
# defying complexity (lessons learned)

Karën Fort & Bruno Guillaume

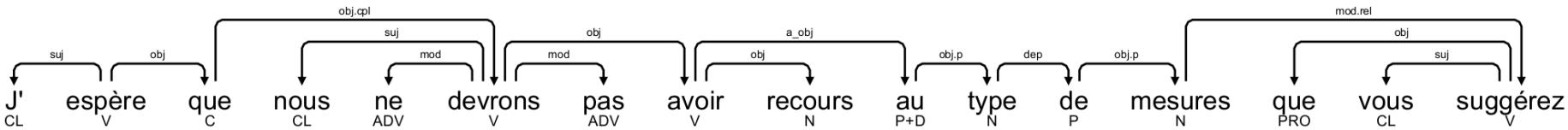
karen.fort@paris-sorbonne.fr / bruno.guillaume@loria.fr

October, 2016


1. Overview of the game
2. Motivating players
3. Behind the curtain
4. Obtained results [Guillaume et al., 2016]
5. Conclusion and future plans

1. Overview of the game
   - Dependency syntax annotation
   - ZombiLingo
2. Motivating players
3. Behind the curtain
4. Obtained results [Guillaume et al., 2016]
5. Conclusion and future plans

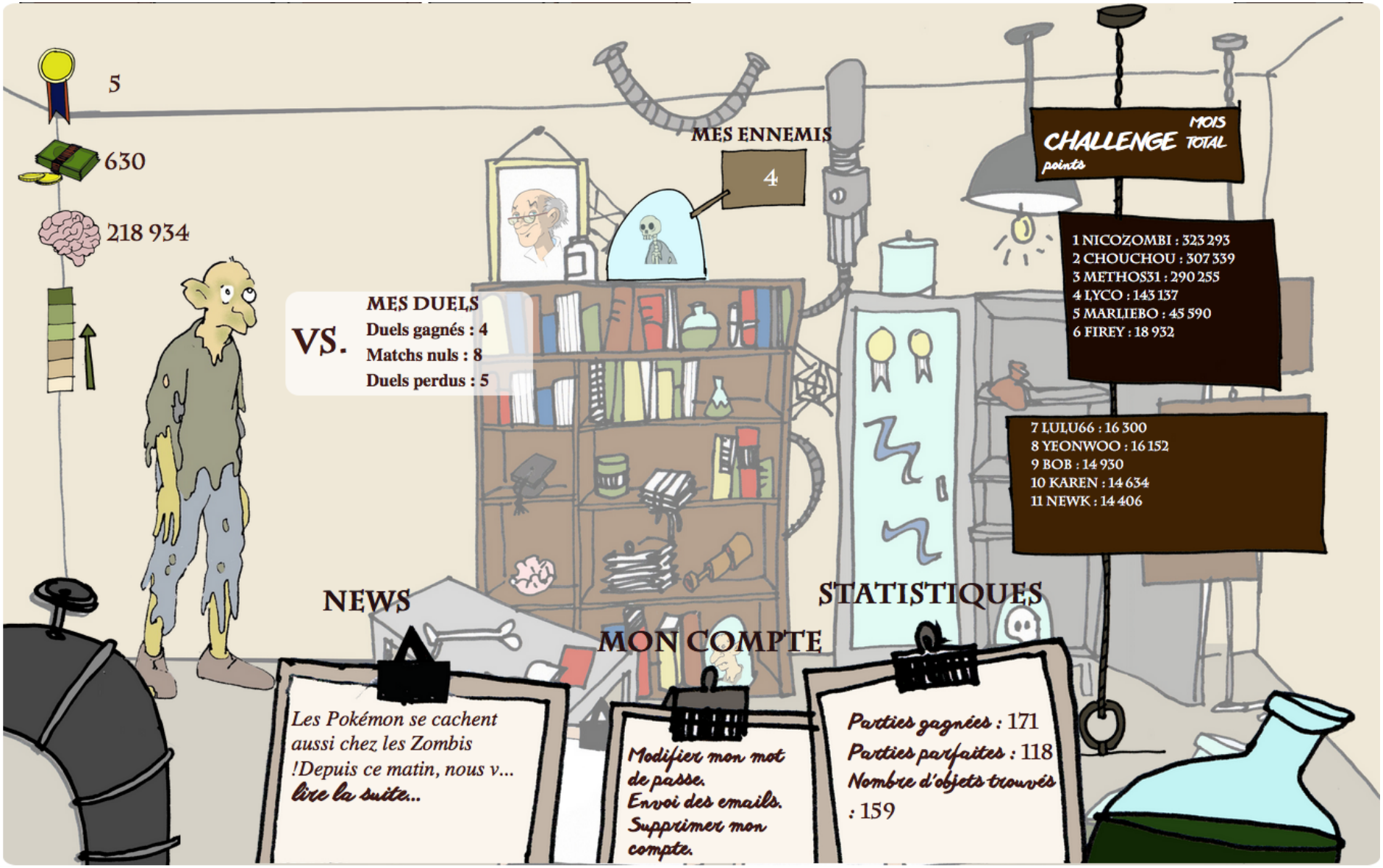
# A complex annotation type



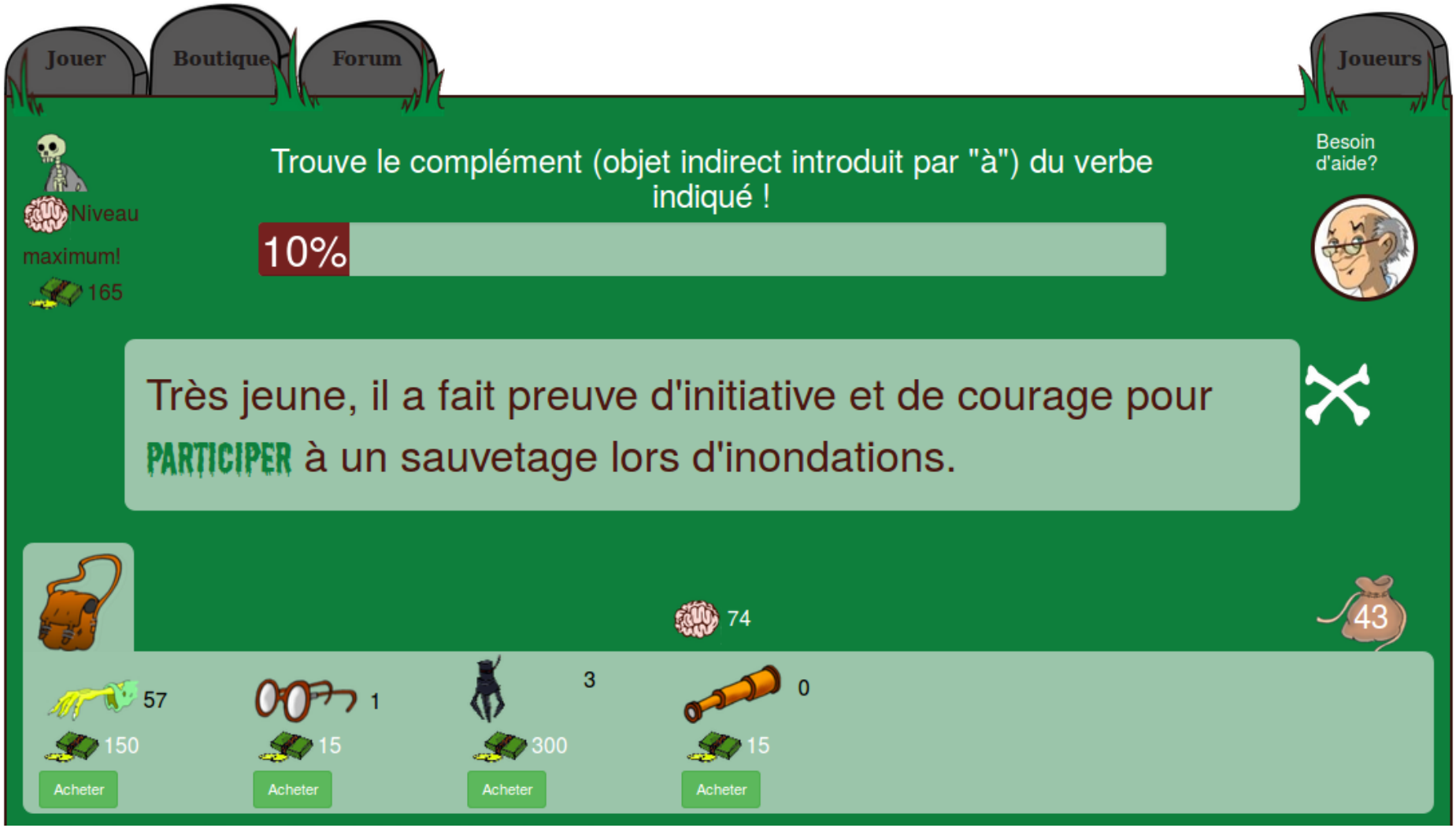

- annotation guidelines:
  - 29 relation types
  - approx. 50 pages
- counter-intuitive decisions

→ decompose the complexity of the task [Fort et al., 2012], not simplify it!

# http://zombilingo.org/


5
630
218 934
MES ENNEMIS
4
CHALLENGE
MOIS
TOTAL
points
1 NICOZOMBI : 323 293
2 CHOUCHOU : 307 339
3 METHOS31 : 290 255
4 LYCO : 143 137
5 MARLIEBO : 45 590
6 FIREY : 18 932
7 LULU66 : 16 300
8 YEONWOO : 16 152
9 BOB : 14 930
10 KAREN : 14 634
11 NEWK : 14 406
VS.
MES DUELS
Duels gagnés : 4
Matchs nuls : 8
Duels perdus : 5
NEWS
MON COMPTE
STATISTIQUES
Les Pokémon se cachent aussi chez les Zombis !Depuis ce matin, nous v... lire la suite...
Modifier mon mot de passe.
Envoi des emails.
Supprimer mon compte.
Parties gagnées : 171
Parties parfaites : 118
Nombre d'objets trouvés : 159


AUXILIAIRE CAUSATIF
Points : 270
Annotations : 6/6
AUXILIAIRE DE TEMPS
Points : 7 986
Annotations : 34/34
CHOIX DU CORPUS
POKEMON
NIVEAU 1
SUJET
JEU
FORMATION
Points : 4 980
Annotations : 22/282
NIVEAU 1
AFFIXE
JEU
FORMATION
Points : 0
Annotations : 0/108
NIVEAU 1
DÉTERMINANT
JEU
FORMATION
Points : 0
Annotations : 0/148
NIVEAU 2
COMPLÉMENT EN « DE »
JEU
FORMATION
Points : 0
Annotations : 0/120
NIVEAU 2
COMPLÉMENT EN « À »
JEU
FORMATION
Points : 0
Annotations : 9/187
NIVEAU 2
OBJET DE LA PRÉPOSITION
JEU
FORMATION
Points : 0
Annotations : 0/213
NIVEAU 2
AUXILIAIRE PASSIF
JEU
FORMATION
Points : 598
Annotations : 9/56


Jouer
Boutique
Forum
Joueurs
Niveau maximum!
165
Trouve le complément (objet indirect introduit par "à") du verbe indiqué !
10%
Besoin d'aide?
Très jeune, il a fait preuve d'initiative et de courage pour PARTICIPER à un sauvetage lors d'inondations.
74
43
57
150
Acheter
1
15
Acheter
3
300
Acheter
0
15
Acheter

1 Overview of the game

2 Motivating players
- Attracting players
- Keeping players playing

3 Behind the curtain

4 Obtained results [Guillaume et al., 2016]

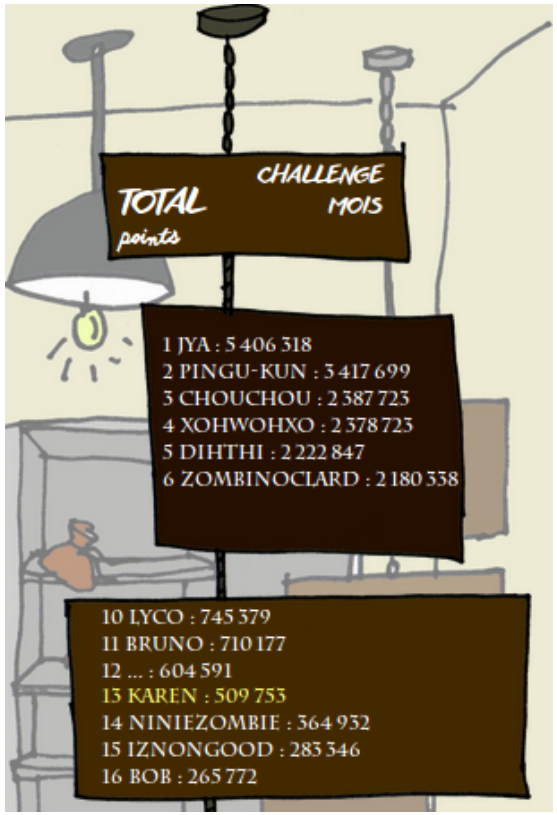
5 Conclusion and future plans

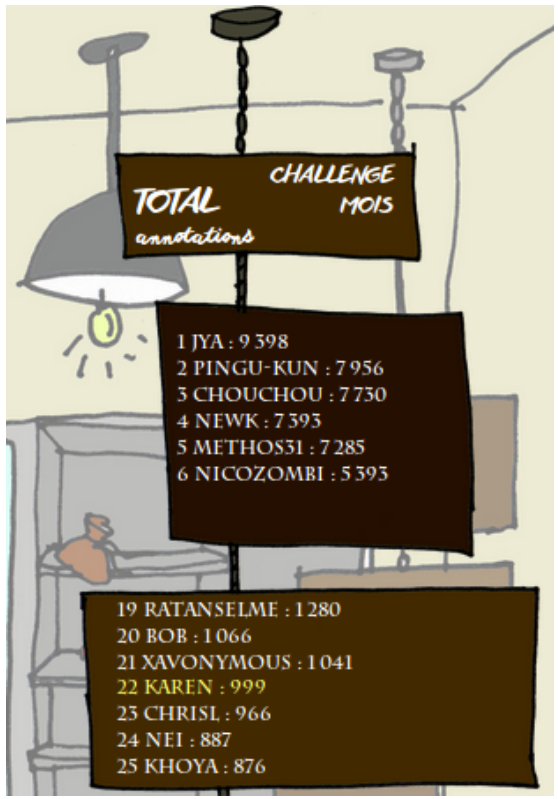
# General features

Bring the fun through:

- zombie design
- use of (crazy) objects
- regular challenges (specific corpus and design) on a trendy topic:
  - Star Wars (when the movie was playing)
  - soccer (during the Euro)
  - Pokemon (well...)

# LeaderboardS (for achievers)

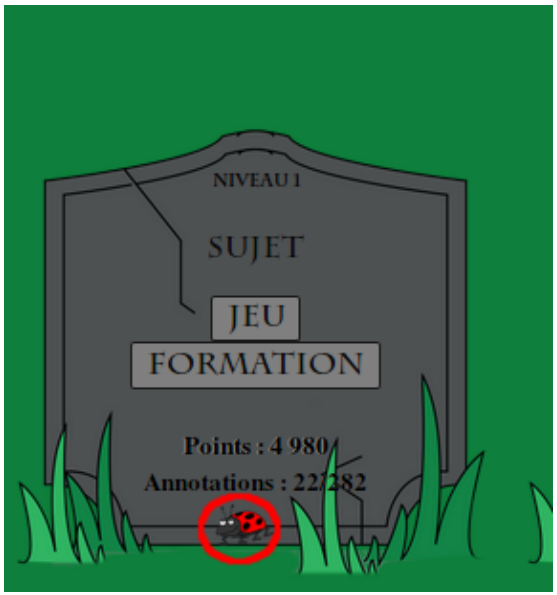



Criteria:

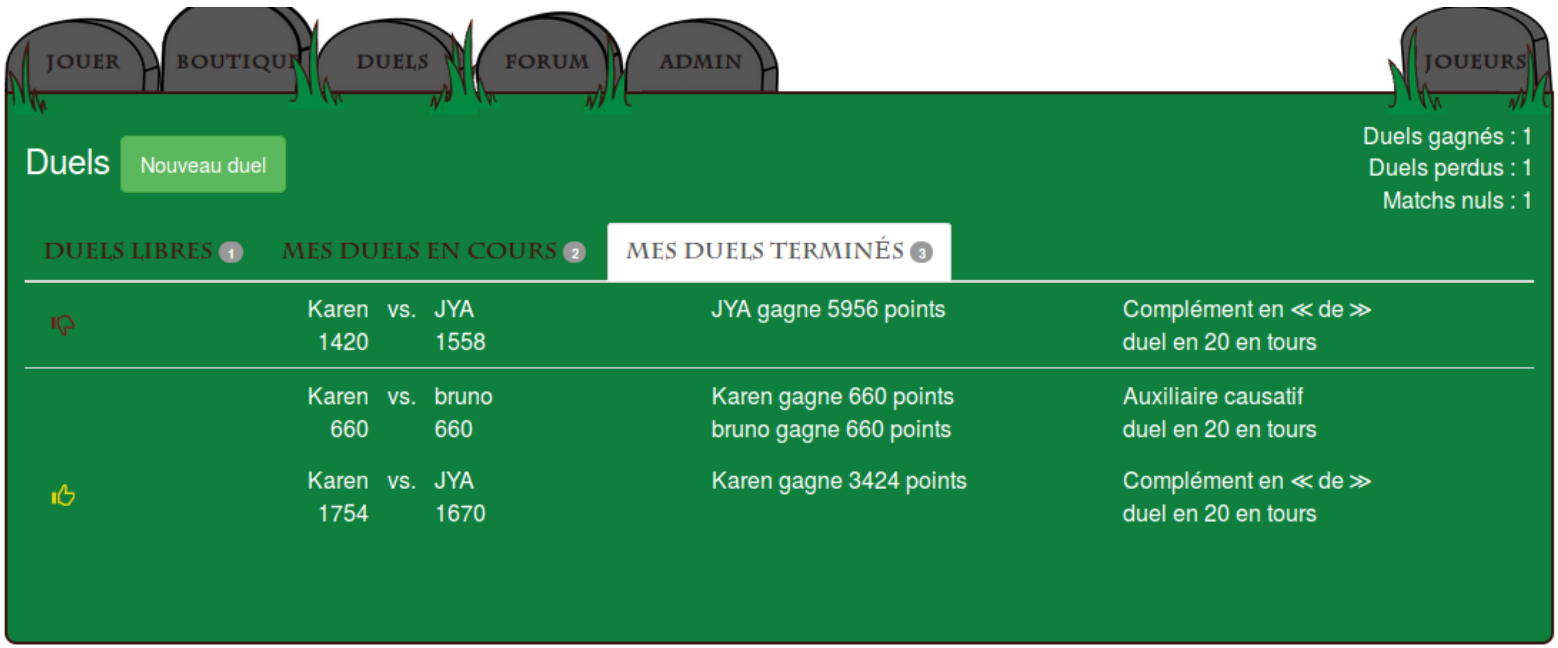
- number of annotations or points
- in total, during the month, during the challenge

# Hidden features (for explorers)



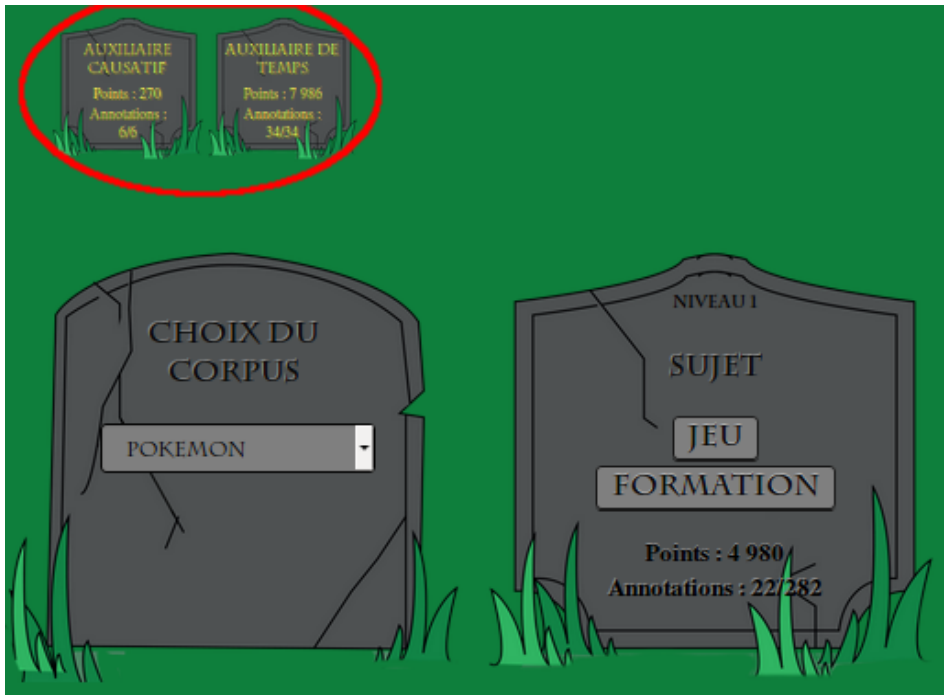

- appearing randomly
- with different effects: objects, other game, etc.

# Duels (for socializers (and killers?))




- select an enemy
- challenge them on a specific type of relation

# Badges (?) (for collectors)




- play all the sentences for a relation type, for a corpus
- play all the sentences from a corpus

1 Overview of the game

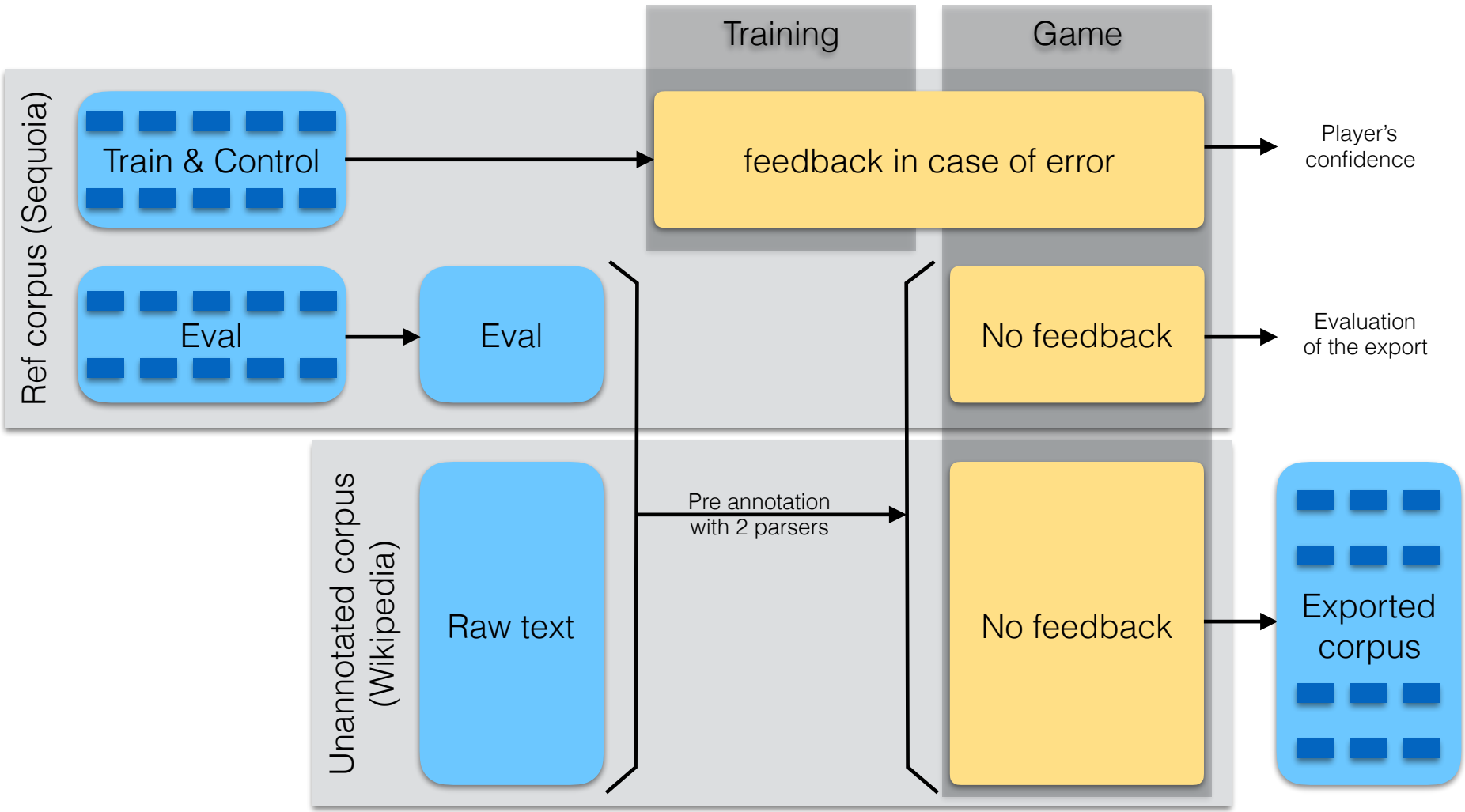
2 Motivating players

3 Behind the curtain
- Overview
- Preprocessing
- Ensuring quality

4 Obtained results [Guillaume et al., 2016]

5 Conclusion and future plans

# Organizing quality assurance



# Preprocessing data (freely available corpora)

Pre-annotation with two parsers:

1. a statistical parser : Talismane [Urieli, 2013]
2. a symbolic parser, based on graph rewriting : FrDep-Parse [Guillaume and Perrier, 2015]

→ play the items for which the two parsers give different annotations

# Training, control and evaluation

Reference: 3,099 sentences of the Sequoia corpus [Candito and Seddah, 2012]

| $\text{REF}_{Train\&Control}$ | $\text{REF}_{Eval}$ | Unused |
|---|---|---|
| 50% | 25% | 25% |
| 1,549 sentences | 776 sentences | 774 sentences |

- $\text{REF}_{Train\&Control}$ is used to train the players
- $\text{REF}_{Eval}$ is used like a raw corpus, to evaluate the produced annotations

# Training the players

## Compulsory for each dependency relation

- sentences are taken from the REF $_{Train\&Control}$ corpus
- a feedback is given in case of error

# Dealing with cognitive fatigue and long-term players

## Control mechanism

Sentences from the REF$_{Train\&Control}$ corpus are proposed regularly:

- if the player fails to find the right answer, a feedback with the solution is given
- after a given number of failures on the same relation, the player cannot play anymore and has to redo the corresponding training

$\rightarrow$ we deduce a level of confidence for the player on this relation

1 Overview of the game

2 Motivating players

3 Behind the curtain

4 Obtained results [Guillaume et al., 2016]
- Quantity
- Quality
- Density

5 Conclusion and future plans

# Production: game corpus size

## compared to other existing French dependency syntax corpora

As of July 10, 2016:

- 647 players
- who produced 107,719 annotations

| | Sequoia 7.0 | UD-French 1.3 | FTB-UC | FTB-SPMRL | Game |
|---|---|---|---|---|---|
| Sentences | 3,099 | 16,448 | 12,351 | 18,535 | 5,221 |
| Tokens | 67,038 | 401,960 | 350,947 | 557,149 | **128,046** |
| Tokens/sent. | 21.6 | 24.4 | 28.4 | 30.1 | 24.5 |

# Production: game corpus size

## compared to other existing French dependency syntax corpora

As of July 10, 2016:

- 647 players
- who produced 107,719 annotations

| | Sequoia 7.0 *free* | UD-French 1.3 *free* | FTB-UC *not free* | FTB-SPMRL *not free* | Game *free* |
|---|---|---|---|---|---|
| Sentences | 3,099 | 16,448 | 12,351 | 18,535 | 5,221 |
| Tokens | 67,038 | 401,960 | 350,947 | 557,149 | **128,046** |
| Tokens/sent. | 21.6 | 24.4 | 28.4 | 30.1 | 24.5 |

# Production: game corpus size

## compared to other existing French dependency syntax corpora

As of July 10, 2016:

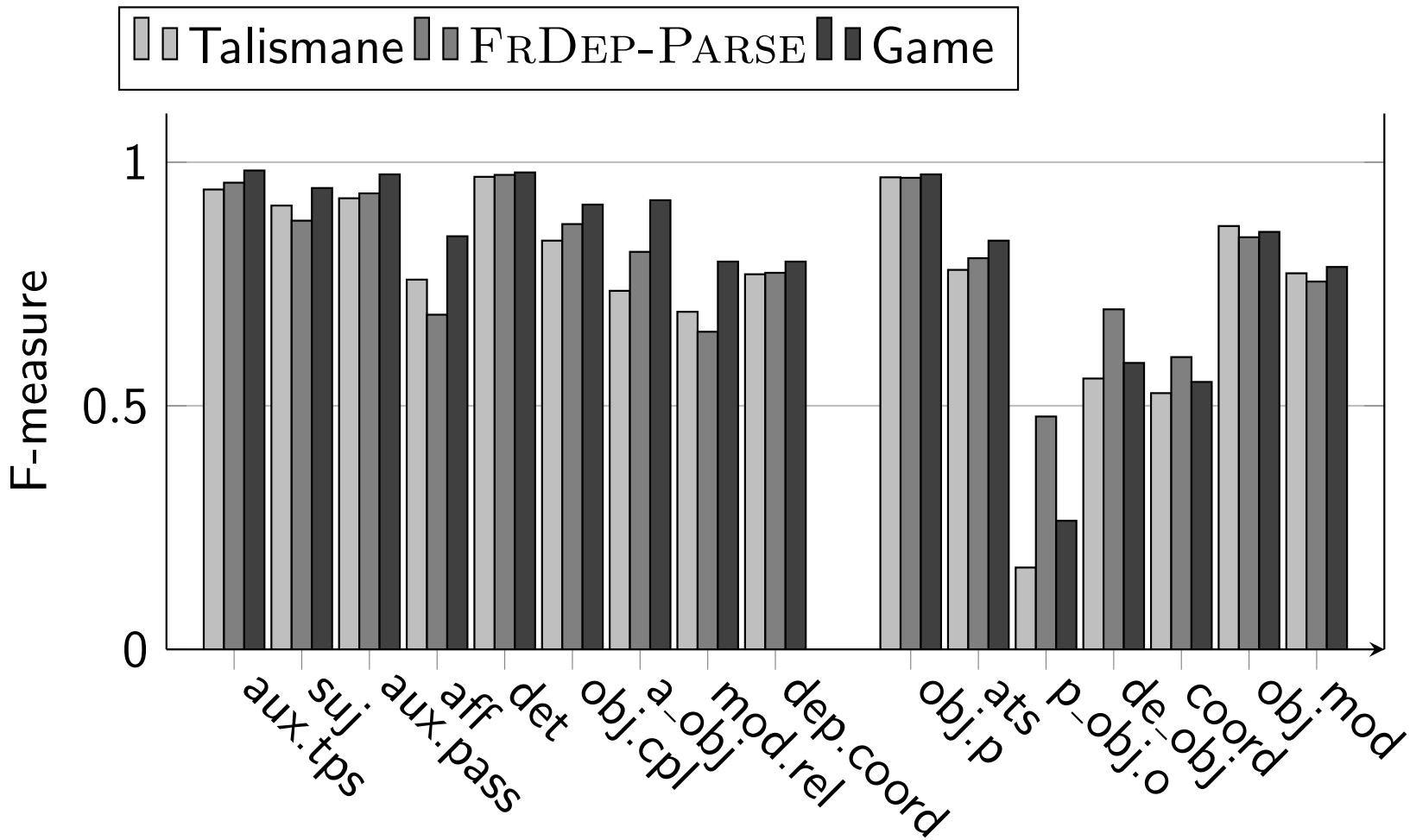
- 647 players
- who produced 107,719 annotations

| | Sequoia 7.0 *free validated* | UD-French 1.3 *free errors* | FTB-UC *not free validated* | FTB-SPMRL *not free validated* | Game *free validated* |
|---|---|---|---|---|---|
| Sentences | 3,099 | 16,448 | 12,351 | 18,535 | 5,221 |
| Tokens | 67,038 | 401,960 | 350,947 | 557,149 | **128,046** |
| Tokens/sent. | 21.6 | 24.4 | 28.4 | 30.1 | 24.5 |

+ (ever)growing resource!

# Evaluating quality

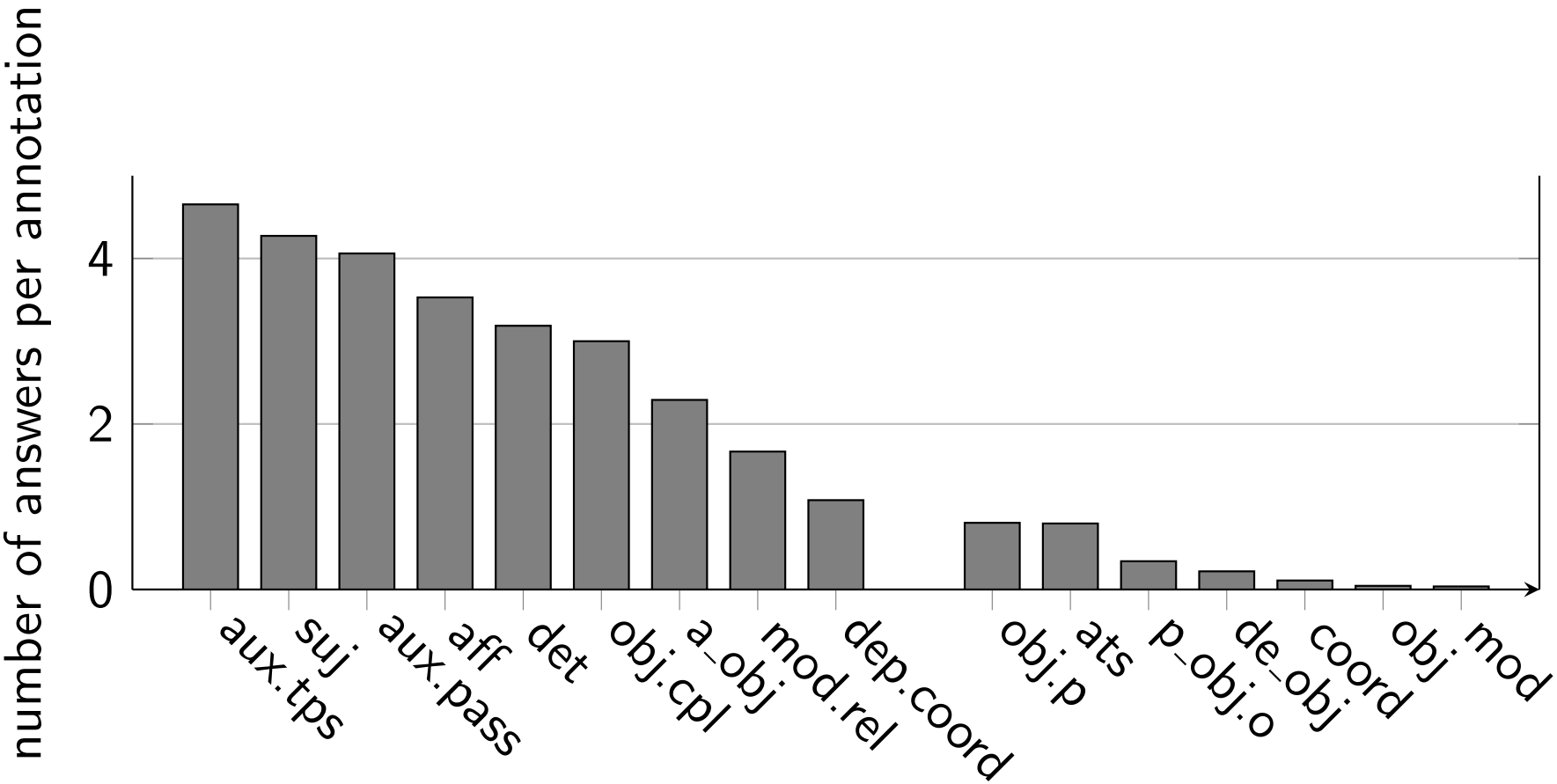
## on the $\text{REF}_{Eval}$ corpus




NB: left part of the figure = density of annotation $> 1$

# Annotation density

## on the $\text{REF}_{Eval}$ corpus




→ need **more** annotations on some relations

1 Overview of the game

2 Motivating players

3 Behind the curtain

4 Obtained results [Guillaume et al., 2016]

5 Conclusion and future plans

# Improving gamification

Give more to:

- explore and collect
- build a real story
- build a sense of community

# Improving the exported resource

Test the influence of:

- the pre-annotation score
- the level of the player in the game
- the confidence we have in the player for the relation type at hand

# Expand to new languages

## and new annotation types

New languages:

- English
- **less-resourced languages**

New annotation types:

- **POS**,
- corpus gathering, etc.

**Alice Millour** (PhD student)

# Building a Community

GWAPs for research should form a network, to:

- attract more players,
- share them,
- share the burden of communication

# Thanks!

Nicolas Lefèbvre (engineer)


Candito, M. and Seddah, D. (2012).
Le corpus Sequoia : annotation syntaxique et exploitation pour l'adaptation d'analyseur par pont lexical.
In Proceedings of the Traitement Automatique des Langues Naturelles (TALN), Grenoble, France.


Fort, K., Nazarenko, A., and Rosset, S. (2012).
Modeling the complexity of manual annotation tasks: a grid of analysis.
In International Conference on Computational Linguistics (COLING), pages 895–910, Mumbai, India.


Guillaume, B., Fort, K., and Lefebvre, N. (2016).
Crowdsourcing complex language resources: Playing to annotate dependency syntax.
In Proceedings of the International Conference on Computational Linguistics (COLING), Osaka, Japan.


Guillaume, B. and Perrier, G. (2015).

Dependency Parsing with Graph Rewriting.
In
Proceedings of IWPT 2015, 14th International Conference on Parsing Te
pages 30–39, Bilbao, Spain.

Urieli, A. (2013).
Robust French syntax analysis: reconciling statistical methods and linguistic knowledge in the Talismane toolkit.
PhD thesis, Université de Toulouse II le Mirail, France.