

*Ce qui n'est pas compté ne compte pas.* [cité-e par A. Desrosières]

## Présence et représentation des femmes dans le traitement automatique des langues en France

Karën Fort et Aurélie Névéol



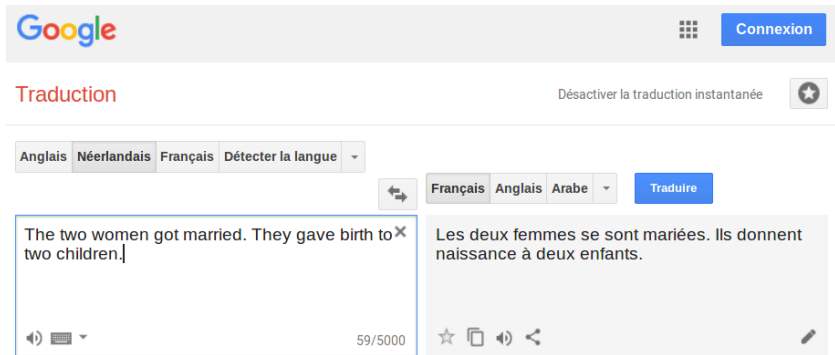
## Le Traitement Automatique des Langues : un domaine hybride

Représentation des femmes dans le TAL en France

Compter les femmes dans le TAL en France

Conclusions et propositions

# Qu'est-ce que le traitement automatique des langues ?



The screenshot shows the Google Translate web interface. At the top left is the Google logo. To the right is a grid icon and a blue button labeled "Connexion". Below the logo is the word "Traduction" in red. To the right of "Traduction" is the text "Désactiver la traduction instantanée" and a star icon in a square. Below this is a horizontal line. Under the line, there are two language selection menus. The first menu has buttons for "Anglais", "Néerlandais", "Français", and "Détecter la langue" with a dropdown arrow. The second menu has buttons for "Français", "Anglais", and "Arabe" with a dropdown arrow. To the right of the second menu is a blue button labeled "Traduire". Below the language menus is a text input box containing the English text: "The two women got married. They gave birth to two children." with a cursor at the end. Below the input box are a speaker icon, a list icon, and a dropdown arrow, and the character count "59/5000". To the right of the input box is a text output box containing the French translation: "Les deux femmes se sont mariées. Ils donnent naissance à deux enfants." Below the output box are a star icon, a copy icon, a speaker icon, a share icon, and a pencil icon.

Google

Connexion

Traduction

Désactiver la traduction instantanée

Anglais Néerlandais Français Détecter la langue

Français Anglais Arabe Traduire

The two women got married. They gave birth to two children.

59/5000

Les deux femmes se sont mariées. Ils donnent naissance à deux enfants.

# Qu'est-ce que le traitement automatique des langues ?

## Définition

*Le traitement automatique [...] des langues (abr. TAL) est une discipline à la frontière de la linguistique, de l'informatique et de l'intelligence artificielle, qui concerne l'application de programmes et techniques informatiques à tous les aspects du langage humain.*

[Wikipédia, consultée le 20 janvier 2018]

## Sections CNU (2015-2016) et CNRS (2014)

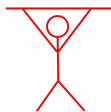
	Maîtresses de conf.	Professeures
CNU 7 (sces du langage)	69 %	50 %
CNU 27 (informatique)	26 %	19 %

[cache.media.enseignementsup-recherche.gouv.fr/file/statistiques/18/6/Section\\_27\\_768186.pdf](http://cache.media.enseignementsup-recherche.gouv.fr/file/statistiques/18/6/Section_27_768186.pdf)

	CR	DR
sect. 34 (sces du langage)	54,5 %	56,9 %
sect. 6 (informatique)	13,6 %	21,3 %
sect. 7 (informatique)	24,5 %	18,9 %

[http://bilansocial.dsi.cnrs.fr/BSP-2014\\_web.pdf](http://bilansocial.dsi.cnrs.fr/BSP-2014_web.pdf)

# Sections CNU (2015-2016) et CNRS (2014)



	Maîtresses de conf.	Professeures
CNU 7 (sces du langage)	69 %	50 %
CNU 27 (informatique)	26 %	19 %

[cache.media.enseignementsup-recherche.gouv.fr/file/statistiques/18/6/Section\\_27\\_768186.pdf](http://cache.media.enseignementsup-recherche.gouv.fr/file/statistiques/18/6/Section_27_768186.pdf)

	CR	DR
sect. 34 (sces du langage)	54,5 %	56,9 % 0/2 DRCE
sect. 6 (informatique)	13,6 %	21,3 % 1/5 DRCE, 6/18 DR1
sect. 7 (informatique)	24,5 %	18,9 % 0/8 DRCE, 8/25 DR1

[http://bilansocial.dsi.cnrs.fr/BSP-2014\\_web.pdf](http://bilansocial.dsi.cnrs.fr/BSP-2014_web.pdf)

Le Traitement Automatique des Langues : un domaine hybride

Représentation des femmes dans le TAL en France

Compter les femmes dans le TAL en France

Conclusions et propositions

# Association savante : l'ATALA

Association pour le traitement automatique des langues

Deux instances de direction :

## 1. comité permanent (CPERM)

- ▶ 9 hommes et 7 femmes
- ▶ Président

Choix par [cooptation](#)

## 2. conseil d'administration (CA)

- ▶ 15 hommes et 5 femmes
- ▶ Président
- ▶ 2 présidentes depuis sa création en 1959

Choix par [élection](#)



# Association savante : l'ATALA

Association pour le traitement automatique des langues

Deux instances de direction :

## 1. comité permanent (CPERM)

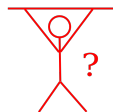
- ▶ 9 hommes et 7 femmes
- ▶ **Président**

Choix par **cooptation**

## 2. conseil d'administration (CA)

- ▶ 15 hommes et 5 femmes
- ▶ **Président**
- ▶ 2 présidentes depuis sa création en 1959 (12,5 %)

Choix par **élection**



# Principale revue nationale : TAL

auto-gérée et en accès libre, depuis 1960

1 numéro *varia* et 2 spéciaux par an

Comité de relecture (CR) :

- ▶ 33 membres (et un·e secrétaire)
- ▶ dont 10 femmes (23 %)

Numéros spéciaux :

- ▶ 30 rédacteurs en chefs
- ▶ 15 rédactrices en cheffes

Choix par [cooptation](#)

# Principale revue nationale : TAL

auto-gérée et en accès libre, depuis 1960

1 numéro *varia* et 2 spéciaux par an

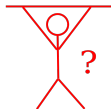
Comité de relecture (CR) :

- ▶ 33 membres (et un·e secrétaire)
- ▶ dont 10 femmes (23 %)

Numéros spéciaux :

- ▶ 30 rédacteurs en chefs
- ▶ 15 rédactrices en cheffes (33 %)

Choix par [cooptation](#)



# Principale conférence nationale : TALN

depuis 1994

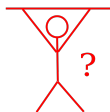
- ▶ Présidence du comité d'organisation :
  - ▶ 2 femmes sur 22
- ▶ Présidence de session :
  - ▶ 2014 : 3 femmes sur 12
  - ▶ 2016 : 5 femmes sur 13
  - ▶ 2017 : 5 femmes sur 14
- ▶ Conférencier·ère·s invité·e·s :
  - ▶ 8 femmes sur 29

Choix par [cooptation](#)

# Principale conférence nationale : TALN

depuis 1994

- ▶ Présidence du comité d'organisation :
  - ▶ 2 femmes sur 22 (8 %)
- ▶ Présidence de session :
  - ▶ 2014 : 3 femmes sur 12 (25 %)
  - ▶ 2016 : 5 femmes sur 13 (38,5 %)
  - ▶ 2017 : 5 femmes sur 14 (35,7 %)
- ▶ Conférencier·ère·s invité·e·s :
  - ▶ 8 femmes sur 29 (27,6 %)



Choix par [cooptation](#)

# Prix

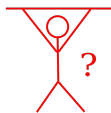
Depuis 2008

- ▶ Jeunes chercheur·se·s :
  - ▶ prix RECITAL : 6 hommes et 6 femmes
  - ▶ prix de thèse ATALA (depuis 2011) : 4 hommes et 3 femmes
  
- ▶ Chercheur·se·s confirmé·e·s :
  - ▶ prix TALN : 10 femmes et 25 hommes

Choix par [sélection](#)

# Prix

Depuis 2008



▶ Jeunes chercheur·se·s :

- ▶ prix RECITAL : 6 hommes et 6 femmes
- ▶ prix de thèse ATALA (depuis 2011) : 4 hommes et 3 femmes

▶ Chercheur·se·s confirmé·e·s :

- ▶ prix TALN : 10 femmes et 25 hommes (28,5 %)

Choix par [sélection](#)

Le Traitement Automatique des Langues : un domaine hybride

Représentation des femmes dans le TAL en France

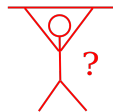
Compter les femmes dans le TAL en France

Conclusions et propositions



# Combien de femmes dans le TAL en France ?

Peut-on vraiment parler de plafond de verre ?



## Autrices d'articles dans les conférences vs ...

TALN :

- ▶ 24 % d'autrices en 1997
- ▶ 29 % d'autrices en 2014

NB : entre temps, la conférence est passée en double aveugle

LREC (international language resources and evaluation conf.) :

- ▶ 34 % d'autrices [Mariani et al., 2014]

NB : soumission non anonyme

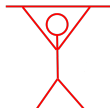
→ Biais en cascade ? [Wenneras and Wold, 1997]

## ... membres de l'ATALA

De 2003 à 2016 :

- ▶ 640 femmes
- ▶ 696 hommes
- ▶ 247 personnes au prénom mixte ou de genre inconnu

→ soit un taux de 47 % de femmes en excluant les inclassables



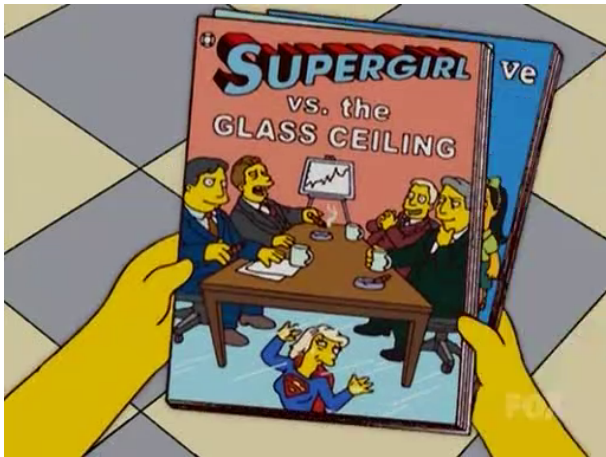
Le Traitement Automatique des Langues : un domaine hybride

Représentation des femmes dans le TAL en France

Compter les femmes dans le TAL en France

**Conclusions et propositions**

# Un plafond de verre effectif. . .



<http://worstsimpsonspageever.tumblr.com/post/59257276442>

... mais difficile à mesurer avec exactitude

Manque de données :

- ▶ historique du CPERM
- ▶ TALN :
  - ▶ comités d'organisation
  - ▶ présidences de session
  - ▶ articles refusés
- ▶ TAL :
  - ▶ articles refusés

## Des outils pour compter

- ▶ *crowdsourcing* pour compter les femmes et les hommes dans les amphithéâtres des conférences :  
<http://itcounts-app.org/>
- ▶ contacter les plateformes d'organisation de conférences (EasyChair, ScienceConf, ...) pour sensibiliser les organisateur·trice·s

Merci !







Mariani, J., Paroubek, P., Francopoulo, G., and Hamon, O. (2014).

Rediscovering 15 years of discoveries in language resources and evaluation : The LREC anthology analysis.

In Actes de International Conference on Language Resources and Evaluation, Reykjavik, Islande.



Wenneras, C. and Wold, A. (1997).

Nepotism and sexism in peer-review.

Nature, 387 :341–343.