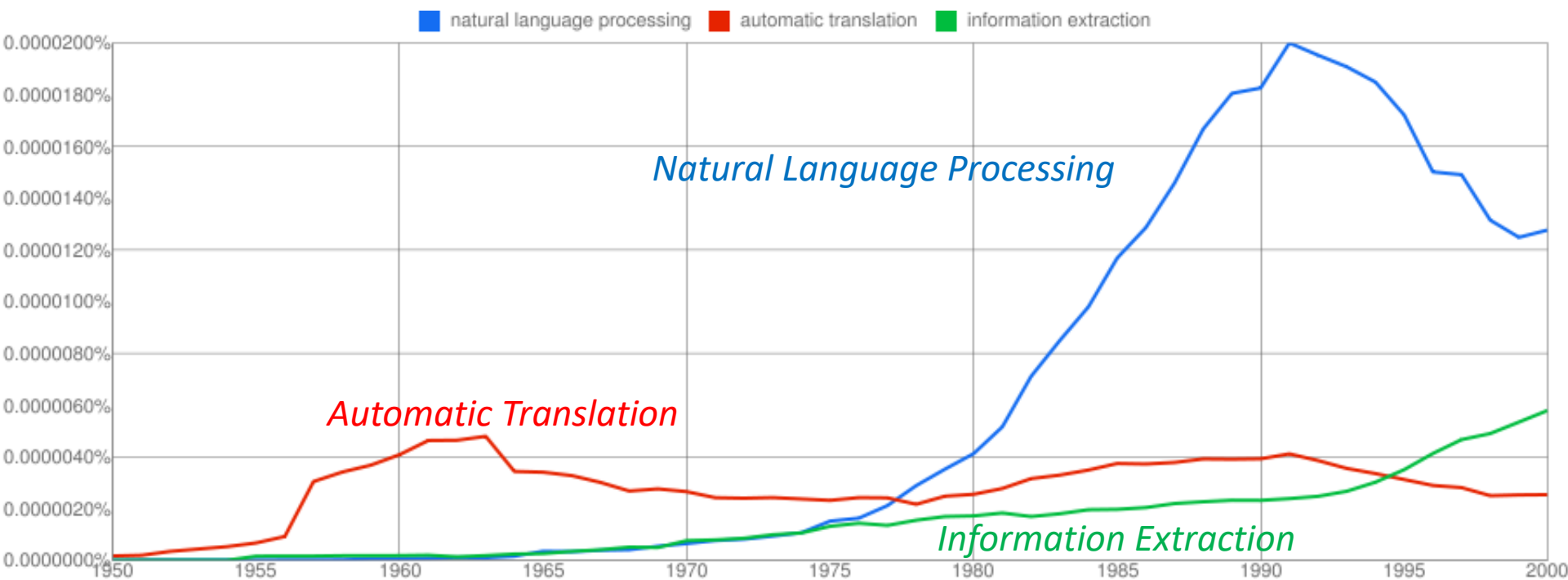


Extraction d'Information (Introduction)

Master DAC, Sorbonne Université

Xavier Tannier

xavier.tannier@sorbonne-universite.fr



Google labs

Books Ngram Viewer

Extraction d'information : définition

- Aux origines du TAL, une barre placée trop haut
 - Des modèles génériques de **compréhension des textes**
 - Des problèmes à la fois théoriques et pratiques
 - **Couverture insuffisante** des grammaires
 - **Ambiguïtés** non levées trop nombreuses
 - Difficultés pour collecter et manipuler les **connaissances sémantiques et pragmatiques** suffisantes
 - L'approche **générique** de compréhension est pour l'instant une **utopie**

Extraction d'information : définition

- L'**extraction d'information** :
 - ne cherche plus à comprendre les textes dans leur ensemble
 - vise à extraire d'un texte donné des éléments pertinents qui répondent à un besoin d'information
- Le type d'information pertinente pour une **application** donnée est en général **défini à l'avance** par le modèle

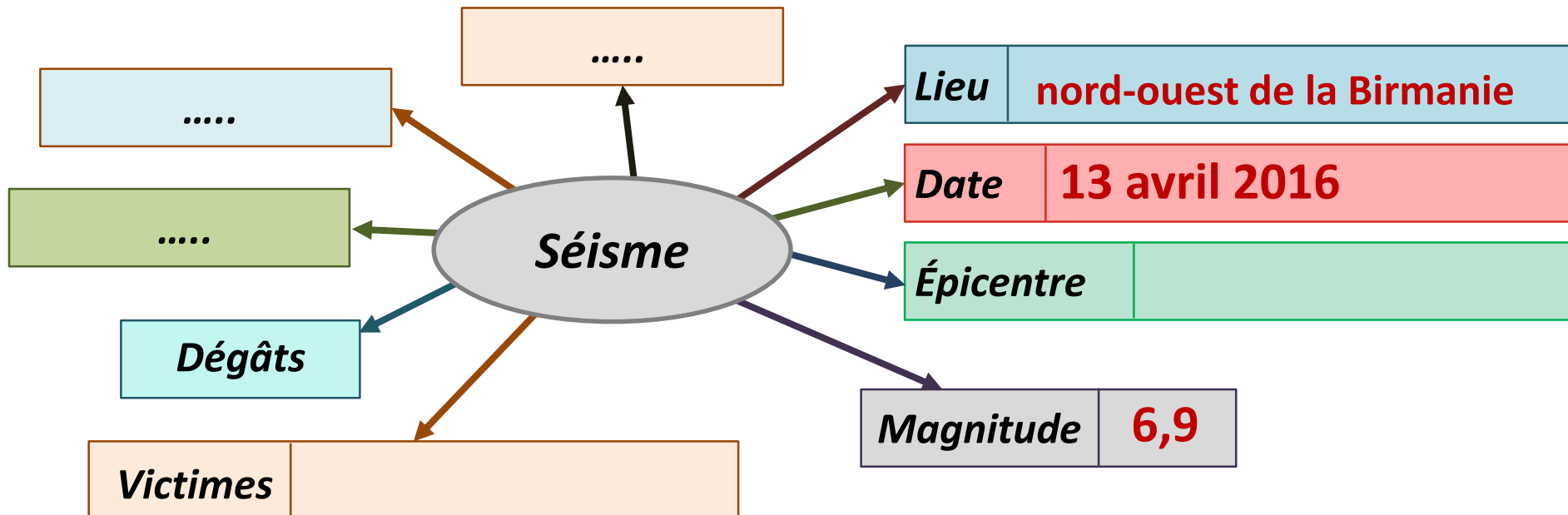
Extraction d'information : définition

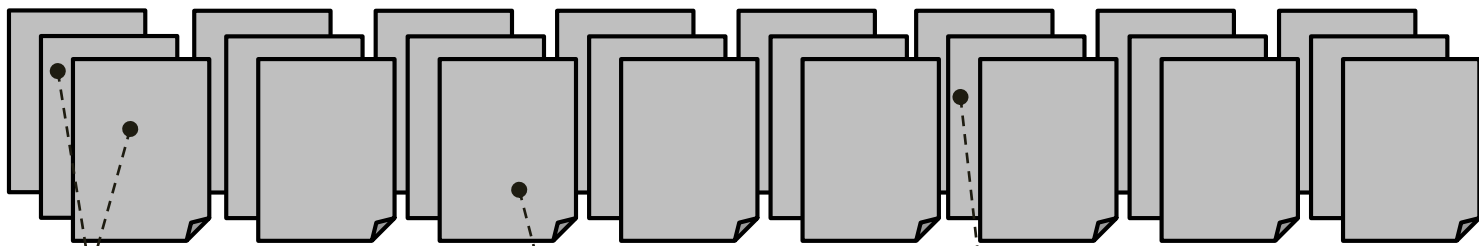
- "L'activité qui consiste à remplir automatiquement une banque de données à partir de textes écrits en langue naturelle" *(T. Poibeau)*
- Une approche **guidée par le but** :
 - Identifier les occurrences d'événements particuliers
 - En extraire les arguments impliqués
 - En donner une représentation structurée
- L'analyse s'effectue au **niveau local**
- Seule une partie du texte est considérée
(10 à 20 % de texte utile pour un tâche spécifique)

Exemple 1

Extraction d'entités, extraction de relations

Un séisme de magnitude 6,9 a secoué hier
le nord-ouest de la Birmanie.





.....

.....

Lieu	nord-ouest de la Birmanie
-------------	---------------------------

Date	13 avril 2016
-------------	---------------

Épicentre	
------------------	--

Magnitude	6,9
------------------	-----

Séisme

Dégâts

Victimes	
-----------------	--

.....

Exemple 2

Cadres :

MATCH
<i>équipe1 :</i> <i>équipe2 :</i> <i>vainqueur :</i> <i>score :</i>

BUTEURS
<i>équipe1 :</i> <i>équipe2 :</i>

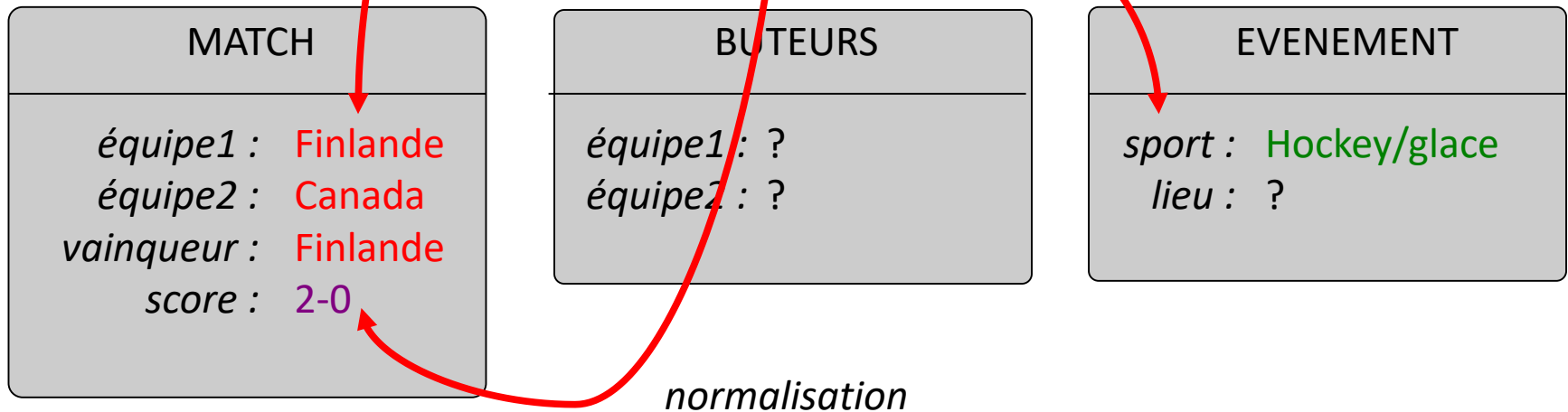
EVENEMENT
<i>sport :</i> <i>lieu :</i>

Exemple 2 (avec des patrons sémantiques)

Entrée :

La Finlande a battu le Canada 2 à 0 lors d'un match du premier tour du tournoi masculin de hockey sur glace.

Résultat :



Exemple 3

Texte :

San Salvador, 19 avril 1989 (ACAN-EFE) – Le président du San Salvador Alfredo Cristani a condamné l'assassinat d'origine terroriste du ministre de la justice Roberto Garcia Alvarado et a accusé du meurtre le Front de Libération National Farabundo Marti.

Cadre :

INCIDENT	
<i>date :</i>	19 avril 1989
<i>lieu :</i>	El Salvador : San Salvador (CITY)
<i>auteur :</i>	Front de Libération National Farabundo Marti
<i>victime :</i>	Roberto Garcia Alvarado

Exemple 4 : information semi-structurée

Responsable artistique web senior (H/F)

Nous recherchons un(e) responsable artistique web senior pour rejoindre l'équipe créative de notre agence de communication intégrée, et prendre en charge son pôle Webdesign. Le rôle de ce responsable artistique web sera de manager l'équipe de webdesigners du pôle, et d'être force de proposition sur l'intégralité de nos projets web.

Vos missions sont :

- Concevoir l'habillage graphique du site, dans le respect de la charte et de l'identité visuelle de Verte-privee, pour toutes les différentes rubriques
- Intervenir sur des projets novateurs et développer les futures versions du site
- Organiser la charge de travail des webdesigners de l'équipe, les orienter dans leurs pistes créatives, et assurer le suivi des créas.

Votre profil :

Détenteur d'une formation artistique, vous avez un parcours créatif d'au moins 5 ans dans l'univers du web, idéalement en agence ainsi qu'une première expérience significative en management d'équipe créée.

Vous savez vous appuyer sur vos qualités relationnelles et sur votre sensibilité marketing pour présenter et défendre en interne les propositions créatives de votre équipe.

Une parfaite maîtrise de Photoshop, une bonne connaissance de Flash et de solides notions d'intégration (Dreamweaver et CSS) sont indispensables sur le poste.

Titre	Responsable artistique Web senior (H/F)
Expérience	<ul style="list-style-type: none"> - au moins 5 ans dans l'univers du web - première expérience significative en management
Qualités	<ul style="list-style-type: none"> - qualités relationnelles - sensibilité marketing
Technique	<ul style="list-style-type: none"> - Photoshop - Flash - Dreamweaver - CSS

ier Tannier

Exemple 5 : information structurée

amazon.com

Hello, **Xavier Tannier**. We have [recommendations](#) for you. ([Not Xavier?](#))

[Xavier's Amazon.com](#)



[Today's Deals](#)

[Gifts & Wish Lists](#)

[Gift Cards](#)

[Shop All Departments](#)

Search

Musical Instruments

[Bestsellers](#)

[Guitars](#)

[Bass](#)

[Keyboards](#)

Customers viewing this page may be interested in this **Sponsored Link** ([What's this?](#))

[Hi Quality Vertical Mills](#) Get the best for your **machine** shop with quality Lagun Vertical Mills

Singing Machine ISM-370 Multi-Function Karaoke System with iPod Docking

Only products by [Singing Machine](#)

★★★★☆ (3,411 customer reviews)

Wish Price: ~~\$199.99~~

Price: **\$129.99** & this item ships for **FREE with Super Saver Shipping**. [Details](#)

You Save **\$70.00 (35%)**

In Stock.

We are not able to ship this item to your default shipping address.

Ships from and sold by **Amazon.com**.

[2 new](#) from **\$97.95** [1 refurbished](#) from **\$75.00**



Gift Ideas for Musicians

Are you trying to figure out what to get the musician in your life for the holidays? We've got you covered with a great selection of



Exemple 5 : information structurée

Buy.com FREE 2 Day Express Shipping when you checkout with PayPal Look for the  PayPal icon for qualifying products. Sign In. New User? [Sign Up.](#)

[Products](#) | [Deals](#) | [BuyTV](#) | [News and Reviews](#) | [Funny TV Ads!](#)

[Home](#) » [Weekly Advertised Specials](#) | [Sell on Buy.com](#) | [Gift Certificates](#) Holiday

Singing Machine ISM-370 Multi-Function Karaoke System with iPod Docking

 **Save 16%**

FREE SHIPPING
Some restrictions apply. [Click here](#) for details.

List Price: [See Details](#) **\$129.99**
You Save: **(16%) \$21.36**
Our Price: \$108.63
Shipping: **FREE**

Buy.com Total Price: \$108.63

Currently Unavailable: [This item is currently unavailable from the Manufacturer.](#)

► [Alert me when this item is in stock.](#)

★★★★★ based on 2 reviews [enlarge image](#)

[Print this Page](#) [Email a Friend](#) [Add to Wishlist](#) [permalink](#)

SQUARE **Full refund** **3 Year Warranty \$17**

Exemple 5 : information structurée

eCOST.com Order Now: 1-877-888-2678
Your Online Discount Superstore!™

Rebate Center Helpdesk

Computers Networking Electronics & Entertainment TV's, Monitors & Projectors Cameras & Camcorders Memory & Storage For the Home


No Payments for up to 6 Months
Minimum purchase required. Subject to credit approval. [Details >](#) [BillMeLater](#)

DAILY DEALS AgfaPhoto 12.0 MegaPixel Compact Digital Camera - COMPACT 10

ELECTRONICS > **PRO**

SINGING MACHINE-PRO SOUND SINGING MACHINE KARAOKE SYS W/IPOD DOCK - ISM-370

Other products by SINGING MACHINE-PRO SOUND



Price: \$116.99

QTY: 1 [+ ADD TO CART](#)

[ShareThis](#) | [f](#) [t](#) [e](#) [★](#) [🖨](#)

[Click here for Canadian Pricing](#)

All mechanical and electrical failures • 100% parts & labor • No deductibles

SQUARE TRADE 3 Yr. Warranty - \$14.99 [ADD WARRANTY](#)

☐ Protect from Drops and Spills [?](#) [details](#)

Available at eCOST.com
Manufacturer: MFG
Item:

Exemple 5 : information structurée

KEENZO **KEEN ON QUALITY**

Search :

Holiday Deals

CATEGORIES

- Accessories
- Apparel
- Appliances
- Automotive
- Aviation
- Barcode / POS
- Barebone Systems
- Batteries
- Boating / Marine


AUDIO / VIDEO DEVICES > STEREO / CD / TURNTABLE / AMPS > ISM-370

SINGING MACHINE KARAOKE - ISM-370

Manufacturer:
Part#: ISM-370
Pkg Size: 16.5 X 9.8 X 10.2
Product Condition: Brand New, Full Warranty
UPC: 47237003701
☆☆☆☆☆ (0 Customer Reviews)

[eMail Item](#) [To Wishlist](#) [Ask Us](#) [Print](#)

At Keenzo.com our customers come first. If a product is out of stock, we will email you the moment this item is restocked.



[Zoom >>](#)

Cart Items:
Sub Total: \$
[My Cart](#) [Check Out](#)

[Contact Us](#) [Contact Us](#) [Newsletter](#) [My Account](#) [Log In](#)

\$109.99
MSRP ~~\$129.99~~
You Save (15.4%) ~~\$20.00~~
Shipping [?](#) **\$14.75**

Sold Out Until : 12/22/2009 [Alert Me](#)

Have our staff email you the moment this

Exemple 5 : information structurée

amazon.com

Buy.com

eCOST[®].com
Your Online Discount Superstore!™



AGRÉGATION

Singing Machine ISM-370 Multi-Function Karaoke System with iPod Docking

Singing Machine ISM-370 Multi-Function Karaoke System with iPod Docking

SINGING MACHINE-PRO SOUND SINGING MACHINE KARAOKESYS W/IPOD DOCK - ISM-370

SINGING MACHINE KARAOKE - ISM-370

NORMALISATION



CHOIX

\$129.99

\$108.63

\$116.99

\$109.99

(minimum)

CALCUL

Exemple 5 : information structurée

Google products [Advanced Product Search](#) [Preferences](#)

Products *Choix*

Show only:
☐ Google Checkout
☐ Free shipping
☐ New items


Enter location for tax and shipping: Sort by:


Any category
[Audio Players & Recorders](#)
[Karaoke](#)
[MP3 Players](#)
[Video Players](#)
[Sheet Music](#)
[More »](#)


Any price
[Under \\$60](#)
[\\$60 - \\$150](#)
[\\$150 - \\$300](#)
[Over \\$300](#)
\$ to \$


Any brand
[The Singing Machine](#)
[VocoPro](#)
[Me](#)
[Err](#)
[Ac](#)

Karaoke at Walmart
[Walmart.com](#) Save on Kids Electronics & More! Holiday Deals are at Walmart.

The Singing Machine ISM-370 - Karaoke system w/ iPod dock

Microphone, Internal, LCD display, Remote control
Top load CD/CD+G player AM/FM with digital tuner Alarm clock with snooze function Wake up to iPod/Radio/Buzzer/CD Microphone echo control Microphone volume control A.V.C (Auto ...
★★★★☆ 8 reviews - [Add to Shopping List](#)

Memorex MKS 8590 Karaoke system

CD/CD+Graphics compatible - displays song lyrics on the monitor or your television screen; LCD display for SD/MMC and USB thumbdrive read out; Two-digit LED display for CD ...
[Add to Shopping List](#)

The Singing Machine STVG-999 - Karaoke system

Microphone, wireless microphone, Display, Internal
This product is a pedestal CD/CD+G **karaoke** system, with 7" monitor, video camera and tower speakers.
★★★★☆ 13 reviews - [Add to Shopping List](#)

The Singing Machine STVG-535 - Karaoke system

Microphone, Cassette deck, Display, Internal

Normalisation

Calcul

\$90 new, \$89 used
from 71 sellers

\$85 new, \$90 used
from 19 sellers

\$170 new
from 5 sellers

\$95 new
from 8 sellers

Exemple 5 : information structurée

Google products [Advanced Product Search](#)
[Preferences](#)

The Singing Machine ISM-370 - Karaoke system w/ iPod dock from The Singing Machine in Combined AV Devices

[Overview](#) - [Compare prices](#) - [Reviews](#) - [Technical specifications](#) - [Similar items](#)



\$90 new from [64 sellers](#)

★★★★☆ [8 reviews](#)

Microphone, Internal, LCD display, Remote control

Top load CD/CD+G player
AM/FM with digital tuner
Alarm clock with snooze function
Wake up to iPod/Radio/Buzzer/CD
Microphone echo control... [more »](#)

[Add to Shopping List](#)

Compare prices

Show only: ☐ Google Checkout ☐ Free shipping ☐ New items

Enter location for tax and shipping:

Relevance ▼	Seller rating	Condition	Tax and shipping	Total price	Base price
Amazon.com	★★★★☆ 3,973 seller ratings	New	No tax + Free shipping	\$109.57	\$109.57
Buy.com	★★★★☆ 72,904 seller ratings	New	No tax + Free shipping	\$108.63	\$108.63
Provantage	★★★★☆ 1,412 seller ratings	New	No tax + Shipping: \$13.45	\$119.44	\$105.99
keenzo.com	★★★★☆ 247 seller ratings	New	No tax + Free shipping	\$107.99	\$107.99
Cymax Stores	★★★★☆ 928 seller ratings	New	No tax + Free shipping	\$111.38	\$111.38
eCOST.com	★★★★☆ 8,099 seller ratings	New	No tax + Shipping: \$17.59	\$134.58	\$116.99
Savingslots.com	★★★★☆ 138 seller ratings	New	No tax + Shipping: \$12.50	\$112.45	\$99.95
Ambient Weather Show all 2	★★★★☆ 1,787 seller ratings	New			\$112.04
FadFusion	★★★☆☆ 16 seller ratings	New			\$105.99
eBay	No rating	New	No tax + Shipping: \$15.99	\$105.98	\$89.99

Agrégation

[View all 64 sellers »](#)

Sellers 1 - 10 of 64 [Next »](#)

Niveaux d'analyse utilisés (ou pas)

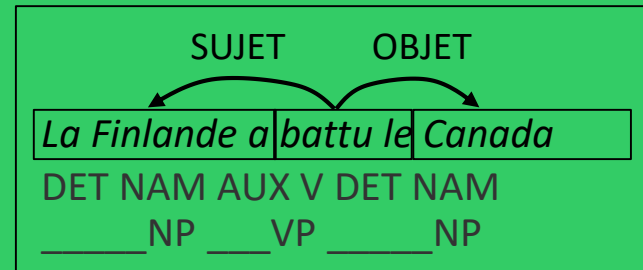
Morphologie

acheter
achètent
a acheté } acheter

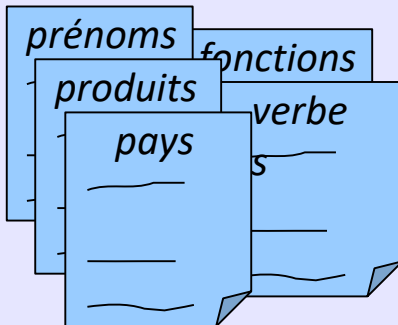
acheter
rachat
acheté } achat

Achats → achat

Syntaxe



Lexiques



Entités nommées

Gaël Montfils a remporté l'open de Metz en 2009

PERSON LIEU(ville) DATE(année)

Sémantique

SUJET(battre, X) & OBJET(battre, Y)
→ match(X, Y), vainqueur(X), vaincu(Y)

Niveaux d'analyse utilisés (ou pas)

Indications de mise en forme

Liste des codes pays :

EE – <i>Estonie</i>

ET – <i>Éthiopie</i>

FK – <i>Maldives</i>

FR – <i>France</i>

Mécanismes d'inférence

*"Federer a réalisé un superbe coup droit gagnant
lors d'une balle de match en sa faveur."*

`point(P) & coup_gagnant(P, J)
→ vainqueur(J, P)`

`balle_de_match(B, P)
→ point(P)`

`match(J1, J2) & balle_de_match(B, J1) & vainqueur(J1, B)
→ vainqueur(J1), vaincu(J2)`