwadeloupéyen

A forest of syntactic trees (\approx rules)

- freely available on GitHub:
`https://github.com/eschang/xmg_GC_metagramar`
- free access to the documentation (book in preparation ©)
- compiler freely available:
`http://xmg.phil.hhu.de/index.php/upload/compile_grammar`
- parser: `https://sourcesup.cru.fr/tulipa/`

RCs: State of the Art

Data from Glaude (2013)

Analysis

TAG Grammar

Conclusion

RCs: State of the Art

Relative Clauses: Description

Relative Clauses in Gwadeloupéyen have been described in the literature as follows:

- They use an enclitic *la*:

"En créole basilectal gwadeloupéen, la est la seule marque formelle de la construction relative, si le représentant nominal est i (et non le pronom relatif ki), et si le nucleus de la relative occupe la position finale de la principale." 🗨️ Ludwig and Pfänder (2003)

- This enclitic *la* appears with a definite antecedent:

"[...] une relative ayant comme antécédent un nom accompagné de l'article défini doit obligatoirement se terminer par un article de rappel." 🗨️ Bernabé (1983)

Relative Clauses: Example

☞ Damoiseau (2012)

- (1) an vwè tifi-la (k)i ka travay èvè 'w la.
1SG see girl-DEF who IPFV work with 2SG DEF
'I saw the girl who works with you'

Note that the enclitic *la* is correct with both *i* and *ki*.

Relative Clauses: Example

☞ Damoiseau (2012)

- (1) an vwè tifi-la (k)i ka travay èvè 'w la.
1SG see girl-DEF who IPFV work with 2SG DEF
'I saw the girl who works with you'

Note that the enclitic *la* is correct with both *i* and *ki*.

Enclitic *-la* is always present in the questionnaires.

Data from Glaude (2013)

- (2) fòs la mal kanalizé [ki ka fè kè tout moun, tout
strength DEF badly channeled that IPFV do that all person all
direksyon politik ka vini plito pou profité]
guidance political IPFV come rather to benefit
'the misused strength making that everyone, every leadership comes as a
profiteer'

- (2) fòs la mal kanalizé [ki ka fè kè tout moun, tout
strength DEF badly channeled that IPFV do that all person all
direksyon politik ka vini plito pou profité]
guidance political IPFV come rather to benefit
'the misused strength making that everyone, every leadership comes as a
profiteer'
- (3) moun la [ki ni enfòmasyon pa ka réperkité y]
person DEF who have information NEG IPFV {pass on} 3SG
'the guy who has information doesn't pass it on to others'

▶▶▶▶ No enclitic *-la*.

Enclitic *-la* found twice (over more than a hundred RC with definite antecedent):

- (4) fo nou osi pa oubliyé transmèt jénérasyon la ki ka vin
must 1PL too NEG forget convey generation DEF that IPFV come
dèyè la
after ENCL
'we must not forget to pass it on the next generation'
- (5) sé té foto a timoun la ki té anlè gato la
it was picture of child DEF that was on cake ENCL
'it was the picture of the child that was on the cake'

Example (5) is quite equivocal.

Analysis

The Definite Hypothesis

As shown previously, the enclitic *-la* is usually presented as a second marking of the definite ("article de rappel"). But:

- it doesn't appear with a demonstrative NP:

(6) nou ni produi **lasa** ki ja kontrolé
1PL have product DEM that already controlled
'we have this product which has been already controlled'

- doesn't appear with *la* and *lé* (definites):

(7) sa k'ay <plus ou moins tué> la sosyété mari-galantèz mèm
it IPFV-go more or less kill the community Marie-Galantese very
ki té on pèp fiyè, on pèp travayè
that PST a people proud a people industrious
'This, more or less, killed the Marie-Galantese community, which is a proud
and industrious people'

(Frenchified example)

The Definite Hypothesis (2)

- (8) an toujou éséyé pran lé désizyon ki fo
1SG always try take DEF.PL decision that need
'I always try to take the right decisions'

The Definite Hypothesis (2)

- (8) an toujou éséyé pran lé désizyon ki fo
1SG always try take DEF.PL decision that need
'I always try to take the right decisions'

▣▶ The arguments in favor of a double definite marking are weak.

☞ Ploog (2006):

*" De nombreux auteurs s'intéressant au français parlé ont évoqué l'omni- présence du marquage d'un constituant nominal par la forme enclitique [la] ; il en est de même en français abidjanais. [...] Dès 1968, LA est également identifié, dans l'emploi par les écoliers abidjanais, comme marqueur emphatique mineur (Herault 1968:99 cité par Hattiger 1983:81), ce qui est corroboré par l'observation de Canut (1998), qui attribue à ce LA - tout aussi massivement présent dans son corpus enregistré au Mali - la fonction d'un **marqueur de clôture**. "*

la has various uses in GC:

- (9) a. a kilè ou ké **la** dimanch ?
at what-time 2SG PROSP there Sunday
'When will you be there on Sunday?'
- b. sé **la** ou ka vwè fyèté an nou
it-is there 2SG IPFV see pride of us
'It's there that you see our pride'
- c. kom an té ka di toutalè **la** nou ka paré dombré
like 1SG PST IPFV say earlier there 1PL IPFV prepare dombré
'As I said earlier, we prepare the dombré (typical dish)'
- d. sé ti boug **la**, yo ka fè moun chié **la**
PL little guy DEF 3PL IPFV make person piss-off there
'These guys, they piss people off'

From the data in Glaude (2013) :

- *la* is not mandatory (anymore?) as RC enclitic ('article de rappel'),
- *la* as discourse-marker?
- can appear with definite RC

TAG Grammar

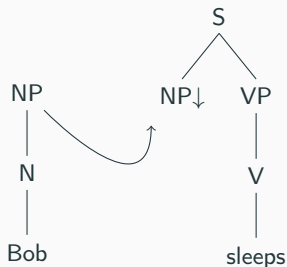
Quick Reminder: TAG in Two Slides

The basic elements of a TAG grammar are *elementary trees* that combine to form sentences. Elementary trees encode information about the co-occurrence restrictions on words. ↗ Joshi and Schabes (1997); Joshi (1987); Abeillé (2002)

Two operations: Substitution \ Adjunction

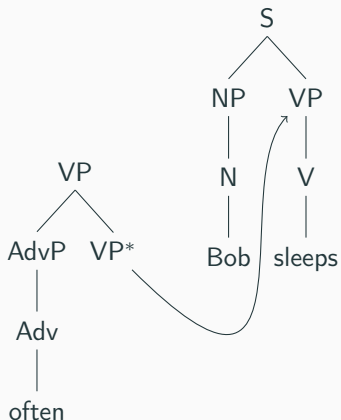
(10) Bob sleeps.

Substitution:



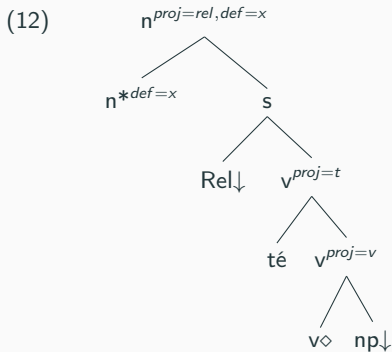
TAG in Two Slides

Adjoining: inserting a tree into another elementary tree.



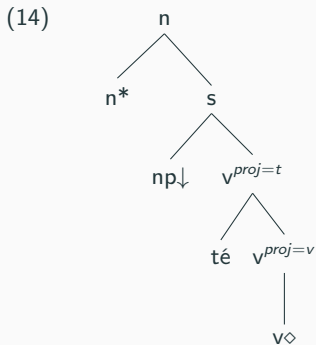
Subject Relative Clauses

- (11) moun [ki té bwè limonad]
 person that PST drink lemonade
 'The person who drank the lemonade'



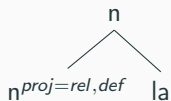
Object Relative Clause

- (13) biten ou té manjé
thing 2SG PST eat
'The thing you've eaten'



Treated as an adjunct to the Subject Relative Clause.

(15)

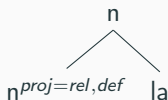


Advantages:

- enclitic *la* treated separately from other markers

Treated as an adjunct to the Subject Relative Clause.

(15)

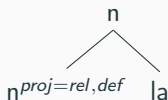


Advantages:

- enclitic *la* treated separately from other markers
- can appear with Relative Clauses with a definite antecedent

Treated as an adjunct to the Subject Relative Clause.

(15)



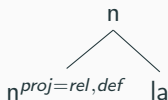
Advantages:

- enclitic *la* treated separately from other markers
- can appear with Relative Clauses with a definite antecedent
- allows for an analysis of Relative Clauses without *la*

Enclitic *-la*

Treated as an adjunct to the Subject Relative Clause.

(15)



Advantages:

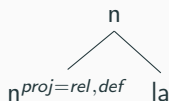
- enclitic *la* treated separately from other markers
- can appear with Relative Clauses with a definite antecedent
- allows for an analysis of Relative Clauses without *la*

Risk:

- may overgenerate (covers Object Relative Clauses)

Treated as an adjunct to the Subject Relative Clause.

(15)



Advantages:

- enclitic *la* treated separately from other markers
- can appear with Relative Clauses with a definite antecedent
- allows for an analysis of Relative Clauses without *la*

Risk:

- may overgenerate (covers Object Relative Clauses)

But ...

can be solved by restricting to Subject Relative Clauses (Restricted Adjoining)

Example

(16)

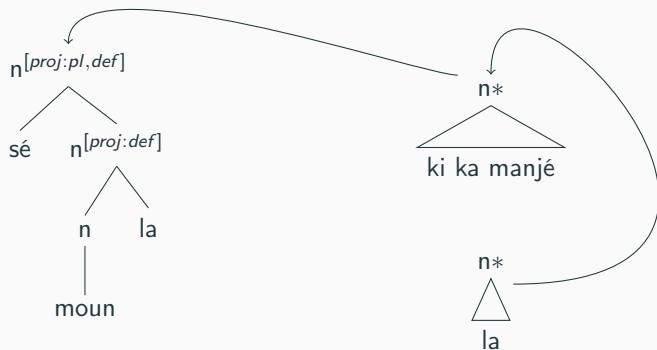


Fig. Assembling Elementary Trees for "Sé moun la ki ka manjé la"

Conclusion

Conclusion

I've presented a formal solution for the description of Relative Clauses with (and without) the enclitic *la* in Gwadeloupéyen:

Conclusion

I've presented a formal solution for the description of Relative Clauses with (and without) the enclitic *la* in Gwadeloupéyen:

- based on a freely accessible corpus (↗ Glaude (2013))

Conclusion

I've presented a formal solution for the description of Relative Clauses with (and without) the enclitic *la* in Gwadeloupéyen:

- based on a freely accessible corpus (↗ Glaude (2013))
- aims at covering GC texts

Conclusion

I've presented a formal solution for the description of Relative Clauses with (and without) the enclitic *la* in Gwadeloupéyen:

- based on a freely accessible corpus (↗ Glaude (2013))
- aims at covering GC texts
- formal description is implemented and freely accessible (and customizable).

Conclusion

I've presented a formal solution for the description of Relative Clauses with (and without) the enclitic *la* in Gwadeloupéyen:

- based on a freely accessible corpus (↗ Glaude (2013))
- aims at covering GC texts
- formal description is implemented and freely accessible (and customizable).

You can:

- improve it (coverage, adequacy, etc.)

Conclusion

I've presented a formal solution for the description of Relative Clauses with (and without) the enclitic *la* in Gwadeloupéyen:

- based on a freely accessible corpus (👉 Glaude (2013))
- aims at covering GC texts
- formal description is implemented and freely accessible (and customizable).

You can:

- improve it (coverage, adequacy, etc.)
- hack it and use another formalism (HPSG, LFG, etc.)

Conclusion

I've presented a formal solution for the description of Relative Clauses with (and without) the enclitic *la* in Gwadeloupéyen:

- based on a freely accessible corpus (👉 Glaude (2013))
- aims at covering GC texts
- formal description is implemented and freely accessible (and customizable).

You can:

- improve it (coverage, adequacy, etc.)
- hack it and use another formalism (HPSG, LFG, etc.)
- use it as a start to build your grammar (for another creole language)

Conclusion

I've presented a formal solution for the description of Relative Clauses with (and without) the enclitic *la* in Gwadeloupéyen:

- based on a freely accessible corpus (👉 Glaude (2013))
- aims at covering GC texts
- formal description is implemented and freely accessible (and customizable).

You can:

- improve it (coverage, adequacy, etc.)
- hack it and use another formalism (HPSG, LFG, etc.)
- use it as a start to build your grammar (for another creole language)
- use it as a grammatical resource and develop tools from it

Thank you for your attention



References

- Abeillé, Anne. 2002. *Une Grammaire électronique du Français*. CNRS Editions, Paris.
- Baptista, Marlyse. 2003. *The syntax of Cape Verdean creole: The sotavento varieties*, volume 54. John Benjamins Publishing.
- Bernabé, J. 1983. *Fondal-natal*. l'Harmattan Paris.
- Damoiseau, Robert. 2012. *Syntaxe créole comparée*, karthala et CNDP-CRDP edition.
- Ferraz, Luiz Ivens. 1978. The creole of São Tomé. *African studies*, 37(1):3–68.

Bibliography II

- Glaude, Herby. 2013. Corpus Créoloral. oai:crdo.vjf.cnrs.fr:crdo-GCF, SFL Université Paris 8 - LLL Université Orléans.
- Joshi, Aravind K. 1987. An introduction to Tree Adjoining Grammars. In A. Manaster-Ramer, editor, *Mathematics of Language*. John Benjamins, Amsterdam, pages 87–114.
- Joshi, Aravind K. and Yves Schabes. 1997. Tree-Adjoining Grammars. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, volume 3. Springer, Berlin, New York, pages 69–124.
- Kihm, Alain. 1994. *Kriyol syntax: the Portuguese-based creole language of Guinea-Bissau*, volume 14. John Benjamins Publishing.
- Kouwenberg, Silvia. 1994. *A grammar of Berbice Dutch creole*, volume 12. Walter de Gruyter.

Bibliography III

- Ludwig, Ralph and Stefan Pfänder. 2003. La particule là/la en français oral et en créole caribéen: grammaticalisation et contact de langues. *Approche de la variation créole et française. Grammaticalisation et réanalyse*, pages 269–284.
- Maurer, Philippe. 1995. *L'angolar: un créole afro-portugais parlé à São Tomé: notes de grammaire, textes, vocabulaires*, volume 16. Buske Verlag.
- Petitjean, Simon, Denys Duchier, and Yannick Parmentier. 2016. Xmg 2: Describing description languages. In *Logical Aspects of Computational Linguistics. Celebrating 20 Years of LACL (1996–2016) 9th International Conference, LACL 2016, Nancy, France, December 5-7, 2016, Proceedings*, pages 255–272, Springer.
- Ploog, Katja. 2006. Du continuum pragmatico-sémantique aux types prosodiques de la en (français) abidjanais. *Hommage à Suzanne Lafage, Le français en Afrique*, (21):303–323.