

Automatic Annotation of Motion Expressions and Place Named Entities

2nd UNITEX/GramLab Workhop

10 octobre 2013

Ludovic Moncla

ludovic.moncla@univ-pau.fr



IAAA
Sistemas de Información
Avanzados
Universidad Zaragoza



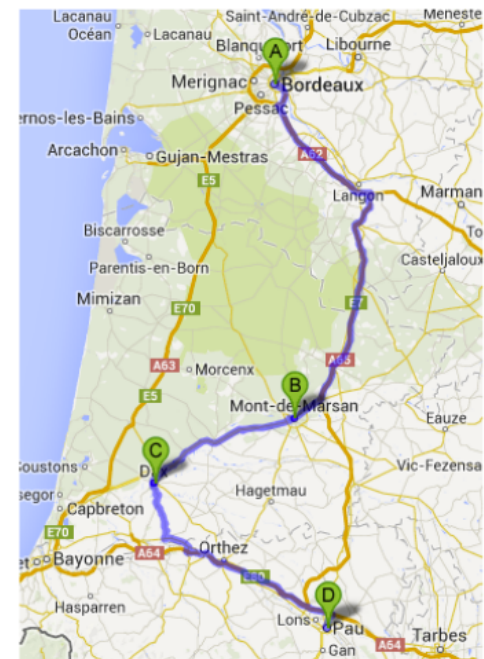
Automatic reconstruction of itineraries from text corpus

Extraction



+

Representation



Named place entities, spatial relation, motion expressions ...

Problems and background

Extraction of motion expressions and place named entities

- Expression of movement is essential in travelogues (Boons, 1987; Sarda, 2000)
- "Travelogue verbs" = verbs of displacement + motion verbs + verbs of perception
- "Extended toponyms" = topographical terms + spatial relation + spatial entity
- "VT structures" = travelogue verbs + extended toponym (Nguyen et al, 2013)

VT structures mark and formalize relations between travelogue verbs and extended toponyms

Method and implementation

Finite-state transducers cascade

- Parse natural language and extract spatial named entities, travelogue verbs and nominal groups having a topographical denotation
- Using Casys available on the Unitex plate-form like for the CasEn system (Maurel et al, 2011)
- We propose a demonstration tool online available here : <http://t2i.univ-pau.fr/GEONTO/>

INPUT

text



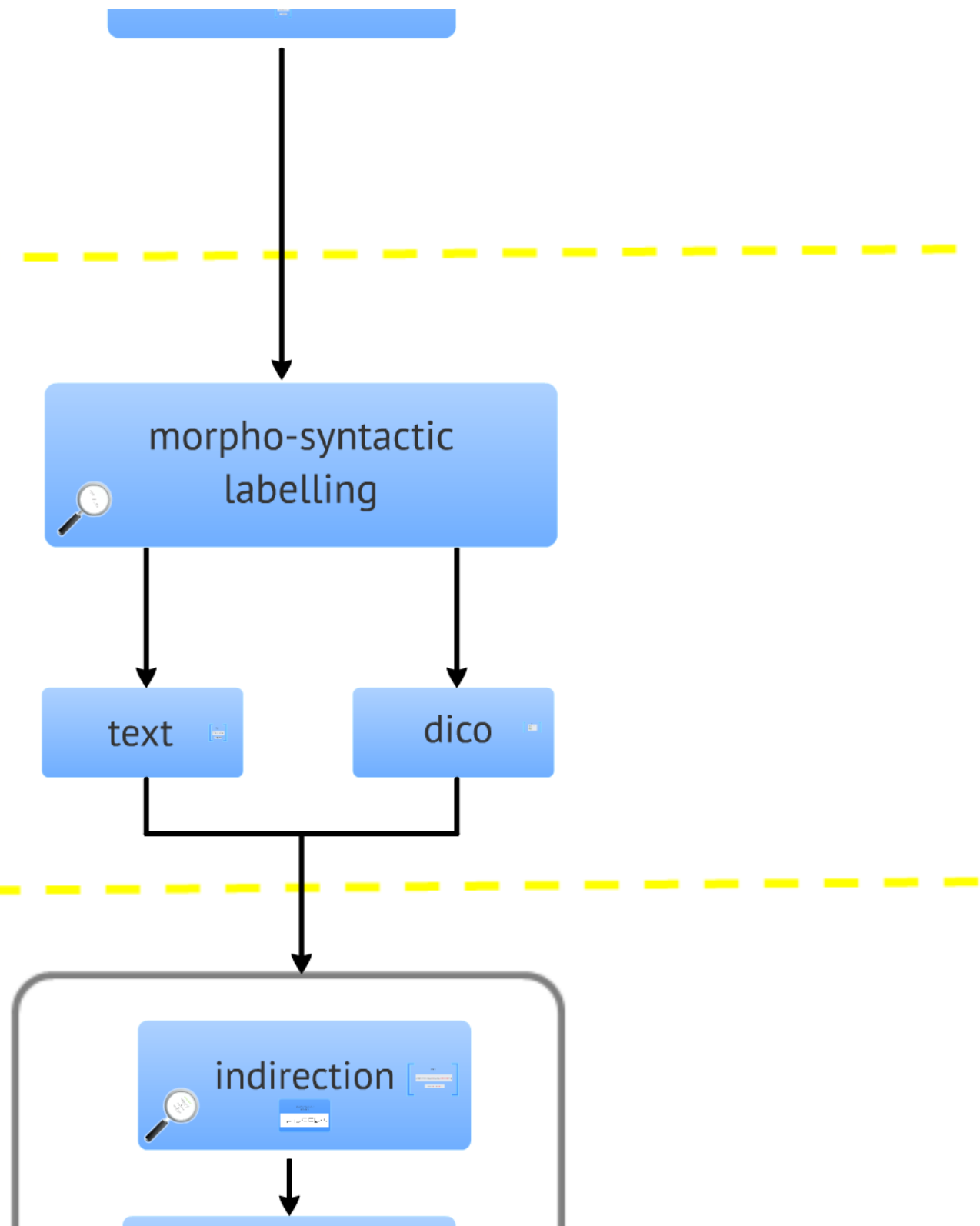
morpho-syntactic
labelling

Example

Je traverse les villages au sud de Pau.

I cross the villages south of Pau.

PRE -PROCESSING

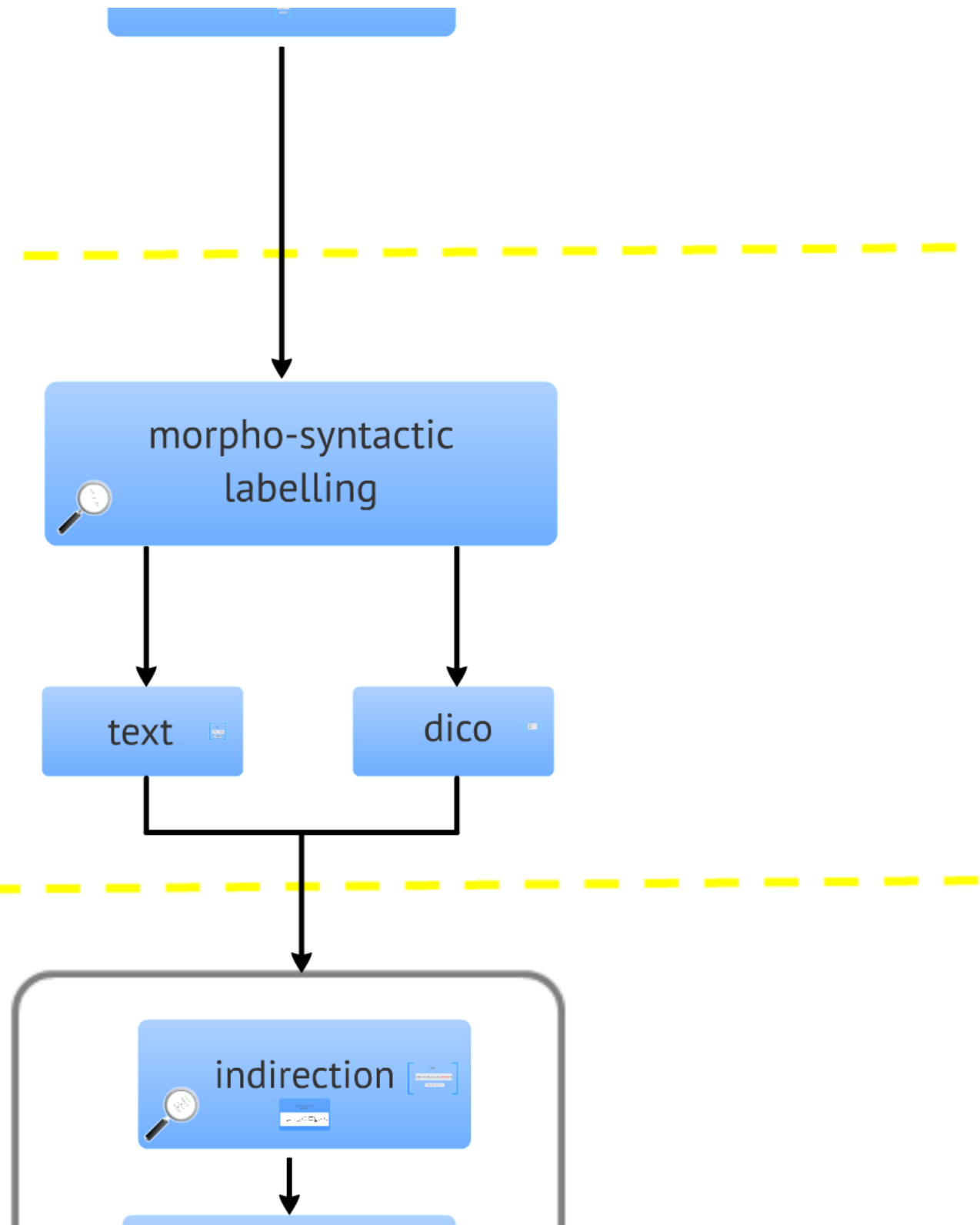




TreeTagger
[Schmid, 1994]

MElt
[Denis & Sagot 2009]

PRE -PROCESSING

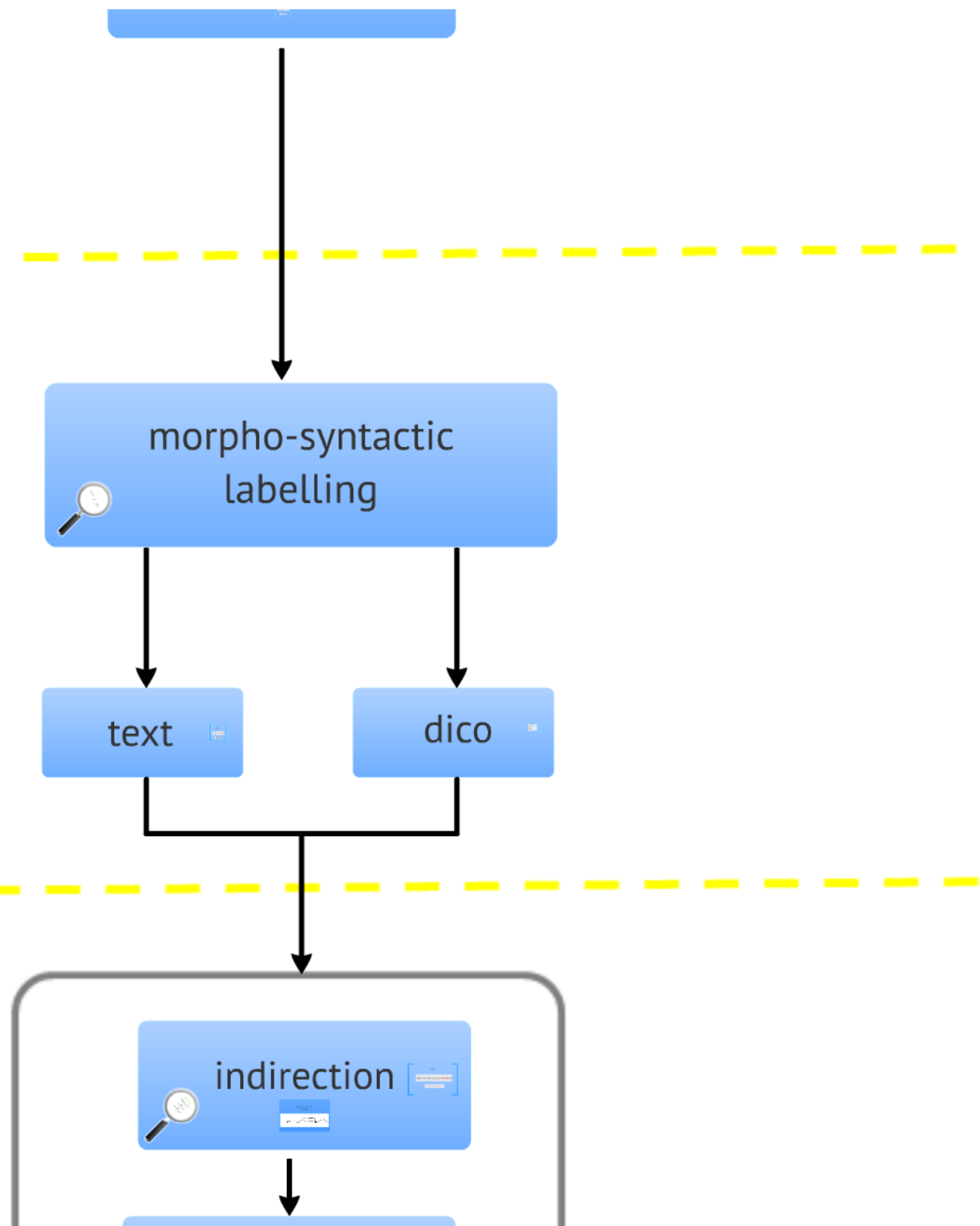


Example

Je/PRO traverse/V les/DET villages/N
au/PREPDET sud/N de/PREP Pau/NPR .

I/PRO cross/V the/DET villages/N south/N
of/PREP Pau/NPR.

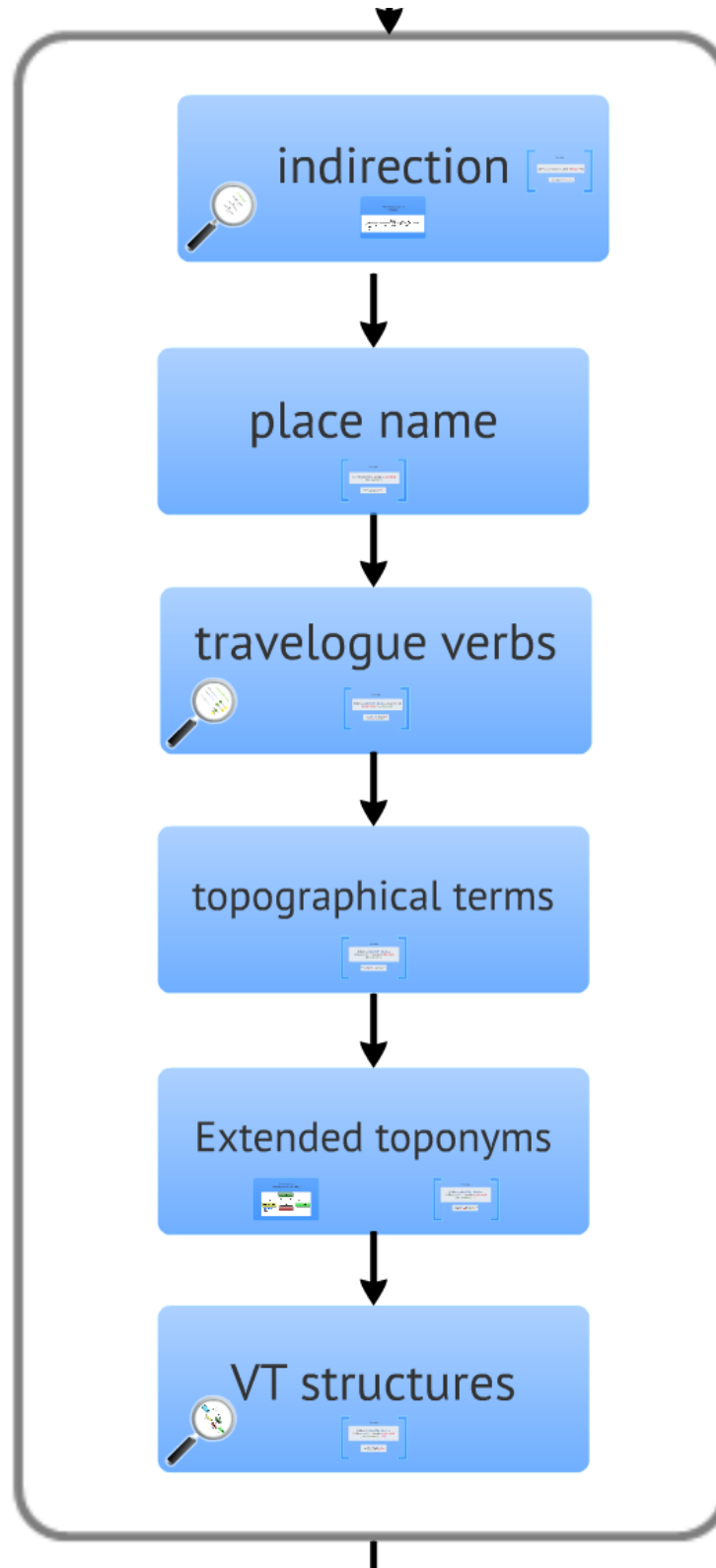
PRE -PROCESSING



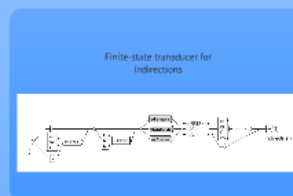
Example

Je,.PRO+PpvIL+Analyseur
traverse,traverser.V+Analyseur
les,.DET+ART+Analyseur
villages,.N+Analyseur
au,.PREPDET+Analyseur
sud,.N+Analyseur
de,.PREP+Analyseur
Pau,.N+PR+NPr+Analyseur

FINITE-STATE TRANSDUCERS CASCADE



indirection



Example

Je traverse les villages (au sud de **indirection**) Pau.

Je traverse les villages (au sud de **indirection**) Pau.

Spatial relation

topological relation

[Egenhofer & Franzosa, 1991]



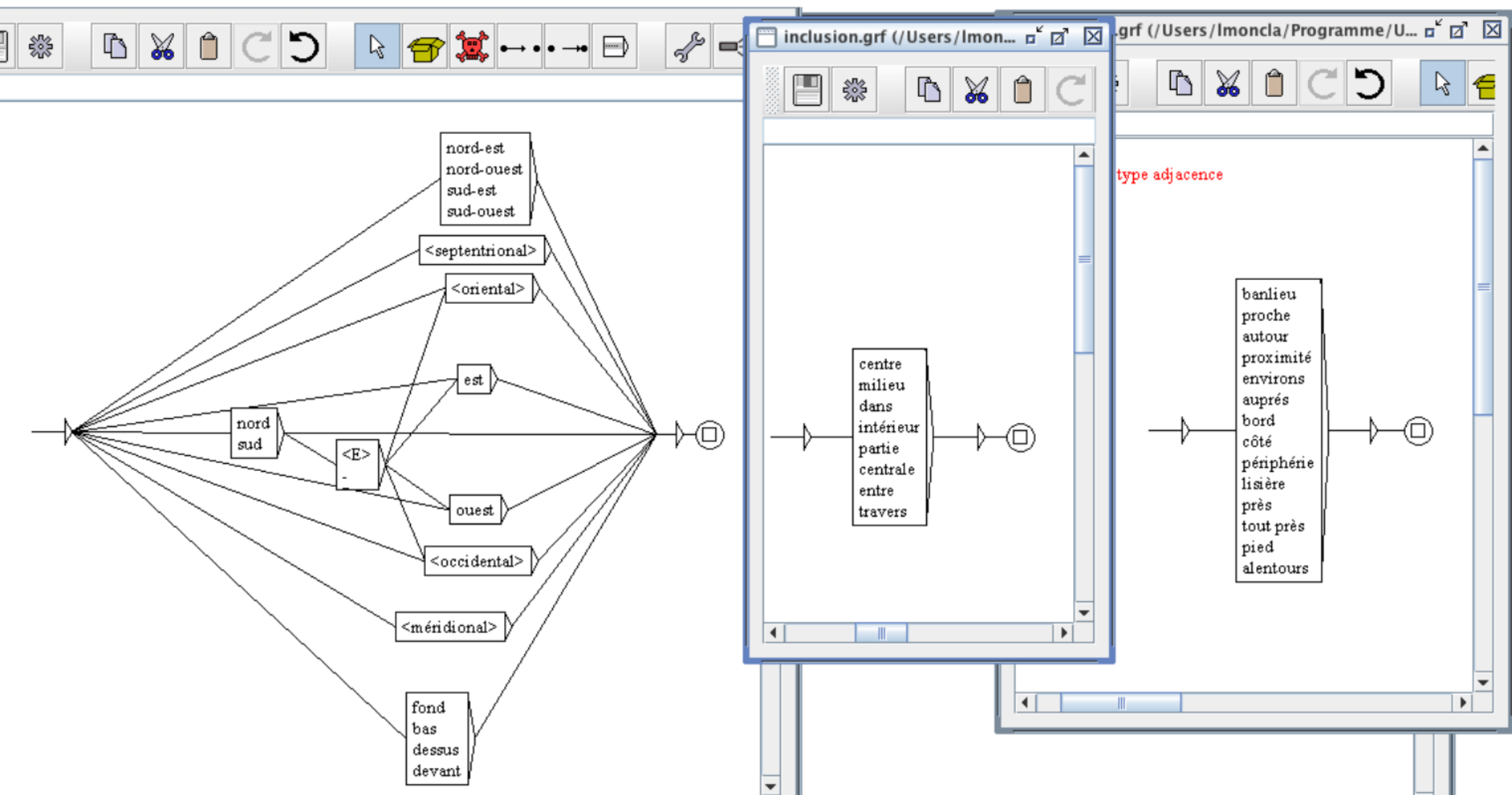
distance relation



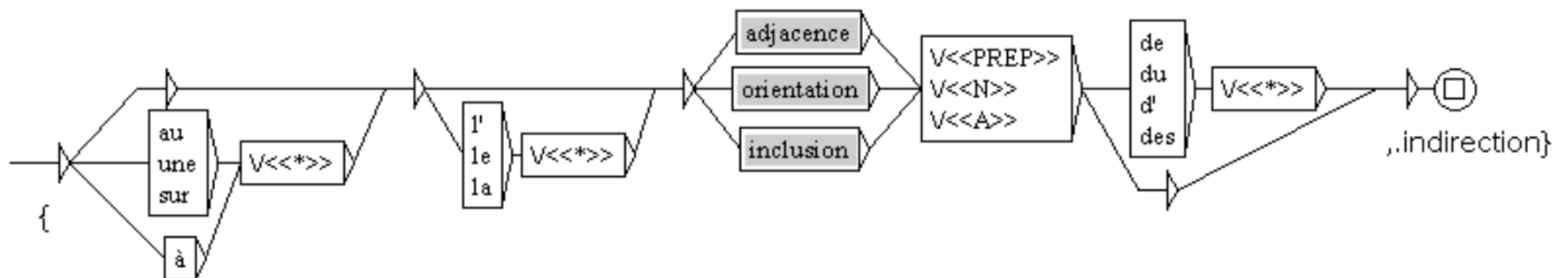
directional relation

[Ligozat, 1998]





Finite-state transducer for indirections



Example

Je traverse les villages {au sud de,.indirection} Pau.

I cross the villages {south of,.indirection} Pau.



place name



Example

Je traverse les villages {au sud de,.indirection}
{Pau,.placeName}.

I cross the villages {south of,.indirection}
{Pau,.placeName}.



travelogue verbs



Polarity of motion verbs

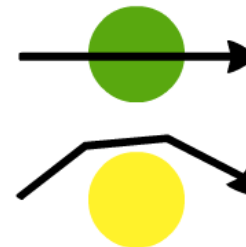
- initial verbs : to leave



- final verbs : to arrive



- median verbs : to cross



Example

Je {traverse,.travelVerb+median} les villages {au sud
de,.indirection} {Pau,.placeName}.

I {cross,.travelVerb+median} the villages {south
of,.indirection} {Pau,.placeName}.



topographical terms

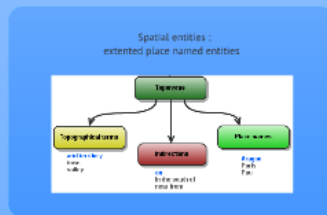


Example

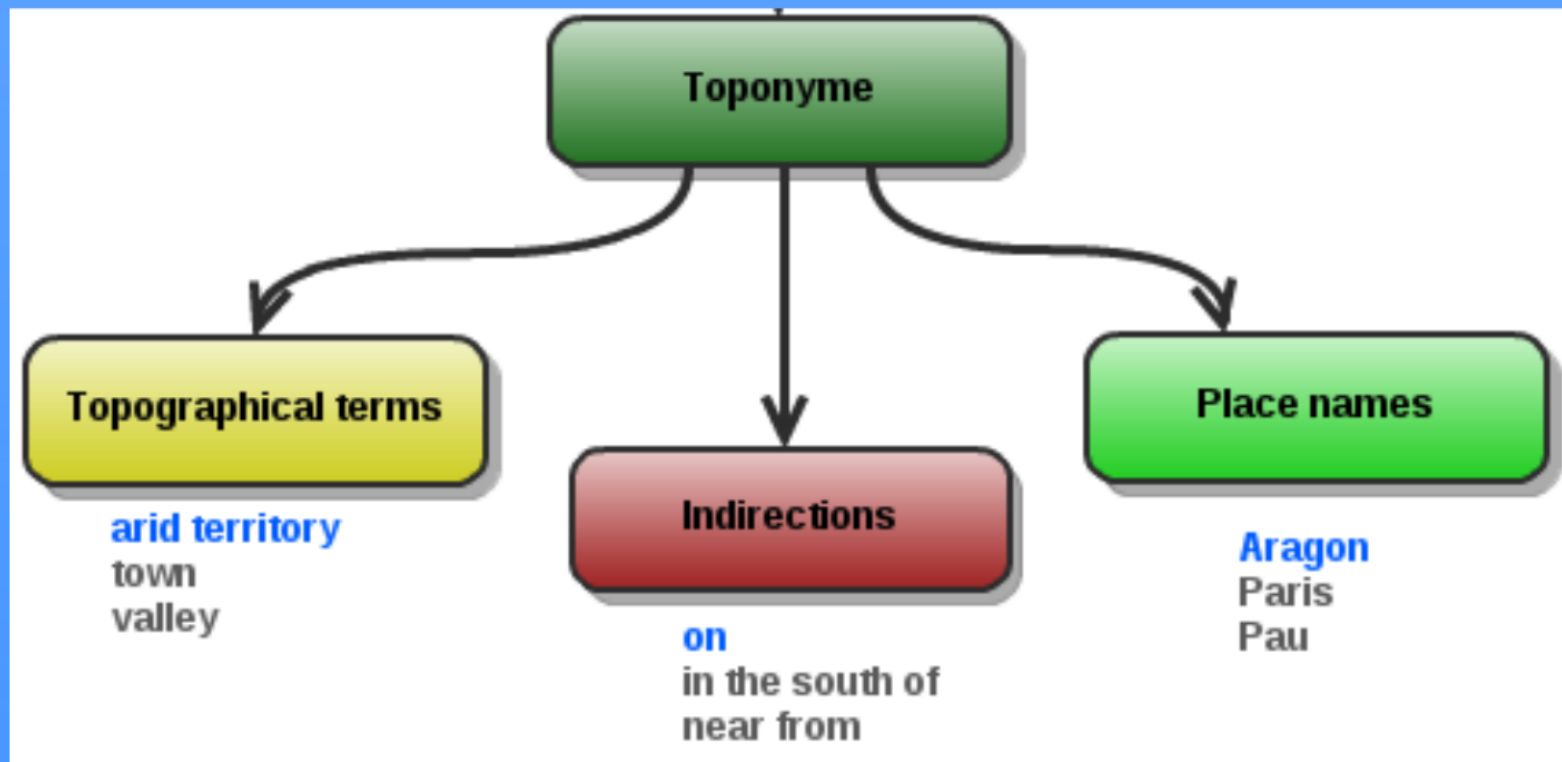
Je {traverse,.travelVerb+median} les
{villages,.topoTerm} {au sud de,.indirection}
{Pau,.placeName}.

I {cross,.travelVerb+median} the {villages,.topoTerm}
{south of,.indirection} {Pau,.placeName}.

Extended toponyms



Spatial entities : extended place named entities



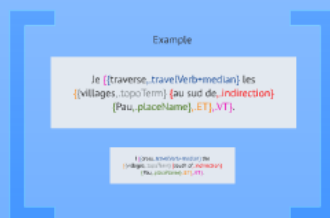
Example

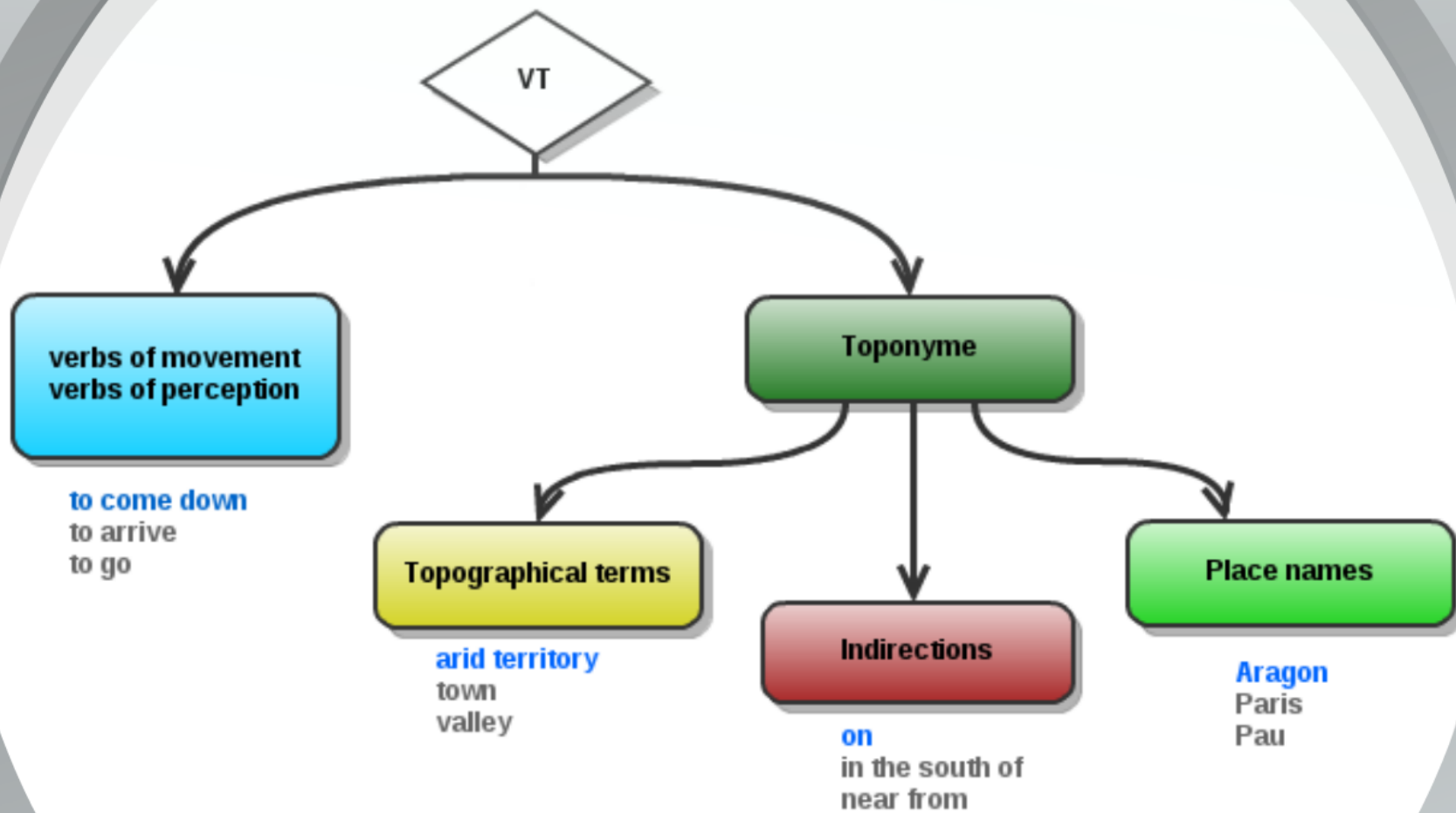
Je {traverse,.travelVerb+median} les
{villages,.topoTerm} {au sud de,.indirection}
{Pau,.placeName},.ET}.

I {cross,.travelVerb+median} the {villages,.topoTerm}
{south of,.indirection} {Pau,.placeName},.ET}.



VT structures



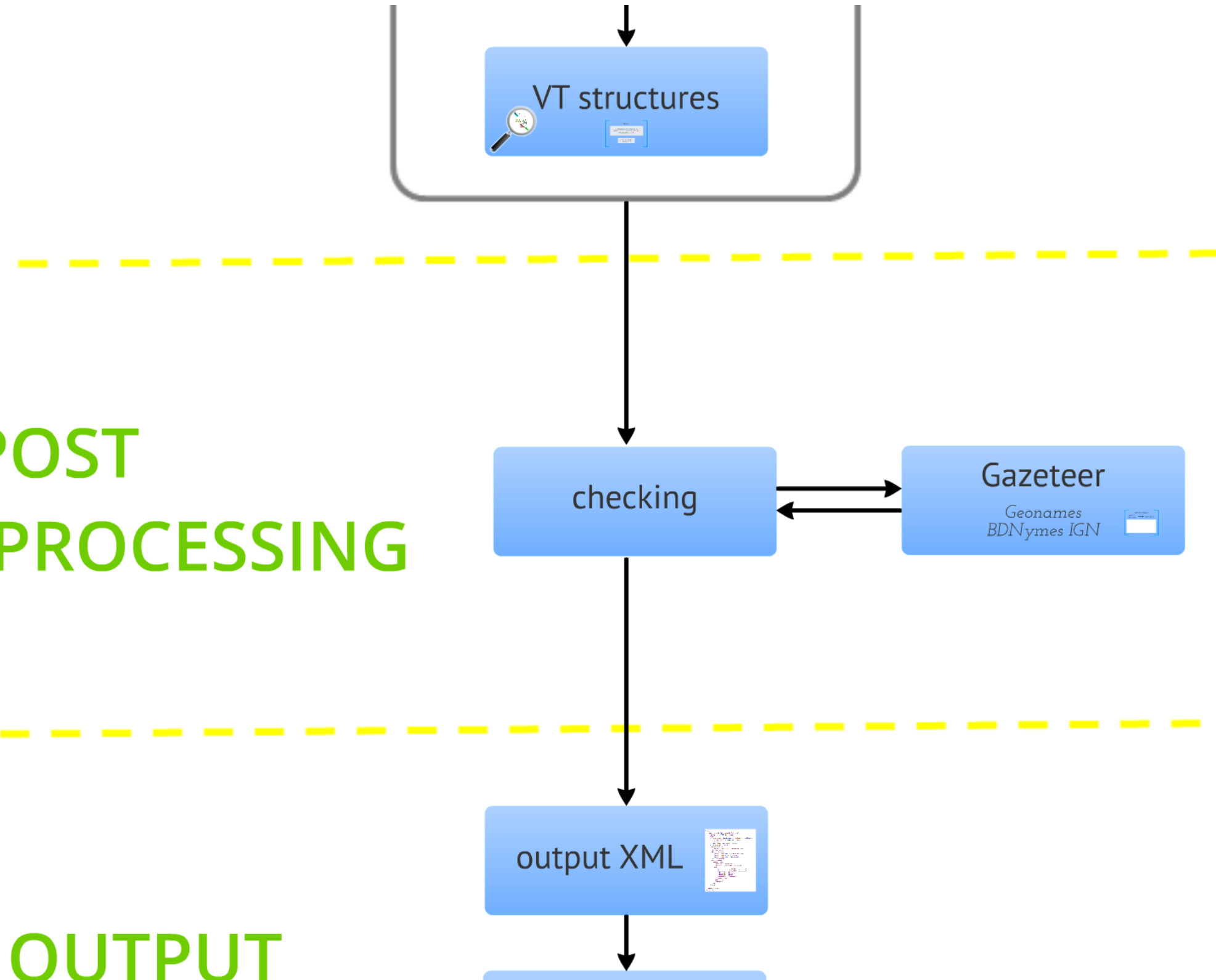


Example

Je {{traverse,.travelVerb+median} les
{{villages,.topoTerm} {au sud de,.indirection}
{Pau,.placeName},.ET},.VT}.

I {{cross,.travelVerb+median} the
{{villages,.topoTerm} {south of,.indirection}
{Pau,.placeName},.ET},.VT}.

POST -PROCESSING



OUTPUT

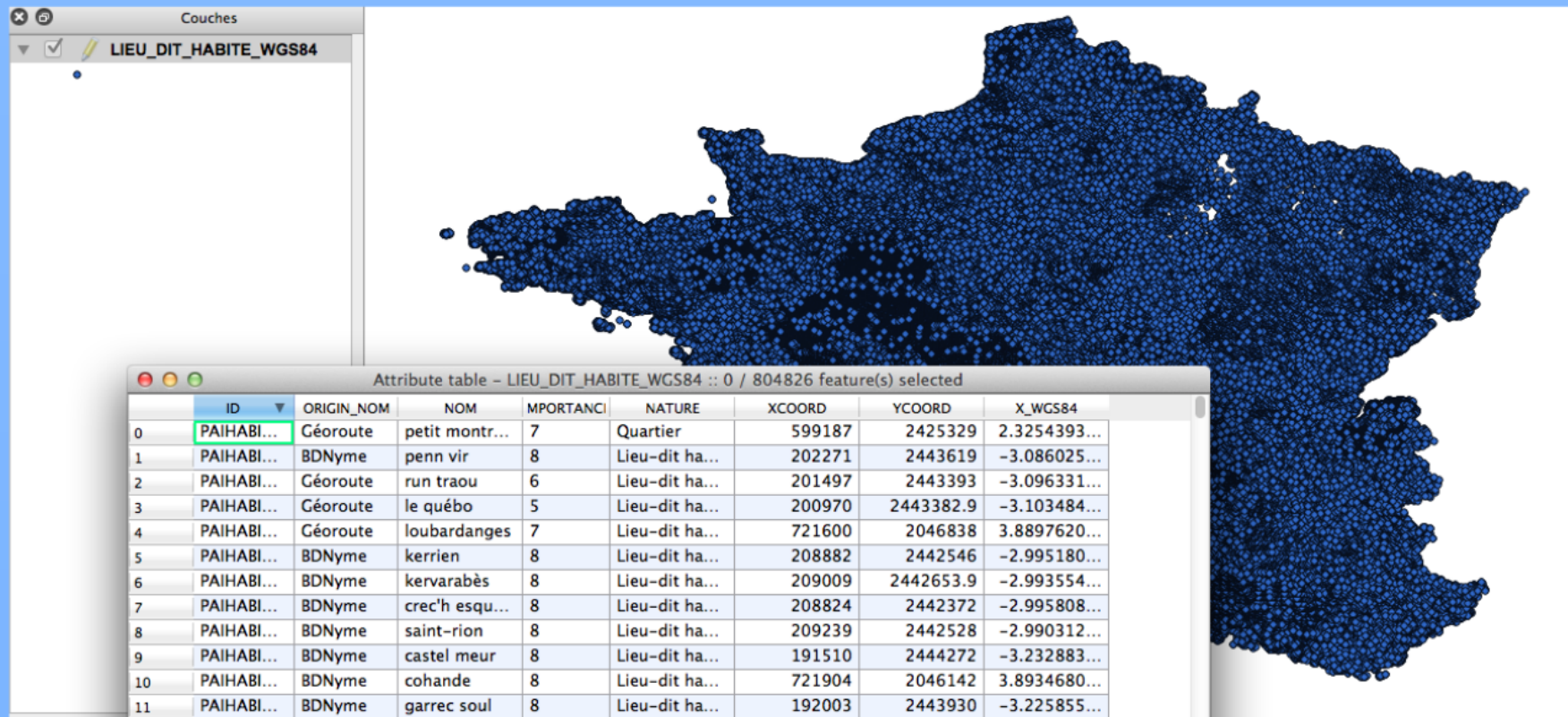
External resources

Geonames

BDnyme (IGN)

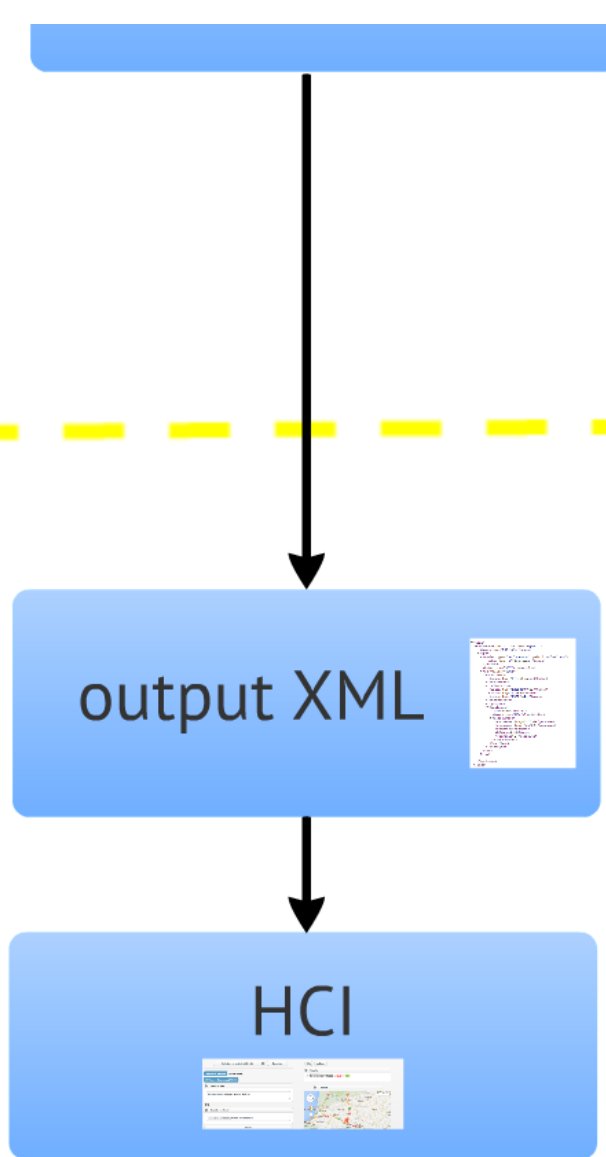


Postgis database



PROCESSING

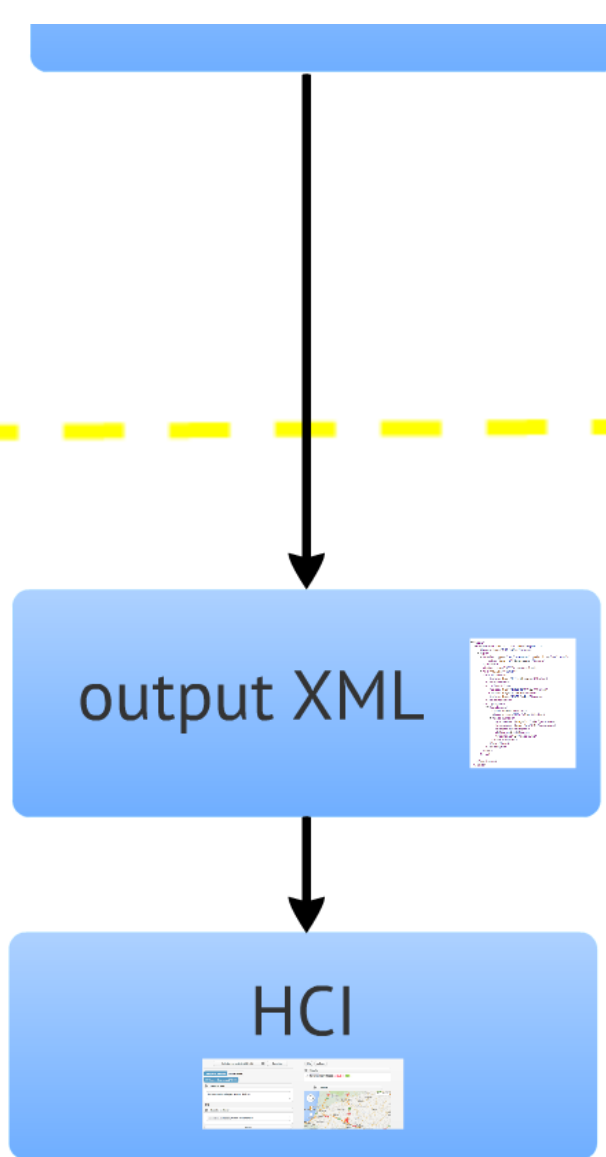
OUTPUT



```
▼<text>
  ▼<sentence id="n1" xml:id="dlpls1">
    <token tag="PRO">Je</token>
    ▼<vpe>
      ▼<verbe type="deplacement" polarite="medians">
        <token tag="V">traverse</token>
      </verbe>
      <token tag="DET">les</token>
    ▼<es valide="true">
      ▼<nomCommun>
        <token tag="N">villages</token>
      </nomCommun>
      ▼<indirection>
        <token tag="PREPDET">au</token>
        <token tag="N">sud</token>
        <token tag="PREP">de</token>
      </indirection>
      ▼<toponyme>
        ▼<confirme>
          <value>Pau</value>
          <token tag="NPr">Pau</token>
        ▼<ressources>
          <geonames lang="fr">3</geonames>
          <geonames lang="es">2</geonames>
          <bdnyme>9</bdnyme>
          <bdtopo>2</bdtopo>
          <georoute>2</georoute>
        </ressources>
      </confirme>
    </toponyme>
  </es>
</vpe>
.
</sentence>
</text>
```

PROCESSING

OUTPUT



Sélectionnez un texte déjà traité

Consulter

TreeTagger + Unitex

MELt + Unitex

BDNyme

Géonames (FR+ES)

Entrer un texte

Je traverse les villages au sud de Pau.

- OU -

Uploader un fichier

Choisissez un fichier

Aucun fichier choisi

Valider

XML

IsoSpace

Résultat

Je traverse les villages au sud de Pau.

Légende



Examples

1. On découvre des paysages extraordinaires dans le sud de la France.
2. Nous descendons sur le territoire aride au sud de la région d' Aragon.
3. Depuis la place de la mairie de Bedous longer la façade sud de l' église
4. Vous rejoignez un chemin goudronné qui franchit un pont sur le Gabarret.
5. Nous descendîmes assez lentement vers le pied de la Maladetta.
6. En se promenant en Corse, on retrouve des paysages extraordinaires : des falaises à vous couper le souffle, des sentiers inoubliables.

Future work

Improve the model and the VT structures

- Adding patterns
- Improve toponyms desambiguation (mix method map-based and data-driven based)

Representation of the itinerary on a map

Make available for Spanish and English

Build a manually annotated corpus of travelogues and hiking for evaluation

Automatic Annotation of Motion Expressions and Place Named Entities

Thank you!

Ludovic Moncla

ludovic.moncla@univ-pau.fr



IAAA
Sistemas de Información
Avanzados
Universidad Zaragoza

