NLP Tools for Mental Hearth – French experience

Maxime Amblard, Michel Musiol

▶ **To cite this version:**

Maxime Amblard, Michel Musiol. NLP Tools for Mental Hearth – French experience. French-Sweden Scientific Day, Nov 2021, stockholm, Sweden. hal-03619164

HAL Id: hal-03619164
https://inria.hal.science/hal-03619164

Submitted on 27 Mar 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NLP Tools for Mental Hearth – French experience

Maxime Amblard - Michel Musiol

November 15th 2021

Main Objectives

- Complete the classic paradigm of Language and Thought Disorders (LTD)

Main Objectives

- Complete the classic paradigm of Language and Thought Disorders (LTD)
- By a semantic-formal model of Speech Disorders (SDD)

Main Objectives

- Complete the classic paradigm of Language and Thought Disorders (LTD)
- By a semantic-formal model of Speech Disorders (SDD)
- Develop computerized tools of aids in the diagnosis of schizophrenia

Main Objectives

- Complete the classic paradigm of Language and Thought Disorders (LTD)
- By a semantic-formal model of Speech Disorders (SDD)

- Develop computerized tools of aids in the diagnosis of schizophrenia
- Adapt and use these tools to address the problem of screening for the onset of the illness and follow-up of the pathology

Understanding language data : semantics and dialogue

1. Formalize

2. Compare with real data

Understanding language data : semantics and dialogue

1. Formalize
   - Logical representations (Montagovian tradition)

$$\exists x.(\exists y.(\textbf{farmer}\ x \wedge \textbf{donkey}\ y \wedge \textbf{own}\ x\ y) \rightarrow \textbf{beat}\ x\ y)$$

[Qian, de Groote, Amblard LiLT 2016]

[Boritchev, Amblard, LCM 2021, SemDial 2020, 2019]

   - Define calculations for semantics $(\!|\lambda|\!)$

[Marsik, Amblard, de Groote TCS 2021]

2. Compare with real data

Understanding language data : semantics and dialogue

1. Formalize
   - Logical representations (Montagovian tradition)

$$\exists x.(\exists y.(\textbf{farmer}\ x \wedge \textbf{donkey}\ y \wedge \textbf{own}\ x\ y) \rightarrow \textbf{beat}\ x\ y)$$

[Qian, de Groote, Amblard LiLT 2016]

[Boritchev, Amblard, LCM 2021, SemDial 2020, 2019]

   - Define calculations for semantics $(\!|\lambda|\!)$

[Marsik, Amblard, de Groote TCS 2021]

2. Compare with real data
   - Failure to understand : schizophrenic patients

G82 l'an dernier euh... j'savais pas comment faire j'étais perdue

V87 Vous savez ça arrive à tout le monde d'avoir des moments biens et des moments où on est perdu

G88 Oui j'ai peur de perdre tout le monde

[Amblard, Musiol, Rebuschi LCM 2021]

Understanding language data : semantics and dialogue

1. Formalize
   - Logical representations (Montagovian tradition)

$$\exists x.(\exists y.(\textbf{farmer}\ x \wedge \textbf{donkey}\ y \wedge \textbf{own}\ x\ y) \rightarrow \textbf{beat}\ x\ y)$$

[Qian, de Groote, Amblard LiLT 2016]

[Boritchev, Amblard, LCM 2021, SemDial 2020, 2019]

   - Define calculations for semantics $(\!|\lambda|\!)$

[Marsik, Amblard, de Groote TCS 2021]

2. Compare with real data
   - Failure to understand : schizophrenic patients

G82 l'an dernier euh... j'savais pas comment faire j'étais perdue

V87 Vous savez ça arrive à tout le monde d'avoir des moments biens et des moments où on est perdu

G88 Oui j'ai peur de perdre tout le monde

[Amblard, Musiol, Rebuschi LCM 2021]

Understanding language data : semantics and dialogue

1. Formalize
   - Logical representations (Montagovian tradition)

$$\exists x.(\exists y.(\textbf{farmer}\ x \wedge \textbf{donkey}\ y \wedge \textbf{own}\ x\ y) \rightarrow \textbf{beat}\ x\ y)$$

[Qian, de Groote, Amblard LiLT 2016]

[Boritchev, Amblard, LCM 2021, SemDial 2020, 2019]

   - Define calculations for semantics $(\!|\lambda|\!)$

[Marsik, Amblard, de Groote TCS 2021]

   ⇒ Wide coverage (semantic calculation and data)
2. Compare with real data
   - Failure to understand : schizophrenic patients

     G82 l'an dernier euh... j'savais pas comment faire j'étais perdue

     V87 Vous savez ça arrive à tout le monde d'avoir des moments biens et des moments où on est perdu

     G88 Oui j'ai peur de perdre tout le monde

[Amblard, Musiol, Rebuschi LCM 2021]

   ⇒ Tooling the analysis, developing resources

Understanding language data : semantics and dialogue

1. Formalize
   - Logical representations (Montagovian tradition)

$$\exists x.(\exists y.(\textbf{farmer}\ x \wedge \textbf{donkey}\ y \wedge \textbf{own}\ x\ y) \rightarrow \textbf{beat}\ x\ y)$$

[Qian, de Groote, Amblard LiLT 2016]

[Boritchev, Amblard, LCM 2021, SemDial 2020, 2019]

   - Define calculations for semantics $(\!|\lambda|\!)$

[Marsik, Amblard, de Groote TCS 2021]

   $\Rightarrow$ Wide coverage (semantic calculation and data)
2. Compare with real data
   - Failure to understand : schizophrenic patients

     G82 l'an dernier euh... j'savais pas comment faire j'étais perdue

     V87 Vous savez ça arrive à tout le monde d'avoir des moments biens et des moments où on est perdu

     G88 Oui j'ai peur de perdre tout le monde

[Amblard, Musiol, Rebuschi LCM 2021]

   $\Rightarrow$ Tooling the analysis, developing resources
3. Integrate digital models (DNN, RNN, transformer, etc.) :
   - identification of the structure of the interaction

[Li, Amblard, Braud TALN 2020]

Cognitive processes and mental pathology

- Identify patients' cognitive difficulties
  - Important clinical issue
  - Difficulty : to access the disorder
  - Discourse as a natural place for the expression of disorder

Cognitive processes and mental pathology

- Identify patients' cognitive difficulties
  - Important clinical issue
  - Difficulty : to access the disorder
  - Discourse as a natural place for the expression of disorder
- Methodology
  - Finely isolating a cognitive disorder using semantic models
  - Link between cognitive and neuro-psychological defects
  - SLAM : explain the decisive breaks with the SDRT

  $\Rightarrow$ Tools for analyzing the disorder through discourse management processes

Cognitive processes and mental pathology

- Identify patients' cognitive difficulties
  - Important clinical issue
  - Difficulty : to access the disorder
  - Discourse as a natural place for the expression of disorder
- Methodology
  - Finely isolating a cognitive disorder using semantic models
  - Link between cognitive and neuro-psychological defects
  - SLAM : explain the decisive breaks with the SDRT

  $\Rightarrow$ Tools for analyzing the disorder through discourse management processes

  $\Rightarrow$ Hypothesis : link between semantic-pragmatic and neuro-pathological disorders
- Contributions
  - Original theory of the disorder
  - Identification of objectivable multimodal clues.
  - Digital tools : transfer to the world of psychiatry

Projects

- SLAM - Schizophrenia and Language : Analyze and Modelling 2010-2017
  - constitution of a corpus - Lyon (18 patients + control corpus)
  - use of formal models
  - automatic analysis of disfluencies

Projects

- SLAM - Schizophrenia and Language : Analyze and Modelling 2010-2017
  - constitution of a corpus - Lyon (18 patients + control corpus)
  - use of formal models
  - automatic analysis of disfluencies

- Action Exploratoire Inria ODiM 2019-2021
  - constitution of a corpus - Aix en Provence (20 patients + control corpus)
  - analysis of dialogical transactions (model and parsing)
  - articulation of neuropsychological and eye-trackers

Projects

- SLAM - Schizophrenia and Language : Analyze and Modelling 2010-2017
  - constitution of a corpus - Lyon (18 patients + control corpus)
  - use of formal models
  - automatic analysis of disfluencies

- Action Exploratoire Inria ODiM 2019-2021
  - constitution of a corpus - Aix en Provence (20 patients + control corpus)
  - analysis of dialogical transactions (model and parsing)
  - articulation of neuropsychological and eye-trackers

- MePheSTO - Inria-DFKI 2020-2022
  - creation of a Franco-German corpus
  - consideration of multimodality
  - use of AI methods

SLAM : Rejouer des ambiguïtés linguistiques

B_{124} Oh ouais et pis compliqué et c'est vraiment très très compliqué la politique c'est quelque chose quand on s'en occupe faut être gagnant parce qu'autrement quand on est perdant c'est fini quoi

A_{125} Oui

B_{126} J. C. D. est mort, L. est mort, P. est mort euh (...)

A_{127} Ils sont morts parce qu'ils ont perdu à votre avis

B_{128} Non ils gagnaient mais si ils sont morts, c'est la maladie quoi c'est c'est

A_{129} Ouais c'est parce qu'ils étaient malades, c'est pas parce qu'ils faisaient de la politique

B_{130} Si enfin

A_{131} Si vous pensez que c'est parce qu'ils faisaient de la politique

B_{132} Oui tiens oui il y a aussi C. qui a accompli un meurtre là il était présent lui aussi qui est à B. mais enfin c'est encore à cause de la politique ça

SLAM : Rejouer des ambiguïtés linguistiques

B_{124} Oh ouais et pis compliqué et c'est vraiment très très compliqué la politique c'est quelque chose quand on s'en occupe faut être gagnant parce qu'autrement quand on est perdant c'est fini quoi

A_{125} Oui

B_{126} J. C. D. est mort, L. est mort, P. est mort euh (...)

A_{127} Ils sont morts parce qu'ils ont perdu à votre avis

B_{128} Non ils gagnaient mais si ils sont morts, c'est la maladie quoi c'est c'est

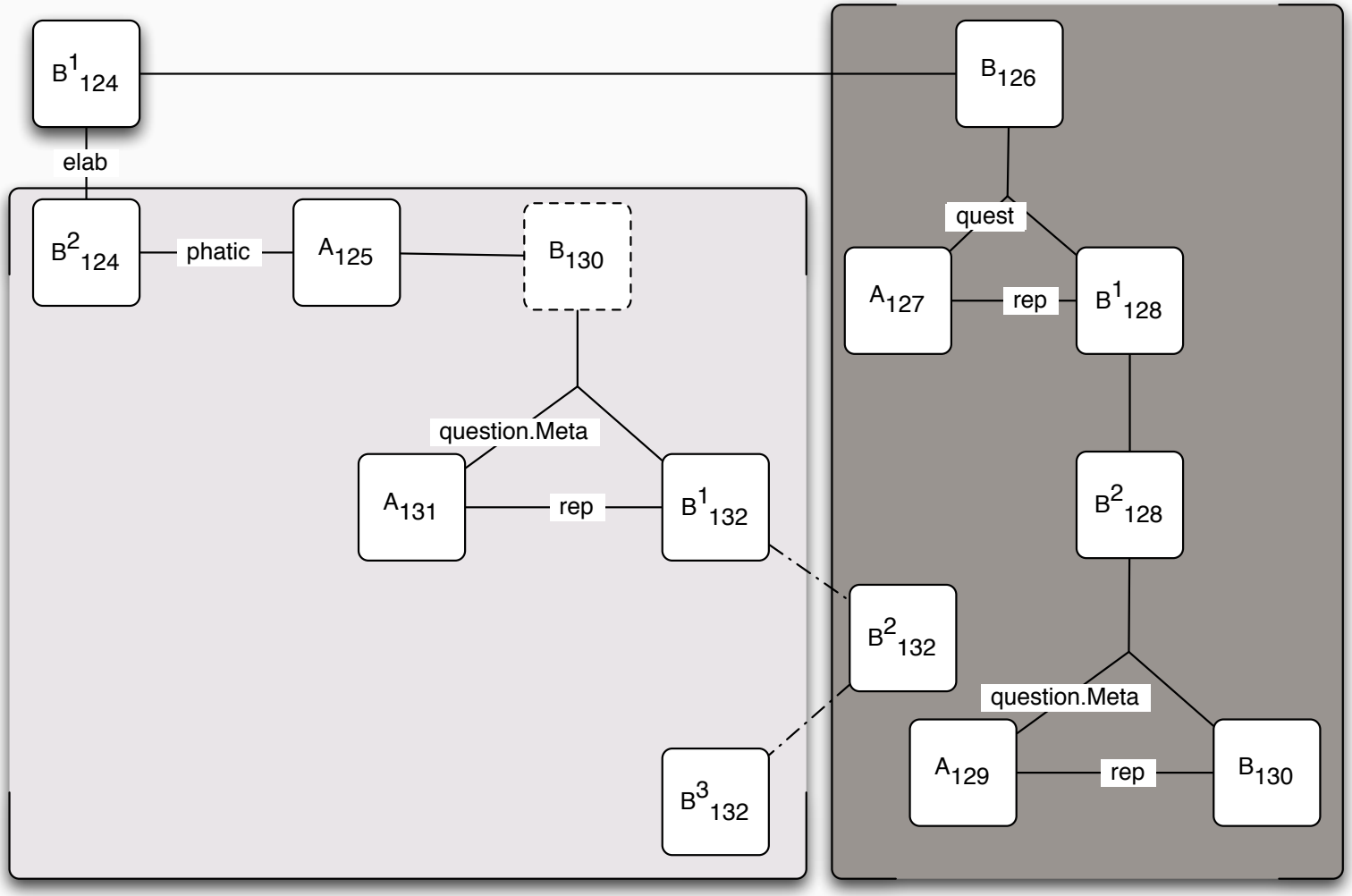
A_{129} Ouais c'est parce qu'ils étaient malades, c'est pas parce qu'ils faisaient de la politique

B_{130} Si enfin

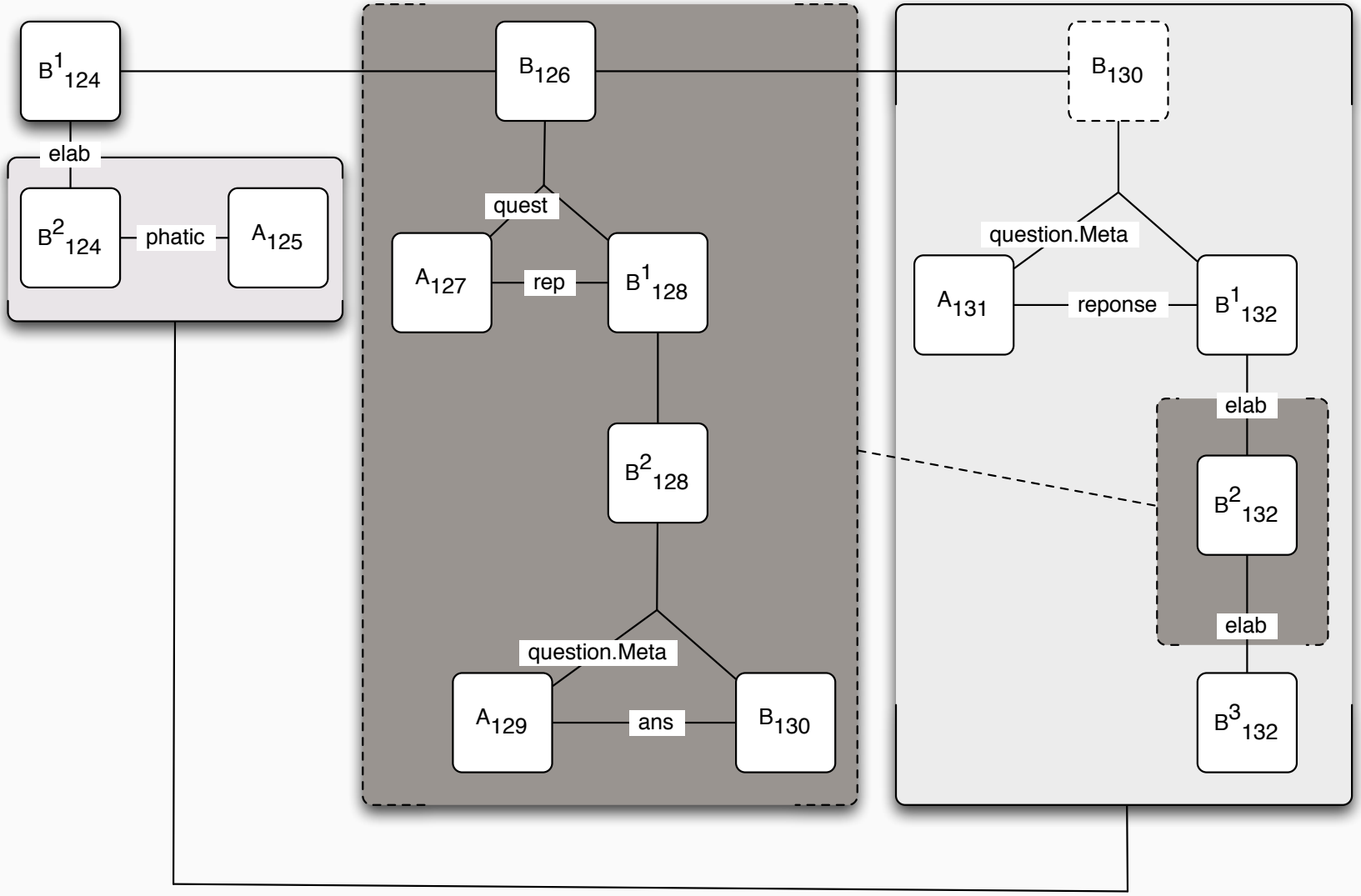
A_{131} Si vous pensez que c'est parce qu'ils faisaient de la politique

B_{132} Oui tiens oui il y a aussi C. qui a accompli un meurtre là il était présent lui aussi qui est à B. mais enfin c'est encore à cause de la politique ça

SLAM : Point de vue du schizophrène



SLAM : Point de vue du psychologue



SLODiM

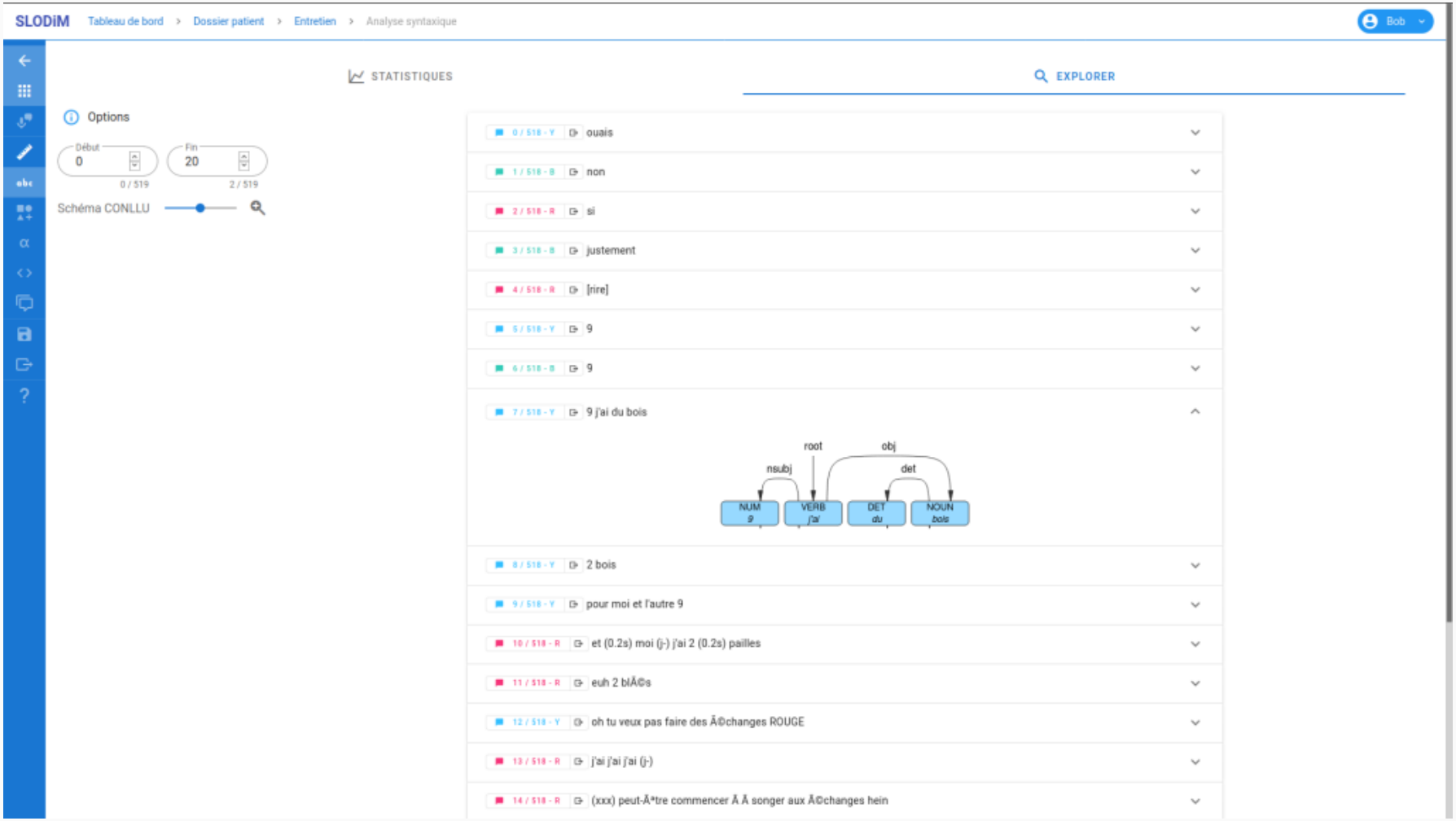
SLODiM
Tableau de bord
Dossier patient
Entretien
Analyse syntaxique
Bob
Participant Y
Distribution longueur des chaines
Participant B
Distribution longueur des chaines
Participant R
Distribution longueur des chaines
Participant O
Distribution longueur des chaines
Distribution des verbal heads
Distribution des verbal heads
Distribution des verbal heads
Distribution des verbal heads
Déictiques
Déictiques
Déictiques
Déictiques
Soi
Autre
Spatiotemporel
Désignation
Quantité
Longueur

SLODiM

ODiM.

CRÉER UN COMPTE

SLODiM


SLODiM
Tableau de bord
Dossier patient
Entretien
Analyse syntaxique
Bob
STATISTIQUES
EXPLORER
Options
Début
0
0 / 519
Fin
20
2 / 519
Schéma CONLLU
0 / 518 - Y ouais
1 / 518 - B non
2 / 518 - R si
3 / 518 - B justement
4 / 518 - R [rire]
5 / 518 - Y 9
6 / 518 - B 9
7 / 518 - Y 9 j'ai du bois
root
obj
nsubj
det
NUM
9
VERB
j'ai
DET
du
NOUN
bois
8 / 518 - Y 2 bois
9 / 518 - Y pour moi et l'autre 9
10 / 518 - R et (0.2s) moi (j-) j'ai 2 (0.2s) pailles
11 / 518 - R euh 2 blÃ©s
12 / 518 - Y oh tu veux pas faire des Ã©changes ROUGE
13 / 518 - R j'ai j'ai j'ai (j-)
14 / 518 - R (xxx) peut-Ãªtre commencer Ã songer aux Ã©changes hein

SLODiM

SLODiM Tableau de bord > Dossier patient

SLAM

ADRESSE

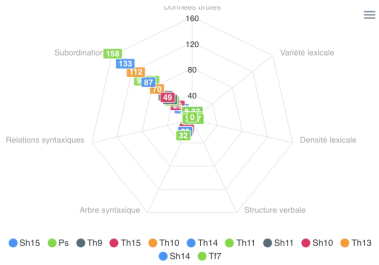
slam@slam.fr

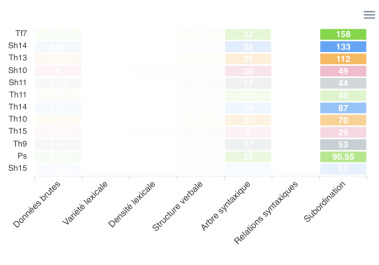
RAPPORT GLOBAL | ENTRETIENS | TESTS NEUROCOGNITIFS

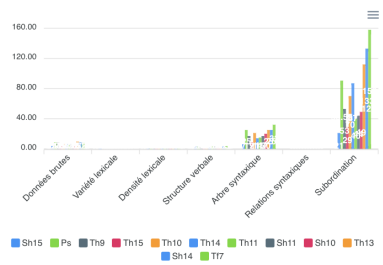
Affichage

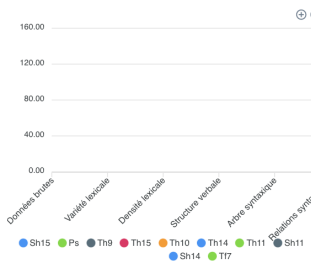
Résumé | Mesures globles | Historique

Résumé







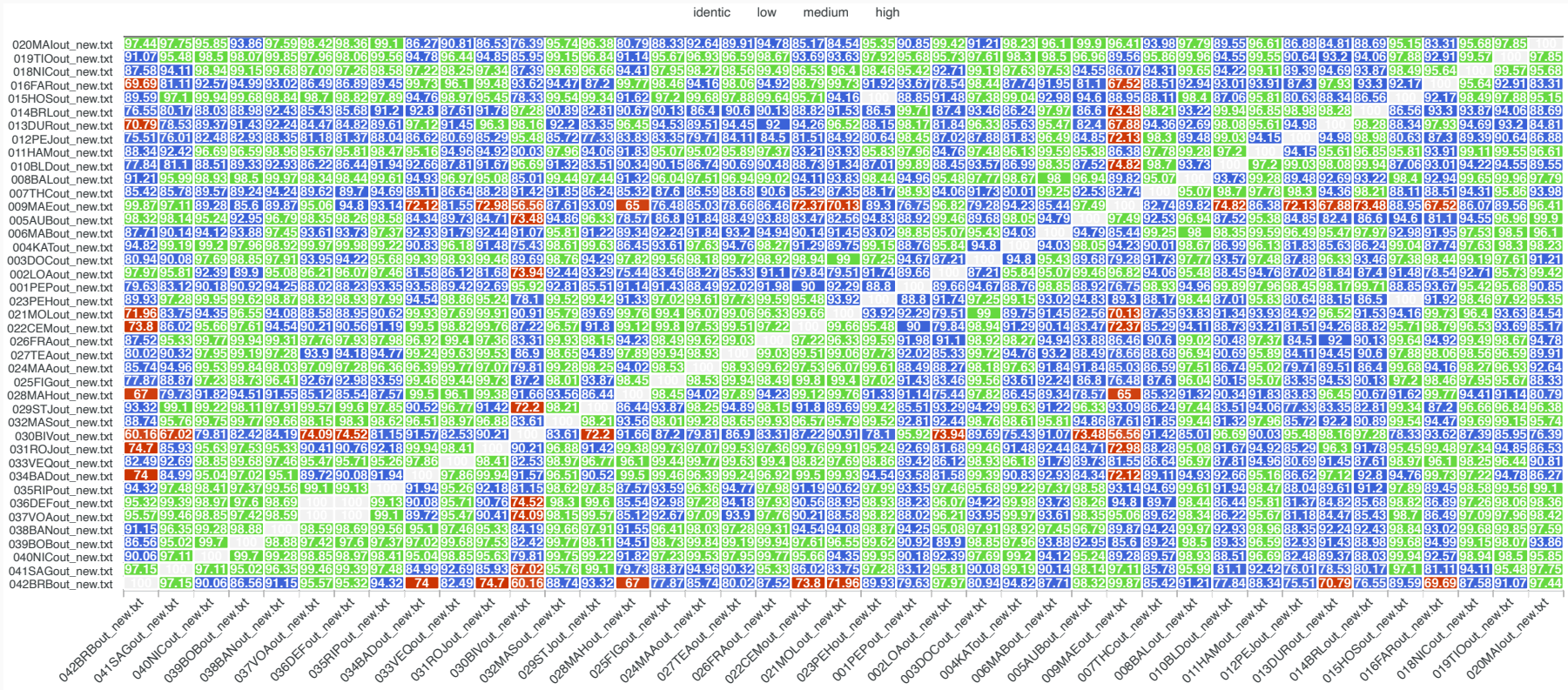


Mesures globales

Search

Donnée	Moyenne	Total
designation	0.01	0.17
spatiotemporal	0.01	0.16
formsTokensRatio	0.02	0.54
other	0.02	0.36
lemmasTokensRatio	0.03	0.6
self	0.04	0.79
averageQuestions	0.14	3.11
stopwordsRatio	0.33	7.26
lexicalDensity	0.45	9.87
arity	3.07	67.51
averageLength	4.24	93.32
flow	8.5	187
average	12.04	264.82
longestDependencyString	21.68	477
questions	32.77	721
subordinates	81.45	1792
verbalRootsRatio	92.41	2033
turns	263.09	5788
words	3193.45	70256

SLODiM


identic
low
medium
high
020MAIout_new.txt
019TIOout_new.txt
018NICout_new.txt
016FARout_new.txt
015HOSout_new.txt
014BRLout_new.txt
013DURout_new.txt
012PEJout_new.txt
011HAMout_new.txt
010BLDout_new.txt
008BALout_new.txt
007THCout_new.txt
009MAEout_new.txt
005AUBout_new.txt
006MABout_new.txt
004KATout_new.txt
003DOCout_new.txt
002LOAout_new.txt
001PEPout_new.txt
023PEHout_new.txt
021MOLout_new.txt
022CEMout_new.txt
026FRAout_new.txt
027TEAout_new.txt
024MAAout_new.txt
025FIGout_new.txt
028MAHout_new.txt
029STJout_new.txt
032MASout_new.txt
030BIVout_new.txt
031ROJout_new.txt
033VEQout_new.txt
034BADout_new.txt
035RIPout_new.txt
036DEFout_new.txt
037VOAout_new.txt
038BANout_new.txt
039BOBout_new.txt
040NICout_new.txt
041SAGout_new.txt
042BRBout_new.txt

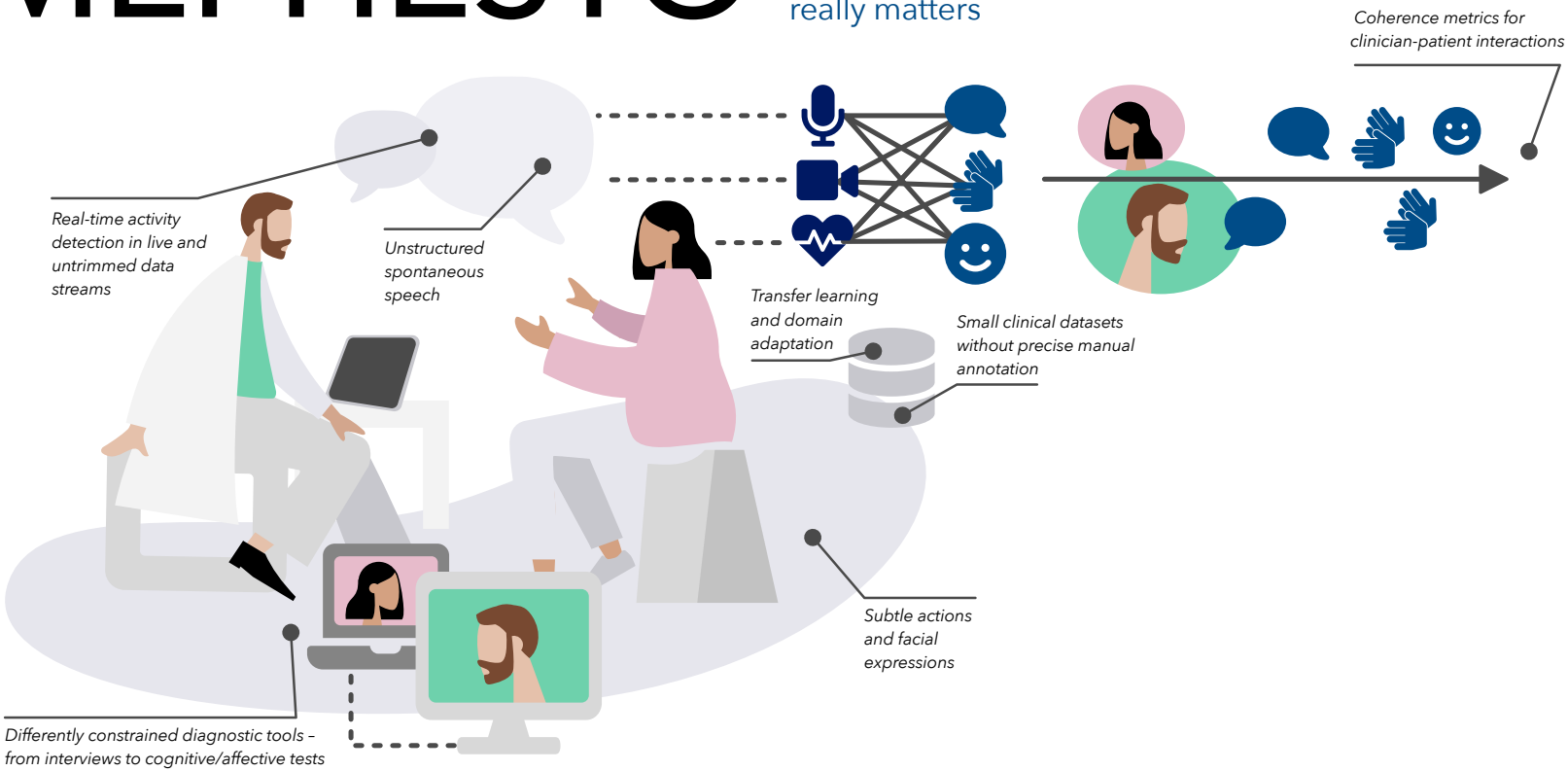
MEPHESTO

"The vision of MEPHESTO is to help breaking the scientific ground for next generation precision psychiatry through AI-based social interaction analysis."

Johannes Tröger, DFKI

MEPHESTO

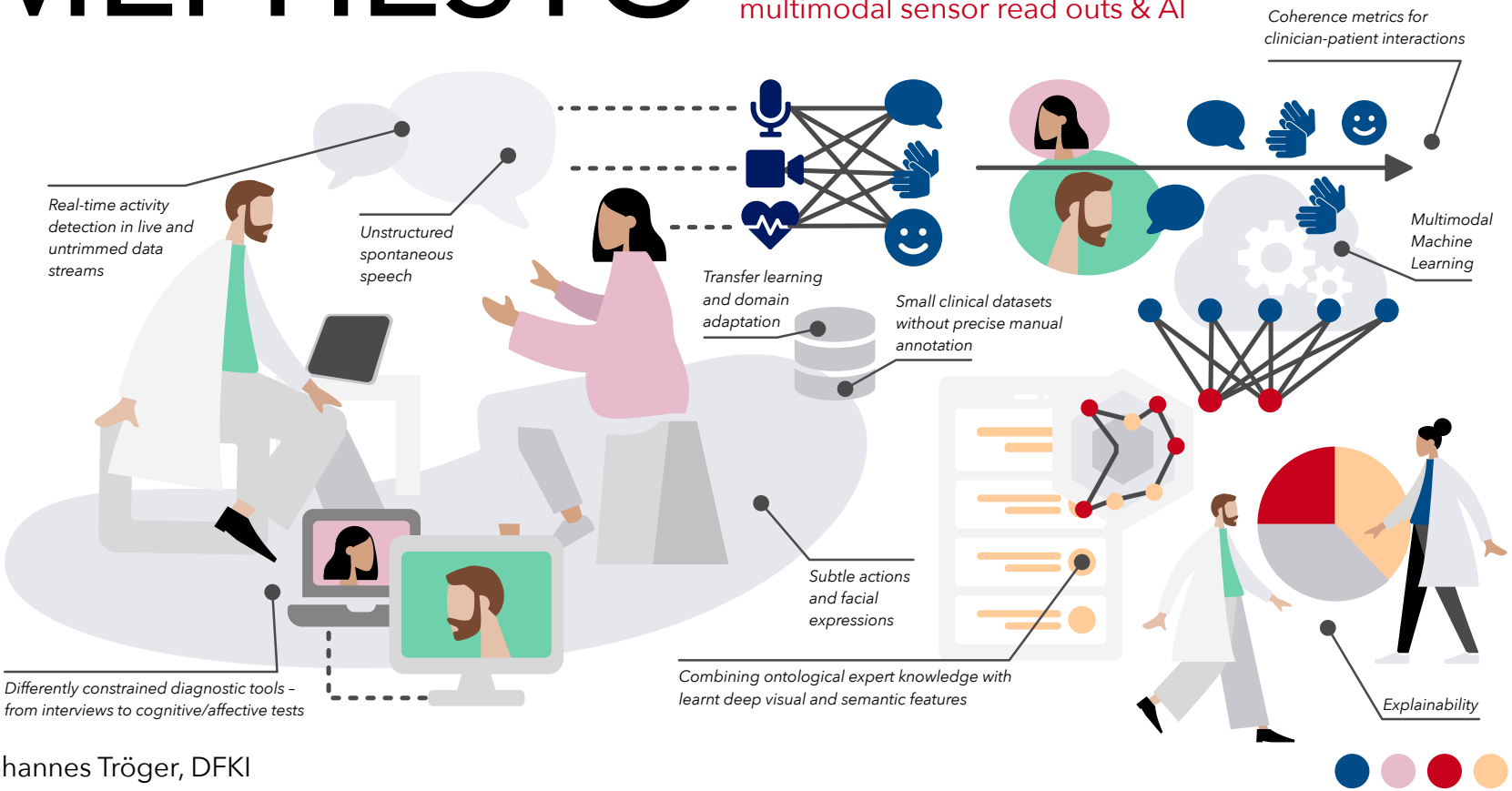
Real-world social interactions | Measuring clinical evidence where it really matters



Johannes Tröger, DFKI

MEPHESTO

Digital phenotyping |
Modelling clinical symptoms through multimodal sensor read outs & AI



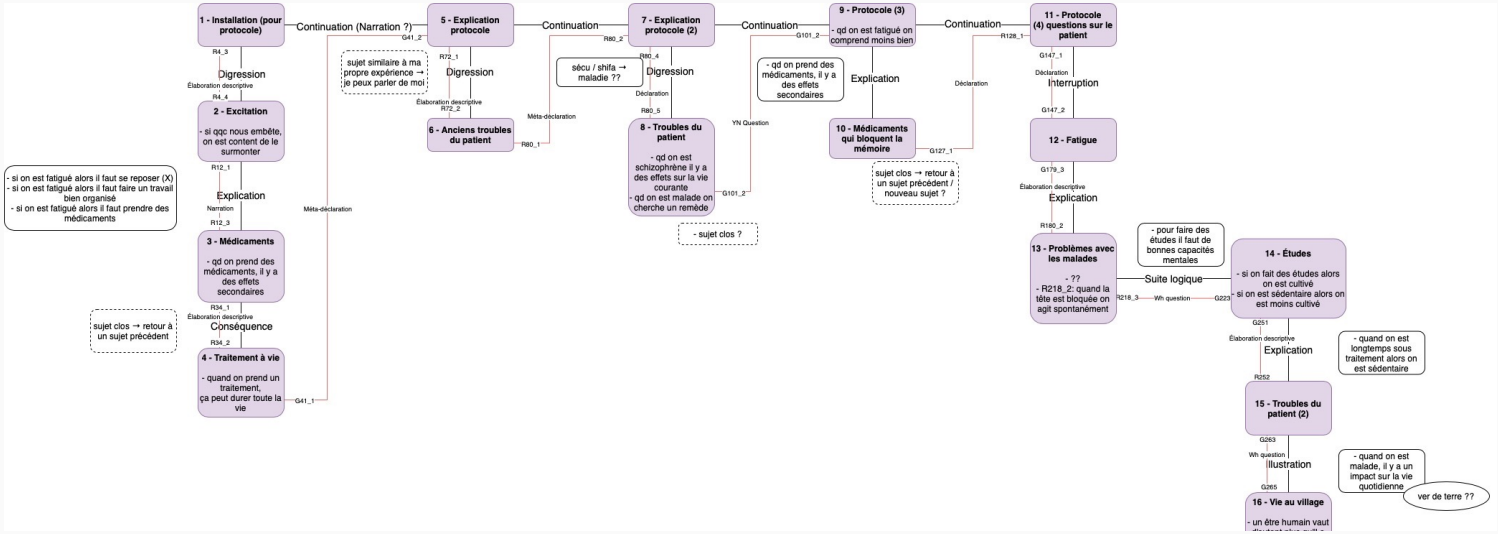
Johannes Tröger, DFKI

Issues

- Treatment of the French language
- Tool for experts
  - Early detection,
  - Help with diagnosis,
  - Relapse prediction,
  - Help with therapy
- Definition of language profiles
- Implementation of tools for the identification of cognitive processes
- Collaborative cycle : data collection - models - production of implementations

French Swedish cooperation

- Works on transcripts of interviews with patients :
  - spontaneous interview (Nancy)
  - balloon task (Göteborg)
- Combining linguistic and meta-conversational aspects



- Use of both formal linguistics and AI tools
  - SDRT
  - topos

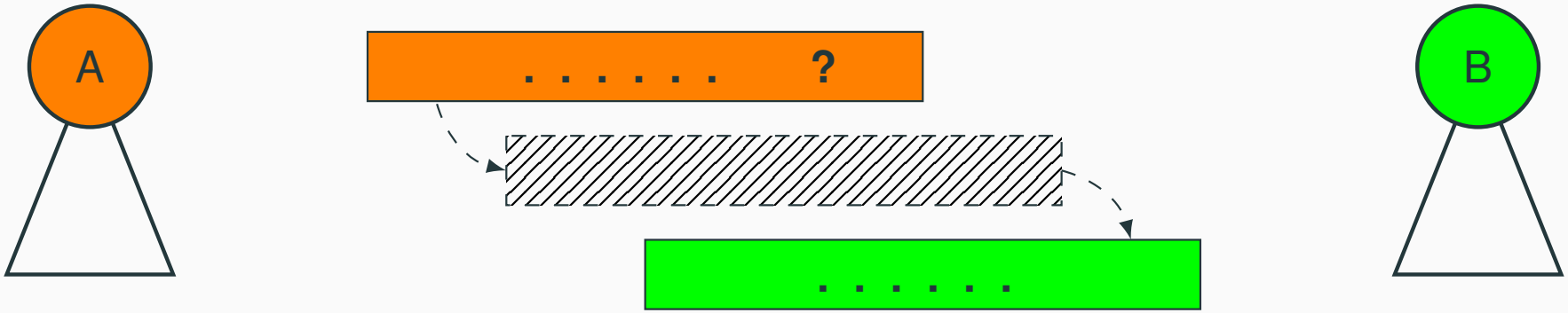
Thank you !

Un corpus difficile à constituer

[Amb. et al journée ATALA 2014]

- Démarches administratives lourdes :
  - CPP de la région de l'institution médicale
  - CNIL
- Les données ne doivent pas être utilisées pour/contre le patient.
- Participation des patients (pertes importantes de participation)
- Protocole lourd

Studying dialogue



- What is a question ?
- What is an answer ?
- Where is the question ?
- Where is the answer ?

- Data is scarce (oral ? French ? non constrained ?)
- Ethics : open access ? anonymity ?

Calculs automatiques

```
In Logical_form:
   CLOS (AGENT (SOME STUDENT) (TIME MIDNIGHT
    (LOCATION (THE PARK) (PATIENT (EVERY TEACHER)
MEET)))) : s

Interpreted by semantics in Semantics as:
   !
    (Ex x.
        (student x) &
        (All x'.
            (teacher x') >
            (! (Exv e. (meet e) & ((patient e x') &
             ((location e (the park)) & ((time e midnight) &
             ((agent e x) & T))))))
        )
    )
    : t
```

Questions

Forms (main, mutually exclusive)	Secondary forms	Functions
Yes/no (polar questions)	Disjunctive (alternative)	Ask (confirmation, feature, performance)
Wh	Interro-neg	
Dialogue management Auxiliary (deontic/epistemic)	Completion suggestion	BC - simple w/ questions
Assertion	Reported/quoted speech	• Phatic
Question tags	Complexity (ID)	• Clarification request
	Focus	• Turn

Réponses

Forms (main, mutually exclusive)	Secondary forms	Functions
Simple answer	Reported speech	Refuse, accept
Strongly exhaustive answer	Yes/no	• Deny the assumption (presupposition)
Question/question	Non Sentential Utterances	BC
Potentially resolving answer		• Phatic
Aboutness	Deflective	Give
		• Uncertain
		• Unknown
		• Unrelated
		• Feature
		Perform