



# NLP model and tools for detecting and interpreting metaphors in domain-specific corpora

## IsoMeta Project

P. Beust, S. Ferrari, V. Perlerin  
GREYC UMR 6072 – University of Caen - France

Interdisciplinary Workshop on Corpus-Based  
Approaches to Figurative Language



# Talk overview

- Goals
- Framework
  - Metaphor
  - Lexical Representation
  - Isotopy
- Tools & Results
  - Browsing and Visualization
  - Examples
- Conclusion and further works



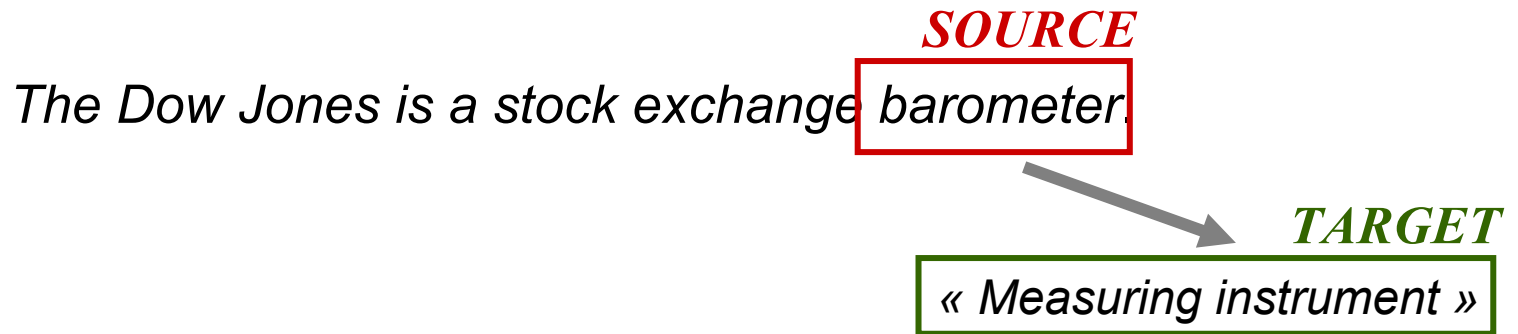
# Goals

- Metaphor detection
- Computer-assisted metaphor interpretation
- ☞ Metaphor study
  - user-centered approach
  - corpus-based approach



# Metaphor

- A metaphor involves two concepts : a source concept and a target concept.



- NLP approaches for metaphor interpretation depend on how the relation is viewed :
  - » as an analogy
  - » as a novelty



# Conventional metaphor

- Conventional conceptual metaphor
- **The corpus**
  - about 600 articles from the French newspaper “Le Monde”, addressing economics and stock market between 1987 and 1989.
  - contains numerous examples of this conventional metaphor
  - also contains meteorological terms not used in a figurative way.
- In order to address the interpretation process of the metaphor, we focus on a conventional conceptual metaphor already observed in the corpus :
  - « *economics is meteorology* »



# Lexical representation

- **The Anadia model :**

- is based on both componential and differential lexical categorization.
- allows a user to produce his own descriptions of meaning components.

Pressure zone	Zone	Pressure
anticyclone	area	high
tropical wave, easterly wave	line	low
depression, cyclone	area	low
	line	high

A *seme* with 2 opposite features

A *table* made from the combinatory of 2 semes features

A *set of tables* describing a domain

Domain objects	Role
stocks, currency (...)	playing a part
charts, ratio, stock indices, curves (...)	studying, analysing

Stock Indices	Geographical Zone
CAC, CAC40	France
Dow Jones, Nasdaq	USA
Nikkei	Japan
Dax	Germany
Footsie	UK



# Isotopy

- Isotopy

- is the recurrence of one *seme* in a linguistic unit.
- is an easy and understandable way of expressing themes in linguistic units.
- is the base of our interpretation and detection process.

After starting off as a formidable **tropical wave** of the coast of Eastern Africa, it quickly developed into a **depression** ...

<http://members.aol.com/windgusts/Alex.htm>

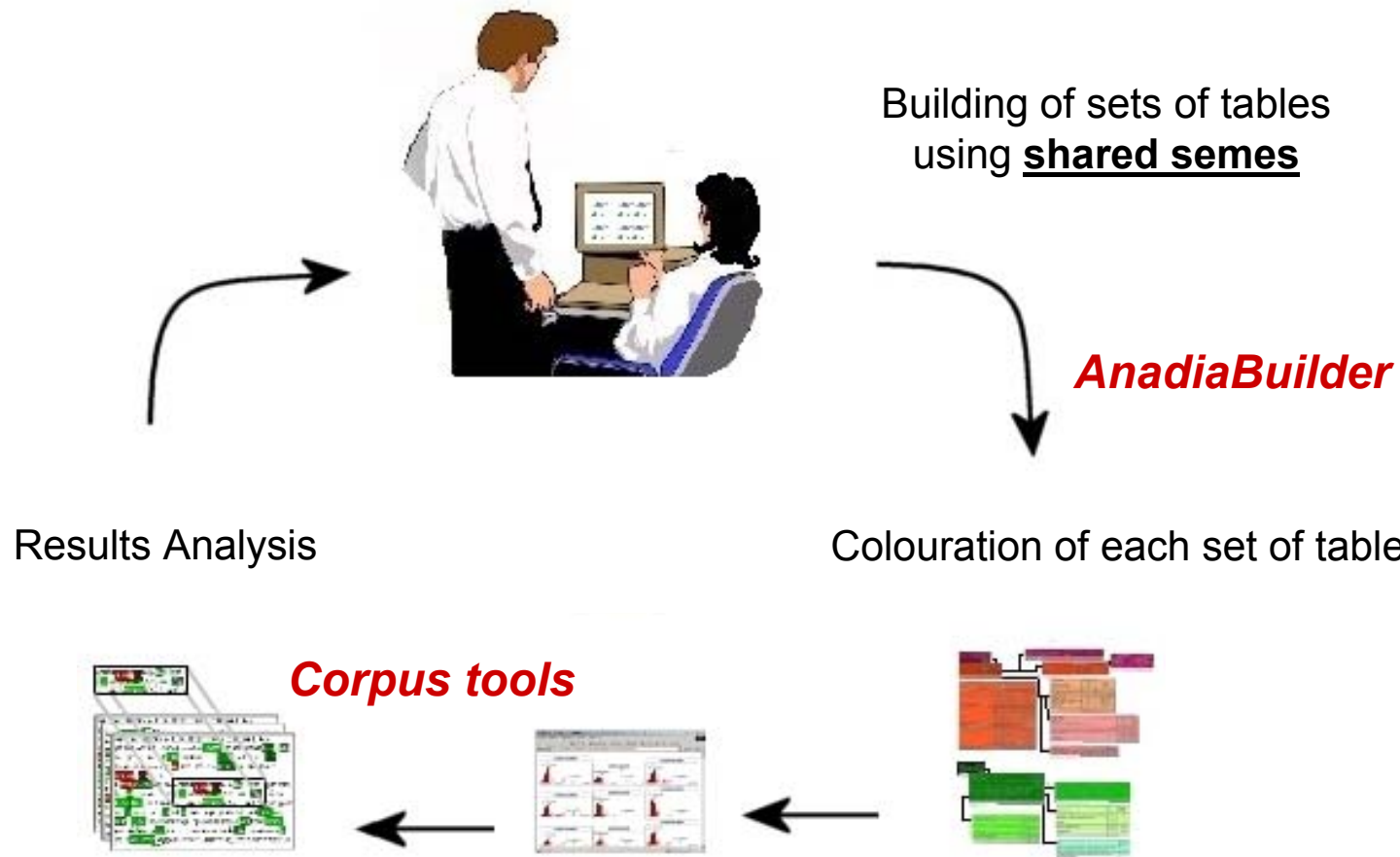
Pressure zone	Zone	Pressure
anticyclone	area	high
tropical wave, easterly wave	line	low
depression, cyclone	area	low
	line	high

☞ The two recurring *semes* involve two isotopies that contribute to the meaning of the sentence.

☞ The recurring features also show that the sentence deals with different kind of pressure zones.

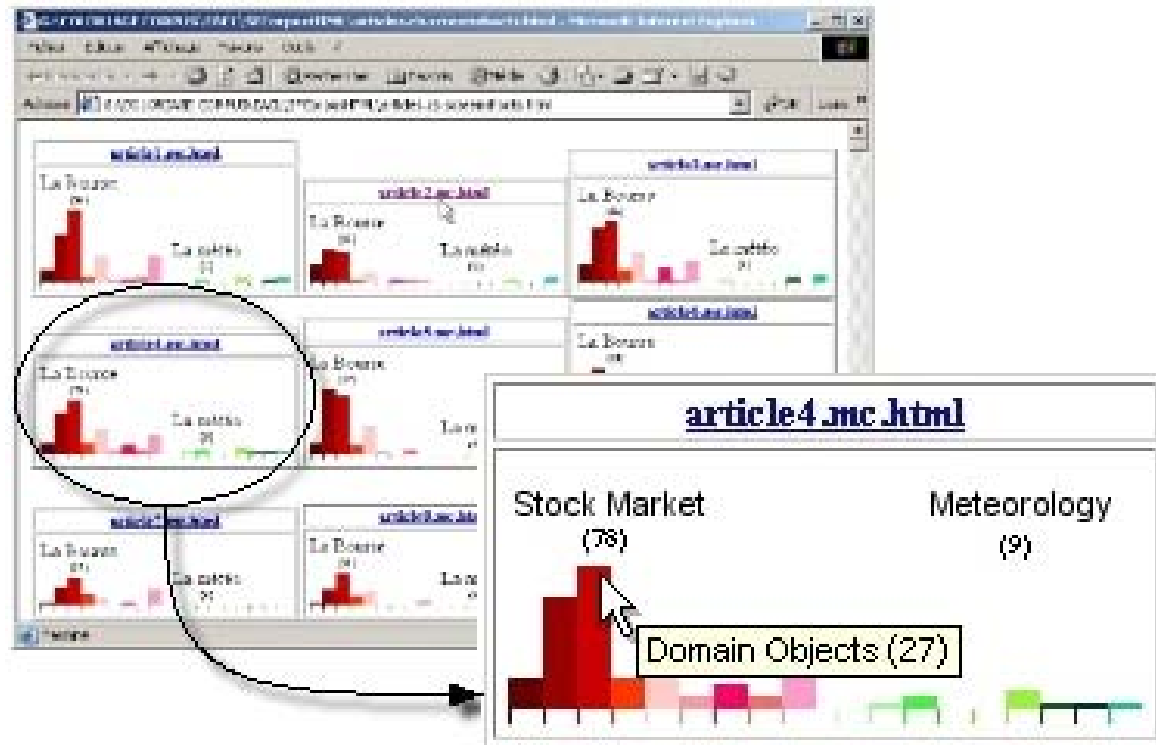


# Analysis





# Corpus study tools (1)



# Corpus study tools (2)

Microsoft Internet Explorer  
G:\COLORIAGE CORPUS\EACL\SFCorpusHTML\article45.mc.html

## Les Saints de **glace** **BOURSE** DE PARIS SEMAINE DU 11 AU 15 MAI

DESSOT ANDRE

Secteur rédaction : *Economie Analytique*  
Sujets en France : **BOURSE**

SAINT Mamert, saint Pancrace et saint Servais, les fameux Saints de **glace**, doivent se frotter les mains. Leur passage annuel a, cette semaine, été couronné de succès. La **température** a très brutalement baissé, non seulement à l'extérieur, obligeant les plus frileux à remettre une petite laine, mais **rue Vivienne** aussi. La **Bourse** de Paris, qui depuis deux mois et demi pataugeait joyeusement dans une interminable consolidation, s'est repliée d'un bloc et à toute allure vers la **cote** 430 de l'**indice CAC**, péniblement atteinte au début du mois de mars dernier. Le **coup de froid** a été sévère (- 5, 5%) et d'autant plus inquiétant pour la "végétation mobilière" que le **thermomètre** a chuté de 4% au **cours** de la seule **séance** de vendredi.

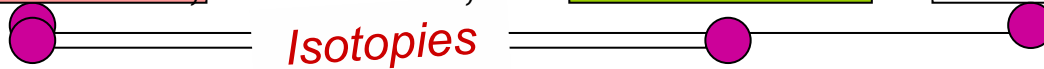
Ce jour-là, pour se redonner du courage devant la débandade des **cours**, des professionnels l'ont entendu : " Appelez la police pour arrêter le massacre! " Inutile de le préciser; aucune hausse **séance** du 15 mai. Cette semaine n'avait pourtant pas mal commencé. Avec l'assagissement de la **tendance** s'était un peu raffermie lundi (+ 0, 3%), ce qui avait paru assez bon signe. Pas longtemps; le lendemain, la **tendance** s'alourdissait un peu (- 0, 74%), ce qui ne sembla guère émouvoir la population du Palais National.

Outils de mesure  
Axe : température

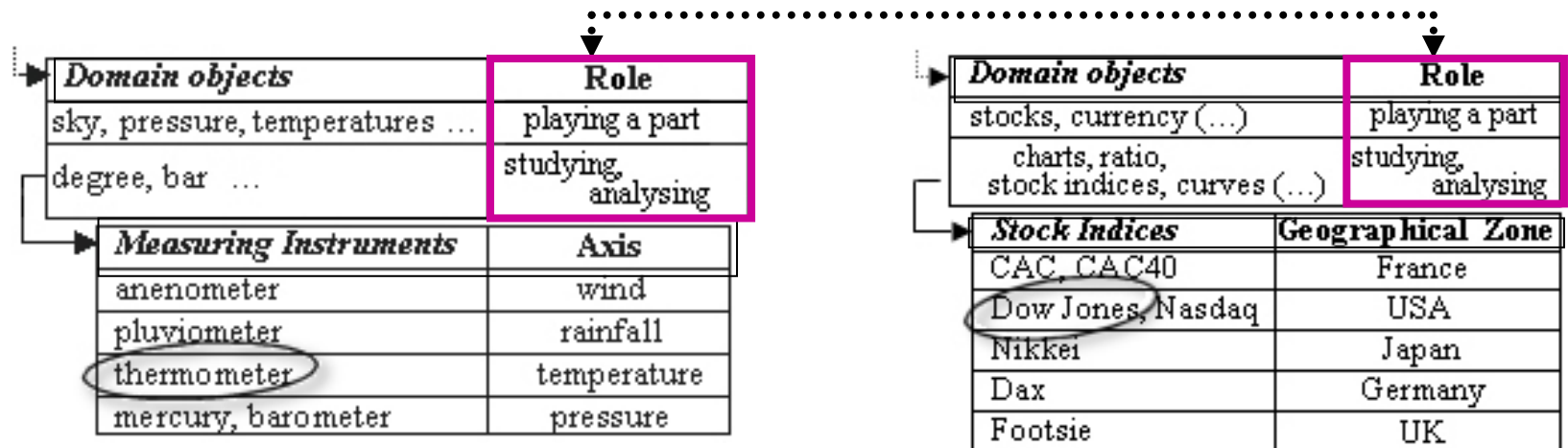


# Results 1

The **Dow Jones**, for instance, the **thermometer** of **Wall Street** which had fallen.



a shared seme



*thermometer* is used in the same way as “graphics”, “ratio”... i.e. to suggest an object for analysis and study in the stock market domain.

👉 analogy point of view



# Results 2

This **crash** was due (...) to the vertiginous and uncontrolled fall of the **dollar**, sig

Isotopies

at the **storm** will henceforth affect the **financial markets**.

shared semes

<i>Dynamic phenomena</i>	Direction	Connotation
depreciation, devaluation, <b>crash</b> , to devalue	down	bad
rise in prices, inflation	up	not connoted
drop in prices, deflation	down	not connoted

<i>Dynamic phenomena</i>	Direction	Connotation	Axis
frost, to freeze	down	bad	temperat
bad weather (...)	up	bad	weath
hull (...)	down	good	weath

<i>Bad weather</i>	Strength
gust, <b>storm</b> , gale, (...)	violent
cyclone, typhon, (...)	fierce

*storm* is used to evoke a not only bad but also violent dynamic phenomenon in the stock market domain.

👉 novelty point of view



# Conclusion and further works

- IsoMeta Project is still in progress :
  - automatic detection
  - interpretation process
  - results evaluation
- Shared semes represent previous user's knowledge on a conceptual metaphor
  - ▶ "new metaphors" are not covered by our approach.





# NLP model and tools for detecting and interpreting metaphors in domain-specific corpora

## IsoMeta Project

P. Beust, V. Perlerin, S. Ferrari  
GREYC UMR 6072 – University of Caen - France

<http://www.info.unicaen.fr/~perlerin/recherche/Anadia.html>

Interdisciplinary Workshop on Corpus-Based  
Approaches to Figurative Language

