

# Evaluation automatique de textes dans un contexte d'apprentissage : Quelques applications de l'analyse de la sémantique latente



Philippe Dessus – Conférence Master Industrie de la  
langue – DIP U. Stendhal Grenoble – 7 oct. 2011



# Plan

- o. Mes questions de recherche
- 1. Le contexte & LSA
- 2. Deux systèmes : *Pensum*, *Résum'Web*
- 3. Boucles, écarts et contradiction (BEC) dans la construction d'un environnement personnel d'apprentissage
- 4. Applications de LSA dans l'éducation : une revue

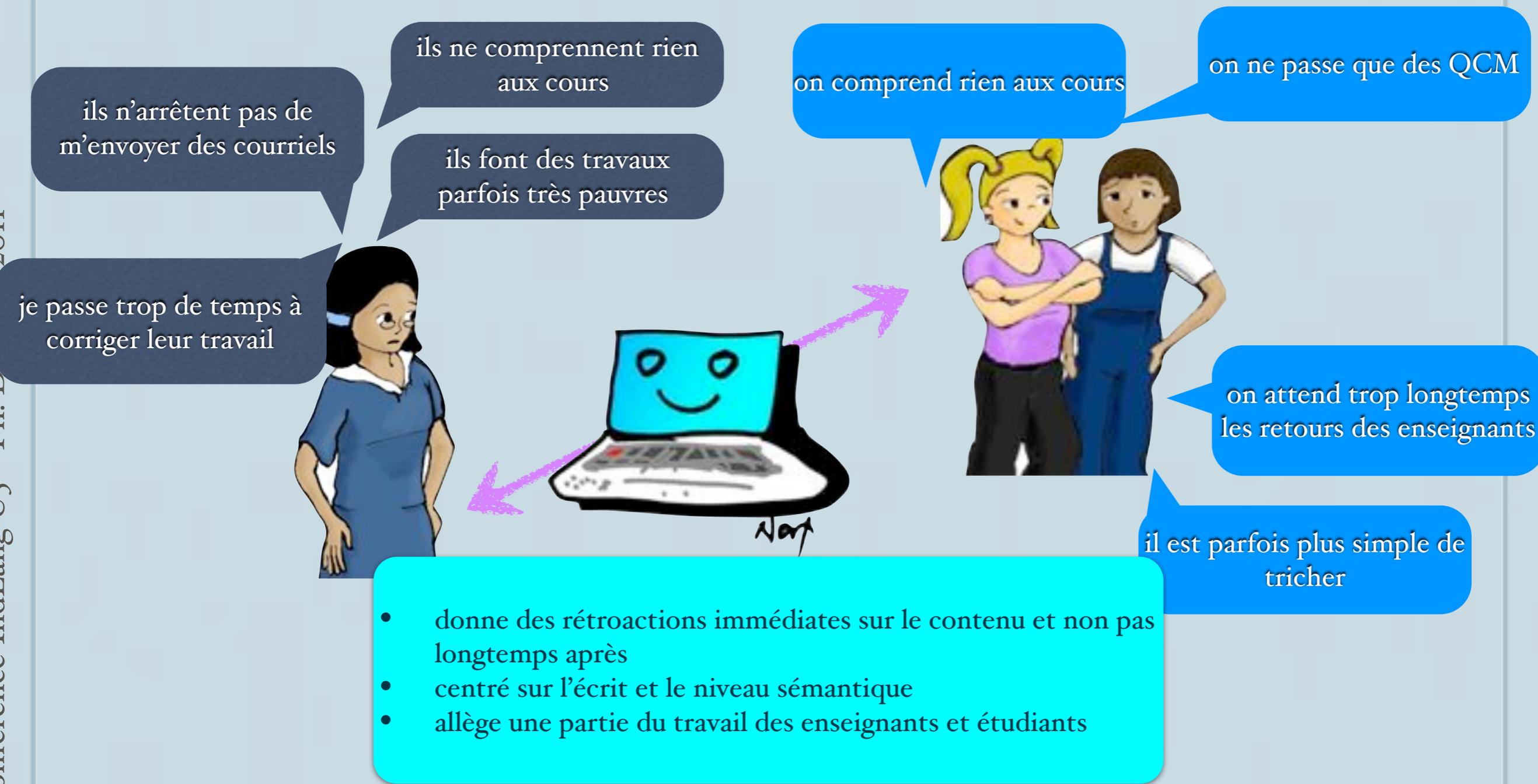
Toutes les références disponibles à <http://www.citeulike.org/user/pdessus/tag/indlang>

# Mes questions de recherche

- ❖ Comment outiller les enseignants et les élèves pour leur procurer des aides automatisées
  - ❖ au sein d'environnements personnels d'apprentissage, EIAH)
  - ❖ dans des contextes d'enseignement à distance ou en présence
  - ❖ leur permettant de comprendre le matériel étudié (cours, textes) et leur donnant des rétroactions sur leur travail ?
- ❖ Quels outils informatisés ? Quel design ? Quels usages ?
- ❖ Travail interdisciplinaire : sciences de l'éducation, psychologie, sciences du langage, informatique (sciences cognitives)

# 1. Le contexte et l'Analyse de la sémantique latente

# Enseignement : Contexte et problèmes [Dupré, Loiseau & Dessus II]



images © <http://daywoodacademy.org/cast>

# Résumer pour comprendre/ en comprenant

- ❖ Redire, à l'oral ou résumer, à l'écrit ou encore expliquer, est un indicateur et/ou un catalyseur de la compréhension des lecteurs (voir Chi et *al.* 94; Kastens & Liben 07), différences selon l'écart au texte-source
  - ❖ fait exprimer l'essentiel de ce qu'on a lu (Kintsch 98; Reed & Vaughn, sous presse)
  - ❖ fait se demander comment on a compris l'information contenue dans le texte lu (autorégulation, Thiede & Anderson 03)
  - ❖ Expliquer, c'est exprimer quelque chose de nouveau (nouveau schéma de la situation)

# 7 Apprendre en étant engagé dans une boucle de 2 activités

Lire



Ecrire



prendre de la distance  
rappeler

récrire

produire

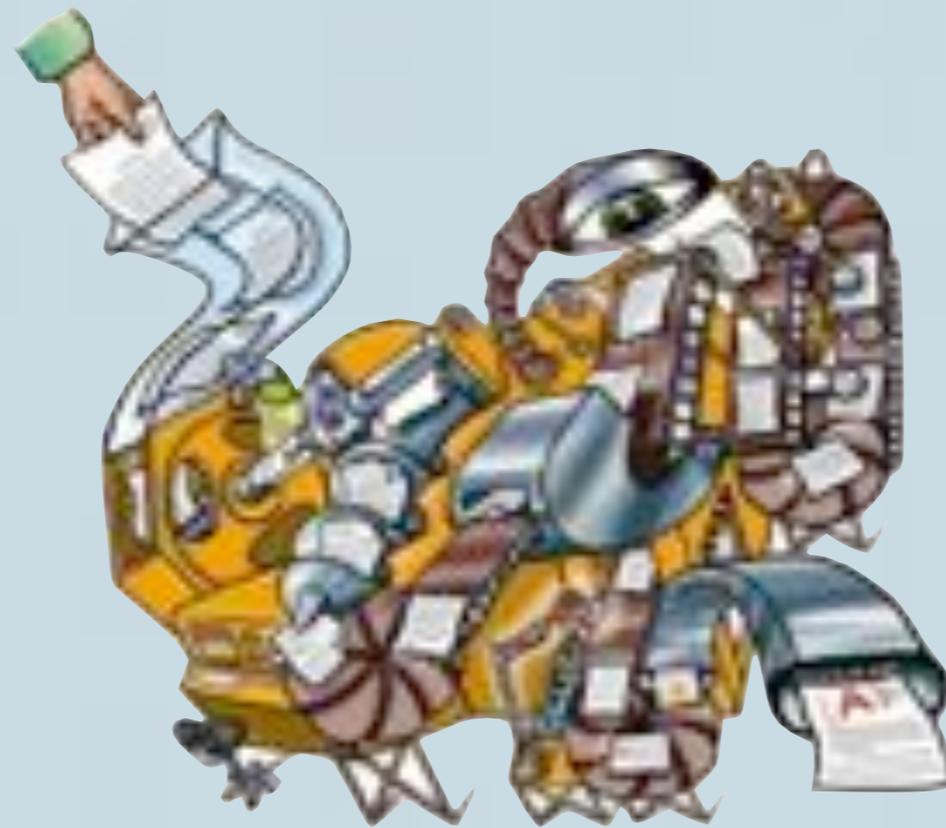
être réflexif

Rechercher  
recupérer  
comprendre  
être informé

(c) <http://www.flickr.com/photos/jamesbt/3299261881/> & <http://www.flickr.com/photos/extraketchup/749315946/#/>

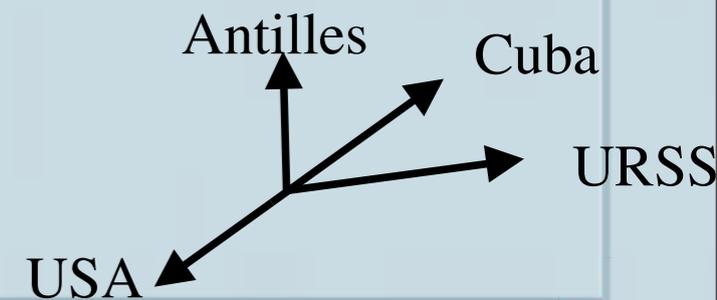
# Il me faut un moyen...

- ❖ ... d'analyser automatiquement et quasi-instantanément les liens sémantiques entre ce qui est lu et produit à l'écrit



# Comment marche l'Analyse sémantique latente ? (1/2)

- ❖ LSA méthode d'analyse statistique de grands corpus textuels (type d'analyse factorielle). Part du principe que
  - ❖ deux mots ont un sens similaire s'ils apparaissent dans des contextes similaires
  - ❖ deux contextes (paragraphes, phrases) ont un sens similaire (contiennent des informations similaires) s'ils contiennent des mots de sens similaire

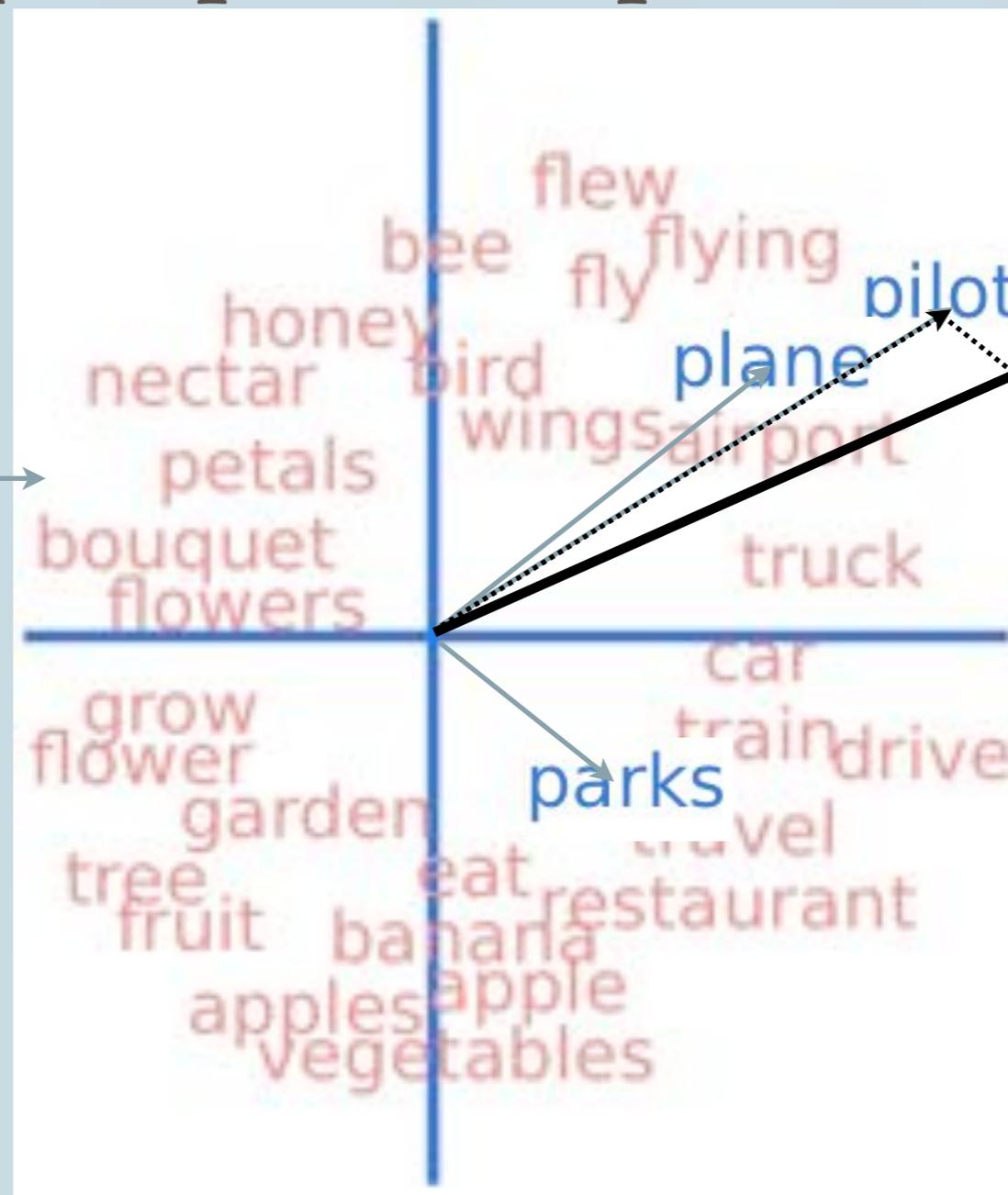
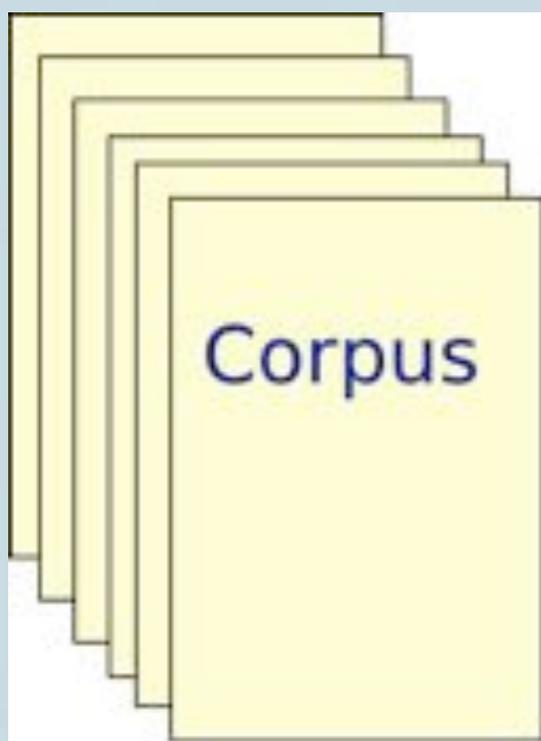


# Comment marche LSA (2) ?

- ❖ Rend compte de la structure *latente* du sens des mots
- ❖ Importance de la composition du corpus pour «entraîner» LSA
- ❖ Pas seulement fondée sur la co-occurrence de mots : 2 mots jamais co-occurents peuvent avoir une grande proximité
- ❖ Approche «paquets de mots», pas de prise en compte de la ponctuation, de la syntaxe, du style, etc.
- ❖ Pas de prise en compte de «mots outils» (déterminants, prépositions)

Une phrase est représentée par la somme des vecteurs des mots qui la composent  
[Lemaire & Denhière '05]

**“The pilot parks the plane”**



# Quelques sorties LSA avec un corpus «enfant» [Denhière *et al.* 07]

I2

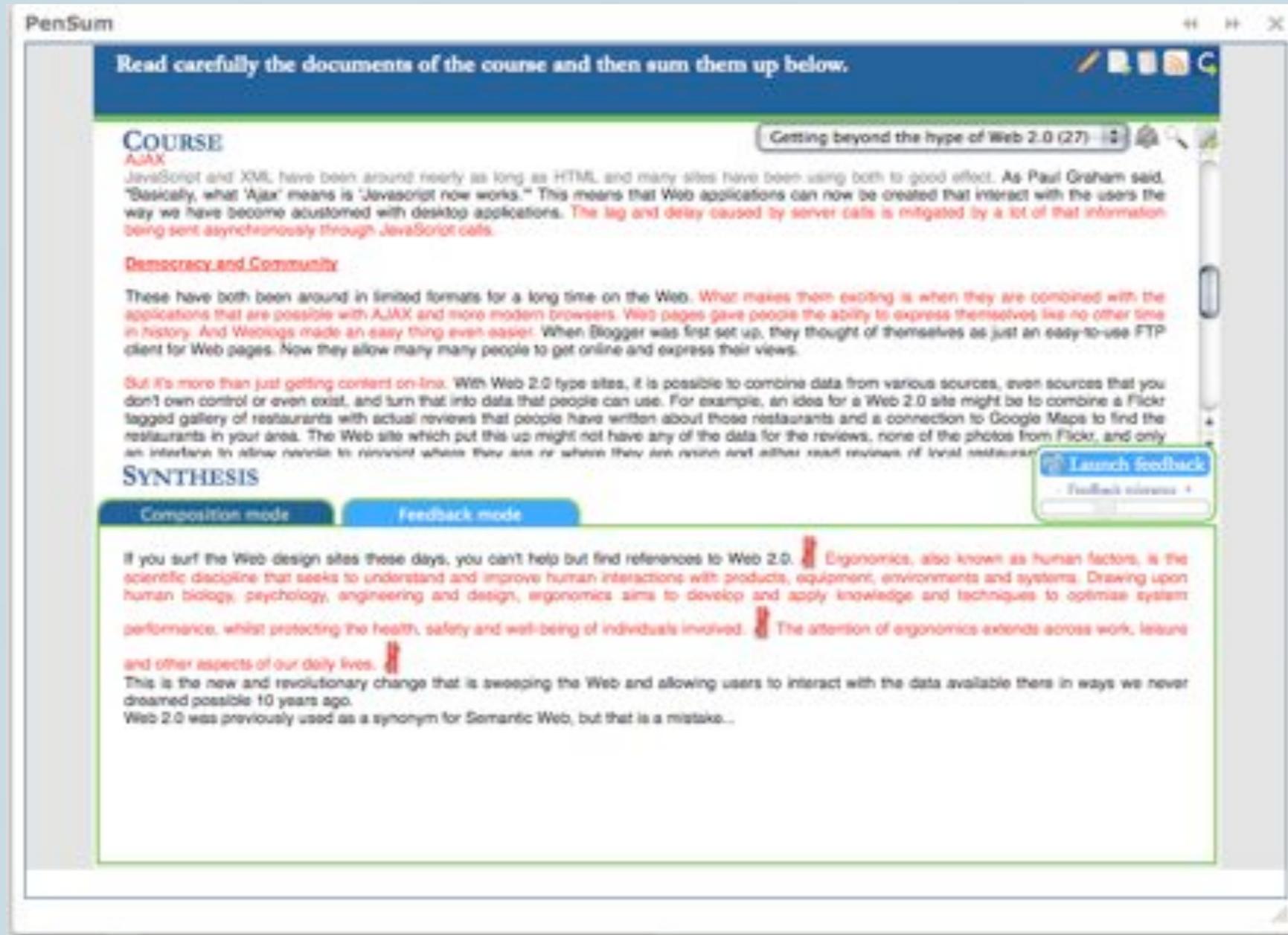
❖ Les 10 mots les plus proches des mots suivants (cosinus)

<b>couteau</b>	prox.	<b>ami</b>	prox.	<b>rire</b>	prox.
coupa	.53	convoite	.80	cynique	.66
prit	.47	charly	.48	éclata	.64
fourchette	.44	grilla	.45	gloussements	.59
précaution	.42	meilleur	.44	riait	.47
lame	.41	motard	.43	pouffer	.47
tranches	.40	accompagnons	.41	clowns	.46
assiette	.40	inquiets	.40	mit	.46
couper	.38	malheureux	.39	trapézistes	.44
assit	.38	louveteau	.39	jongleurs	.43
poche	.38	indolent	.38	éclater	.43

II. Deux systèmes :  
*Pensum & Résumé'Web*  
et leur validation en  
contexte

*Pensum* : Un système d'évaluation automatique de synthèses et résumés de cours. 3 types de rétroactions [Dupré *et al.* 11]

I4



# *Pensum : Contexte de validation*

- ❖ 17 étudiants (dont 15 femmes) de M2 pro sciences de l'éducation (âge moyen : 33 ans), appariés en «étudiant» et «tuteur» :



réaliser en 1 h une synthèse d'un ensemble de 3 documents de cours



lire et comprendre le cours, observer la fin de la rédaction de la synthèse de l'«étudiant», sans intervenir, puis l'évaluer via l'interface

- ❖ Données recueillies : *bugs*, opinion et traces

# *Pensum* : Résultats

## [Dupré *et al.* 11]

- ❖ Traces
  - ❖ Nombreux *feedback* demandés, de manière régulière, en ajustant sa sévérité. *Feedback* le plus remis en question : cohérence entre phrases (les autres l'ont très peu été)
- ❖ Opinion
  - ❖ 5 participants intéressés ; 10 plus critiques (notamment sur la validité des rétroactions) ; 2 négatifs
- ❖ *Bugs*
  - ❖ Quelques *bugs* relevés et corrigés (compatibilité moteur recherche et d'interface)

# *Pensum* : Conclusion

- ❖ Ce qui a été implicitement attendu de *Pensum*
  - ❖ c'est un outil donnant des réponses toujours exactes et des conseils de reformulation
- ❖ Ce que *Pensum* permet
  - ❖ les rétroactions (cohérence, pertinence, hors-sujet) mais aussi la simple remise en question d'un *feedback* (même imparfait) peut permettre une meilleure compréhension du cours
- ❖ Ce à quoi il faut veiller
  - ❖ meilleure information sur les fonctionnalités de *Pensum* (ce qu'il peut ou ne pas faire)
  - ❖ Démo @ <http://webu2.upmf-grenoble.fr/sciedu/ltfl-pensum/>

# Résum'Web, un outil diagnostiquant les opérations cognitives pendant l'activité de résumé

[Mandin *et al.* 07; Mandin 09]

18

Résumez le texte ci-dessous en X mots.

**TEXTE À RÉSUMER**

Ah, le cauchemar des premiers cheveux blancs, des rides au coin des yeux, des muscles qui ne suivent plus nos envies. Un mauvais rêve d'autant plus angoissant que l'on sait que ce ne sont que les préliminaires. La sélection naturelle, normalement, fait tout pour fabriquer des individus au top, armés pour survivre aux conditions les plus extrêmes. Mais pourquoi a-t-elle conservé au cours de l'évolution un mécanisme qui rend les êtres vivants de plus en plus fragiles avec l'âge ? Et ce, dans la quasi-totalité des espèces.

La première idée qui nous vient, c'est que la vieillesse a été inventée que pour nous faire mourir. En effet, le vieillissement est inscrit dans nos gènes. Il fait qu'avec l'âge nos cellules s'usent, arrêtent de se diviser, de se réparer. A un moment donné, il n'y en a plus assez qui fonctionnent correctement pour faire tourner nos organes vitaux (le cœur, le

**RÉSUMÉ**

Ah, le cauchemar des premiers cheveux blancs, des rides au coin des yeux, des muscles qui ne suivent plus nos envies. La vieillesse est inévitable et conduit à la mort à cause de la dégénérescence de nos cellules.

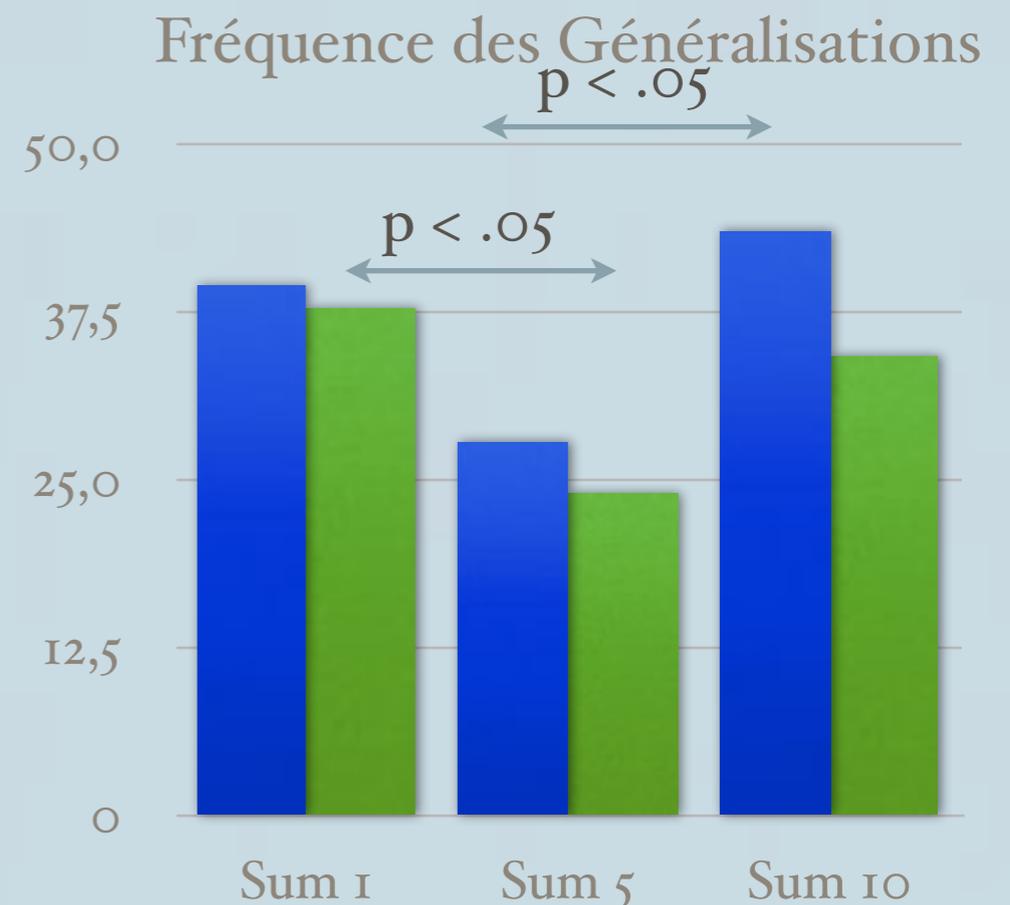
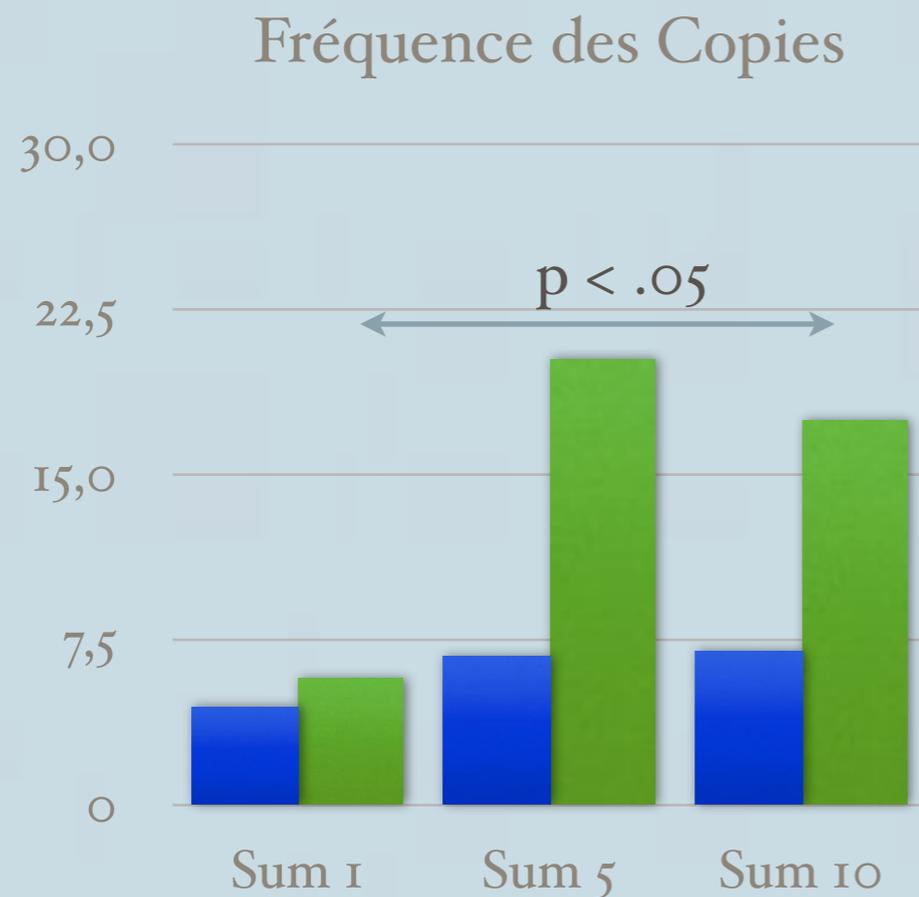
Nbre max. de mots restant : 13

# *Résum'Web* : Le diagnostic des processus de résumé [Lemaire *et al.* 05]

- ❖ 2 processus cognitifs interreliés dans l'activité de résumer
  - ❖ sélectionner les idées les plus importantes du texte source (**TS**)
  - ❖ appliquer sur les phrases du **ST** quelques macro-règles pour former le résumé (**Rés**) [Kintsch & van Dijk 78]
- ❖ *Copie* : une phrase **Rés** est très proche d'une phrase du **ST**
- ❖ *Suppression* : une phrase du **ST** est très éloignée de chaque phrase du **Rés**
- ❖ *Généralisation* : une phrase du **Rés** est proche de plusieurs phrases du **ST**
- ❖ *Construction* : une phrase du **Rés** n'est pas très proche de toutes les phrases du **ST**, mais plutôt proche de certaines d'entre elles
- ❖ Voir une démo @ <http://webu2.upmf-grenoble.fr/sciedu/smandin/demos/resumwebdemo.swf>

# Effet de *Resum'Web* sur l'écriture de résumés (*élèves de lycée, 10 résumés*)

■ Resum'Web ■ Contrôle



Pas d'effet sur la fréquence des hors-sujet et des constructions  
Pas d'effet sur les notes humaines (par enseignants)

Boucles, écarts et  
contradictions (BEC)  
*Section d'après [Dessus II]*

# Se construire un environnement personnel d'apprentissage (EPA)

22



- ❖ Quels outils pour quelle activité dans quelle partie du flux global de travail ?
- ❖ Comment choisir et différencier ces outils?

c) <http://www.flickr.com/photos/briansolis/520051406/>

# Un exemple : *Twitter* comme un EPI (pour information)

23

Une boucle d'activité dans *Twitter* et ses indications de guidage

Qu'est-ce que j'ai fait ?

Boucle lecture/  
écriture

Lecture

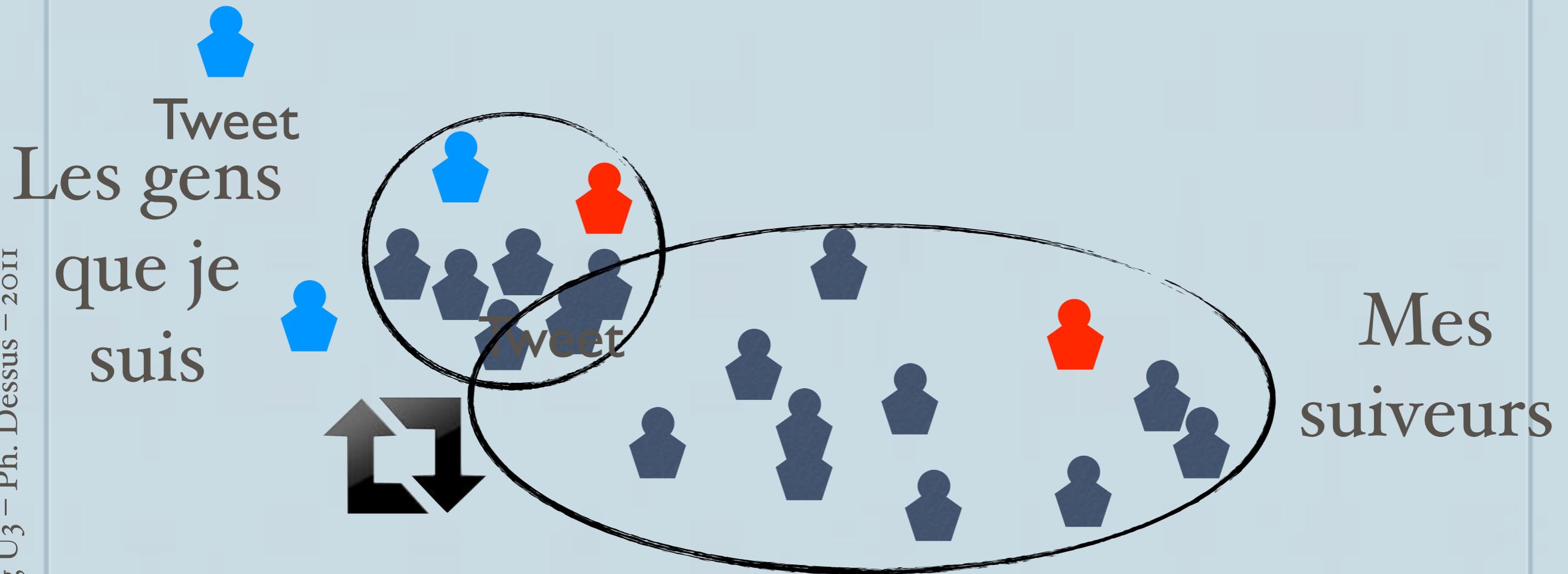
Ecriture



Que faire ensuite ?

# Comment réduire l'écart d'information ?

24



Est-ce que je diffuse simplement de l'information ou bien est-ce que j'essaie de *réduire l'écart* d'information dans ma propre communauté?

# 25 *Twitter* peut-il devenir un EPA (pour apprentissage) ?

- ❖ Des activités riches (boucle lecture/écriture), mais
  - ❖ manque de *rétroactions* immédiates ou rapides sur les *tweets* (pas seulement sur les personnes à suivre)
  - ❖ manque d'interactions plus intensives et d'attention partagée entre personnes (discussion d'idées, argumentations, contradiction, etc.)
- ➔ Exactement ce qu'on peut obtenir au travers du bavardage pertinent ou «pertinage» ou «*backchannelling*»

# La définition d'un EPA

[Johnson *et al.*, 2011, pp. 30-1]

- ❖ Ensemble d'outils indépendants, choisis par l'apprenant
  - ❖ qui peuvent être *connectés* ou utilisés conjointement de manière transparente
  - ❖ qui donne à leurs utilisateurs un certain *contrôle* sur leur apprentissage
  - ❖ et qui les encourage à appréhender l'apprentissage pour satisfaire au mieux leurs *besoins individuels*

# Questions posées dans la construction d'un EPA



BOUCLE

- ❖ *Connexion et «customisation»* des outils pose la question du flux de travail et des boucles (quelle combinaison pour quel apprentissage?)

- ❖ La question des *besoins individuels* pose la question de la manière dont ils sont diagnostiqué et gérés par l'outil (quel affichage, prise en compte?)

ECART

- ❖ Le *contrôle par l'apprenant* pose la question du flux de travail et de la prise en compte du feedback (quel type de contrôle ou de contradiction)

~~CONTRA-  
DICTION~~

# Quelques caractéristiques des boucles d'apprentissage [Carver & Scheier II; Goetz II]

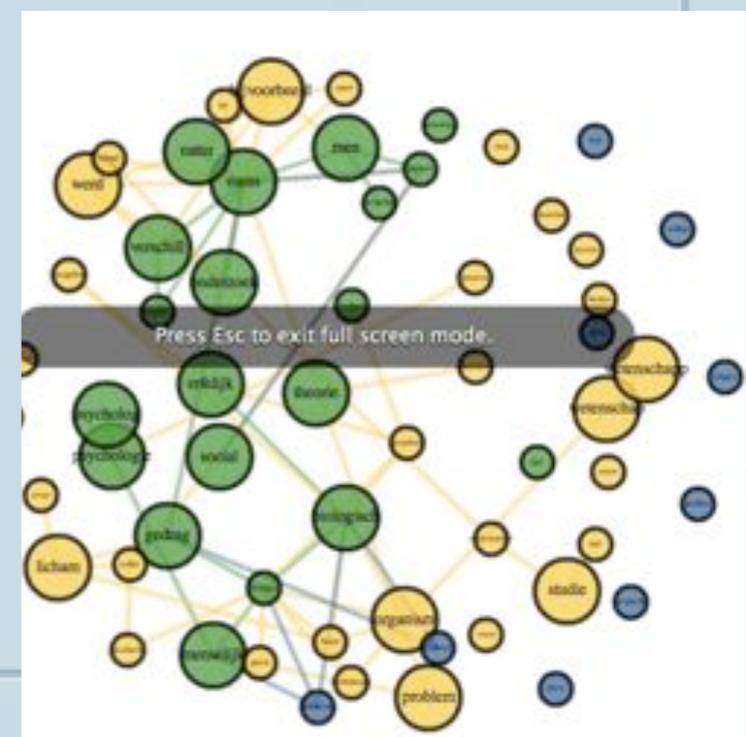
- ❖ Rendre l'activité complexe, pas l'interface
- ❖ Implication libre dans le travail (commencer/arrêter à n'importe quel moment)
- ❖ Flux d'activités libre (changer d'activité librement)
- ❖ Le feedback (humain ou machine) peut aider à l'apprenant dans son flux
- ❖ Promeut l'apprentissage autorégulé



<http://www.flickr.com/photos/reinvented/4825115719/sizes/l/in/photostream/>

# 2,9 Enseigner & apprendre : atténuer des écarts... [Wild *et al.* 10]

- ❖ ... entre apprenants (novices) et tuteurs/enseignants (experts)
- ❖ ... entre points de vue de pairs
- ❖ ... entre du contenu déjà appris et à apprendre...
- ❖ etc.
- ❖ Présenter des écarts permet à l'utilisateur de comprendre comment le système fonctionne



# La contradiction : un ingrédient de l'apprentissage [Stahl 06]

- ❖ Part essentielle de nombreux processus de haut niveau dans l'apprentissage : argumentation, résolution de problèmes
- ❖ Lorsqu'on est dans l'erreur : pas de vue alternative (peut-être vraie) sans contradiction
- ❖ Aide à comprendre le point de vue des autres
- ❖ Aide à construire des connaissances par le débat
- ❖ Le *feedback* n'est jamais sûr à 100 % et un *feedback* trop personnalisé gêne l'apprentissage [Pariser 11]

# BEC illustré

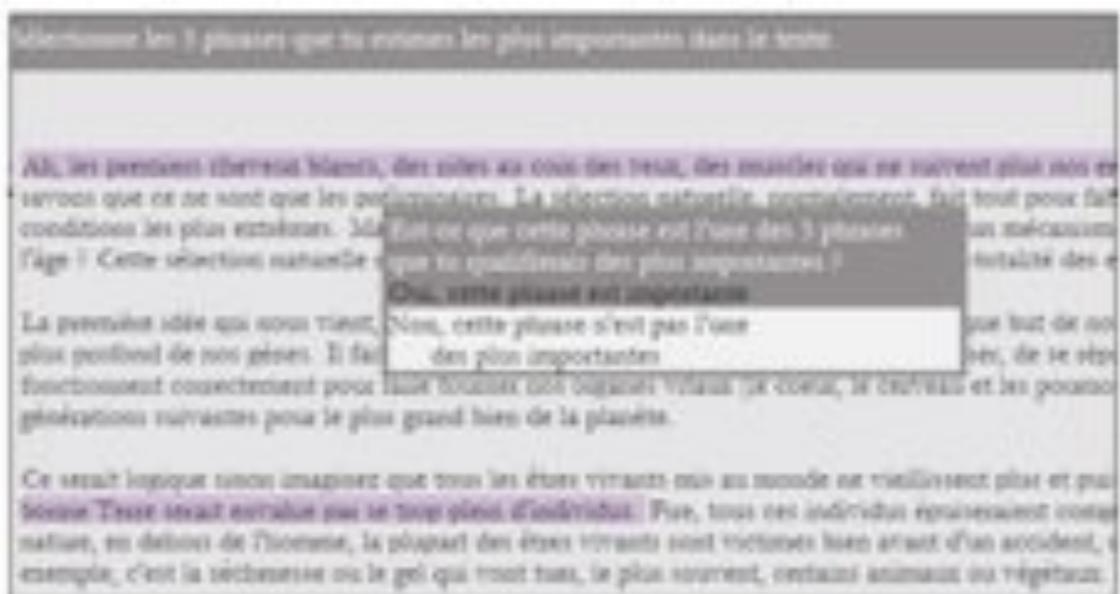
## I. Boucles

- ❖ Boucle lecture/écriture avec des rétroactions

# Les Résum'Web

[Lin 09, 2011]

32



Lire le  
texte



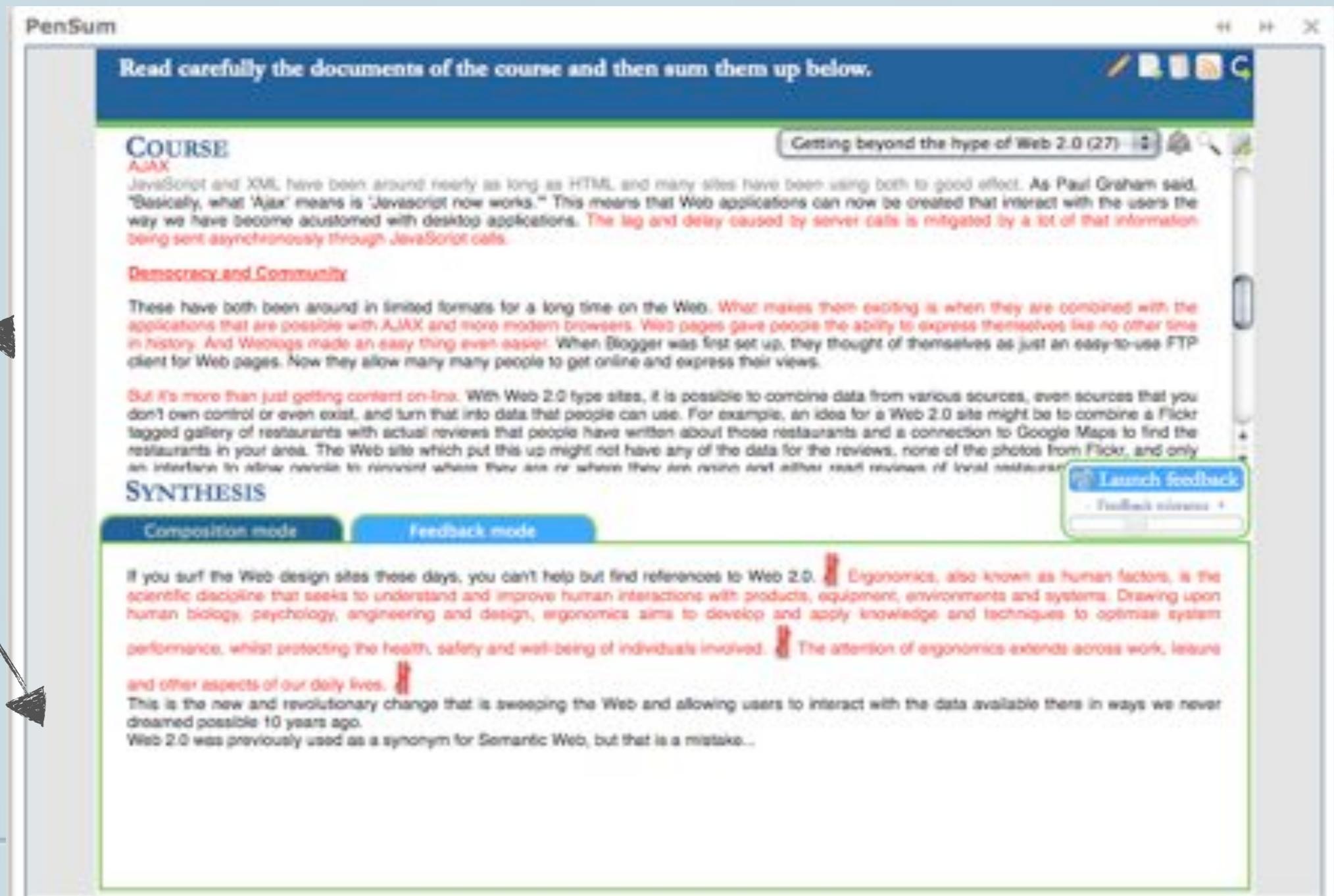
Lire le  
*feedback*

Ecrire

# Boucle dans *PenSum*

[Villiot-Leclercq *et al.* 10]

33



Conférence IndLang U3 – Ph. Dessus – 2011

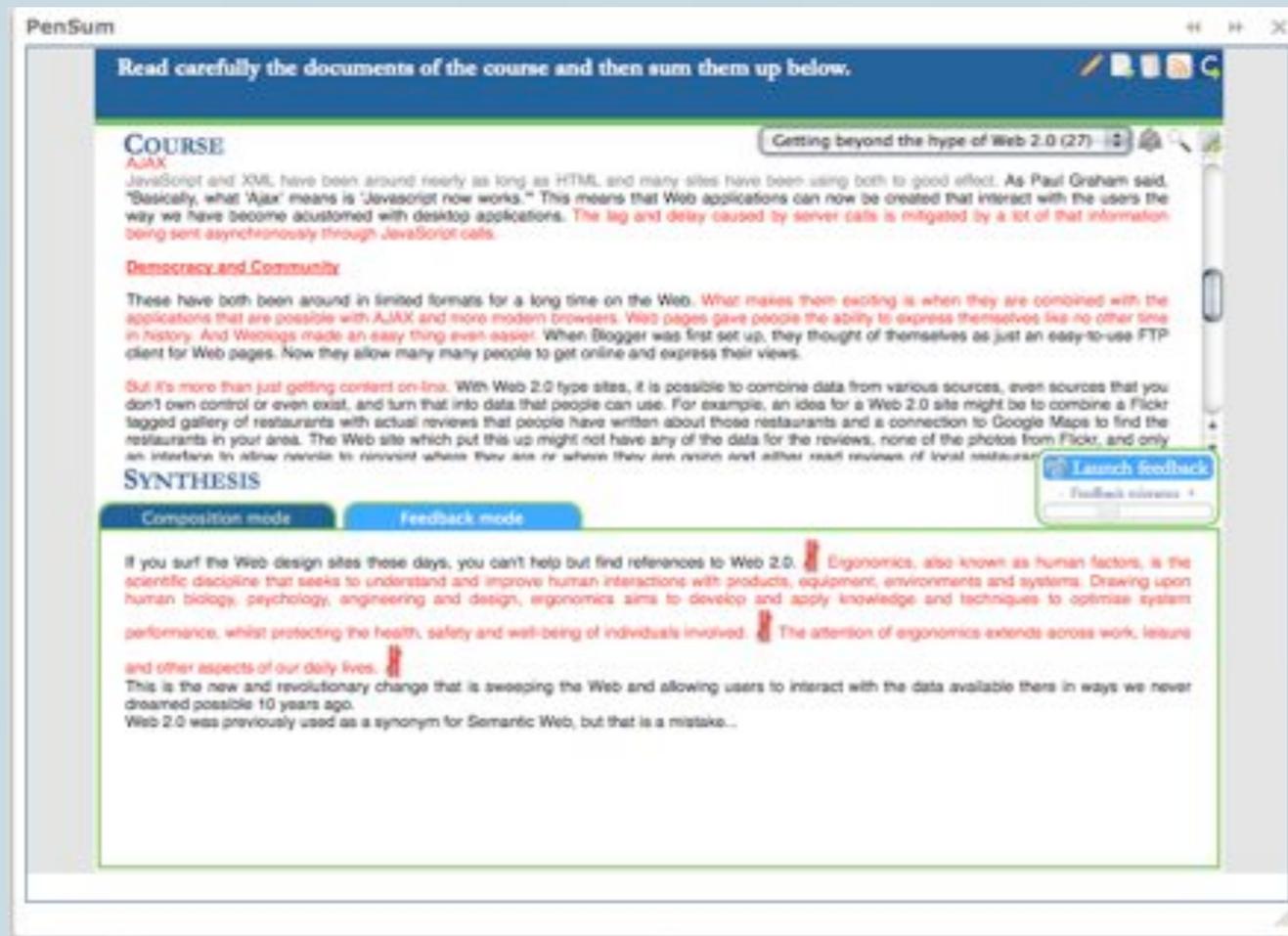
# BEC Illustré

## 2. Ecart

- ❖ Les écarts sont partout. Se focaliser sur eux aide à apprendre
  - ❖ dans des productions d'élèves (sauts de cohésion)
  - ❖ entre une production attendue et réalisée
  - ❖ entre apprenants, etc.

# Ecart dans *Pensum*

Sauts de cohésion : indices de mauvaise compréhension  
[Tapiéro 07]



# BEC Illustré

## 3. Contradictions

Lorsqu'un  
ordinateur vous  
contredit...

... et *vice versa*



# 37 Contradiction dans *Resum'Web*

Commentaires sur ton travail

Ces quelques contradictions que t'ai remarquées dans ton travail :

- Tu estimes que la phrase 2 de ton résumé est une Copie. Il me semble qu'elle reprend la (les) phrase(s) 10 du texte, phrase(s) que tu considères secondaire(s). Si tu as raison, ne serait-il pas plus judicieux de supprimer cette phrase ou de la regrouper avec une autre ?
- La phrase 5 du texte ne me semble pas reprise dans le résumé. Pourtant tu classes cette phrase parmi les plus importantes.
- La phrase 8 du texte ne me semble pas reprise dans le résumé. Pourtant tu classes cette phrase parmi les plus importantes.
- La phrase 13 du texte ne me semble pas reprise dans le résumé. Pourtant tu classes cette phrase parmi les plus importantes.
- La phrase 15 du texte ne me semble pas reprise dans le résumé. Pourtant tu classes cette phrase parmi les plus importantes.
- La phrase 17 du texte ne me semble pas reprise dans le résumé. Pourtant tu classes cette phrase parmi les plus importantes.

Résumés réalisés : 0  
Résumés restants : 10

# Contradiction dans *PenSum*

Rejet du *feedback* sur la cohésion interphrases

The screenshot shows the PenSum web application interface. At the top, a blue header bar contains the text "Read carefully the documents of the course and then sum them up below." Below this, the document content is displayed. The document has sections for "COURSE" (with sub-sections "AJAX" and "Democracy and Community") and "SYNTHESIS". The "SYNTHESIS" section is currently selected, and it contains text about Web 2.0 and ergonomics. A feedback interface is overlaid on the document, featuring a "Launch feedback" button and a "Refuse this feedback" button. The "Refuse this feedback" button is highlighted, and a tooltip below it says "Go back over your refusal of this feedback".

# 39 Contradiction dans *PenSum*

## Reject du *feedback* hors-sujet

The screenshot shows a web browser window titled "PenSum". The main content area displays a document with the following sections:

- Read carefully the documents of the course and then sum them up below.**
- COURSE**
  - AJAX**: JavaScript and XML have been around nearly as long as HTML, and many sites have been using both to good effect. As Paul Graham said, "Basically, what 'Ajax' means is 'Javascript now works.'" This means that Web applications can now be created that interact with the users the way we have become accustomed with desktop applications. *The lag and delay caused by server calls is mitigated by a lot of that information being sent asynchronously through JavaScript calls.*
  - Democracy and Community**: These have both been around in limited formats for a long time on the Web. *What makes them exciting is when they are combined with the applications that are possible with AJAX and more modern browsers. Web pages gave people the ability to express themselves like no other time in history. And Weblogs made an easy thing even easier. When Blogger was first set up, they thought of themselves as just an easy-to-use FTP client for Web pages. Now they allow many many people to get online and express their views.*
  - But it's more than just getting content on-line. With Web 2.0 type sites, it is possible to combine data from various sources, even sources that you don't own control or even exist, and turn that into data that people can use. For example, an idea for a Web 2.0 site might be to combine a Flickr tagged gallery of restaurants with actual reviews that people have written about those restaurants and a connection to Google Maps to find the restaurants in your area. The Web site which put this up might not have any of the data for the reviews, none of the photos from Flickr, and only an interface to allow people to comment where they are or where they are visiting and either read reviews of local restaurants.*
- SYNTHESIS**
  - Composition mode**: If you surf the Web design is a scientific discipline that seeks human biology, psychology, performance, whilst protecting and other aspects of our daily life. This is the new and revolutionary dreamed possible 10 years ago. Web 2.0 was previously used
  - Feedback mode**:
    - Feedback**: ergonomics aims to develop and apply knowledge and technology to design products and environments that optimize human factors, is the main. Drawing upon it optimize system across work, leisure in ways we never
    - Refuse this feedback**
    - Refuse this feedback (and link this sentence to a text)**
    - Go back over your refusal of this feedback and delete links**

At the bottom right of the document, there is a "Launch feedback" button and a "Feedback sources" dropdown menu.

*Processus cognitifs en  
éducation et LSA:*

*Ce qui est fait et ce qui peut être fait,  
Section reprise de [Dessus 09]*

# Fonctionnement (très) simplifié

41

Corpus de  
domaines

Productions  
d'apprenants

Productions  
d'enseignants

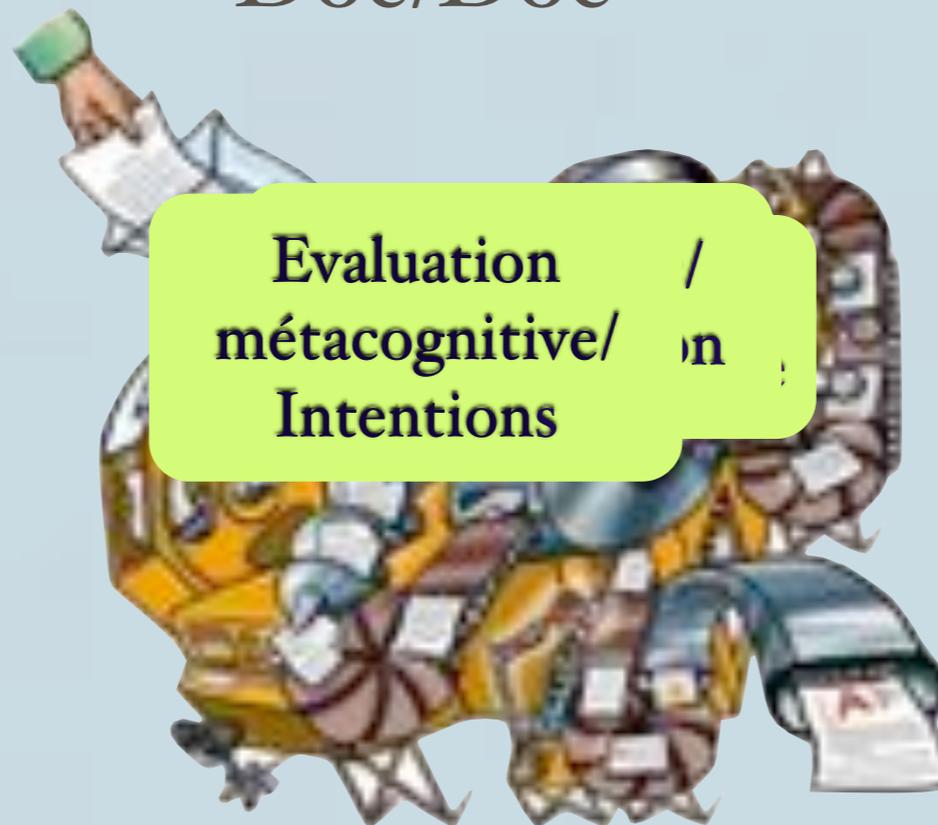
Stimuli de bas  
niveau

Comparaisons

Mot/Mot

Mot/Doc

Doc/Doc



3. Output

1. Contexte &  
Input

2. Type de comparaison

# I. Contexte & Activités

## ENSEIGNER

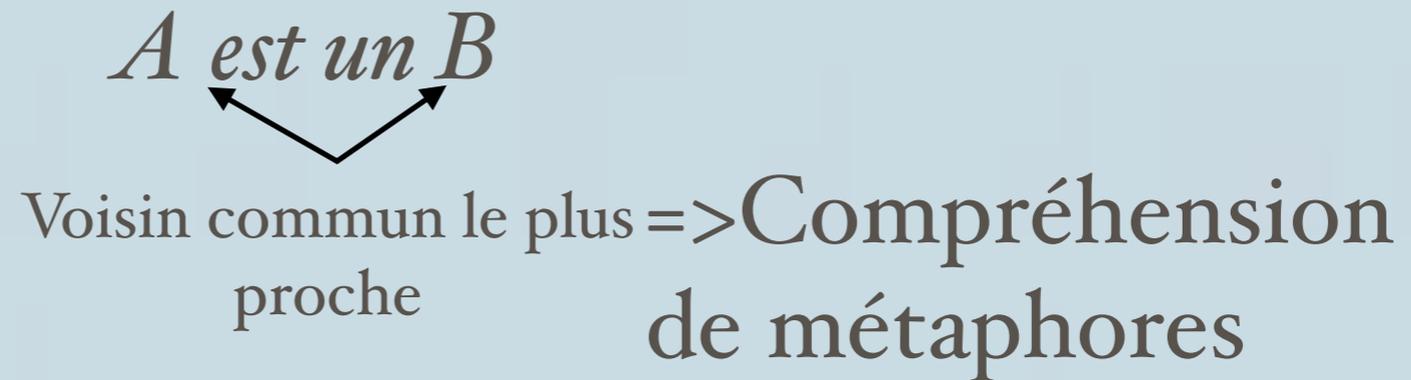
## APPRENDRE

- ❖ Sélectionner le cours approprié
- ❖ Inférer la compréhension d'un El.
- ❖ Représenter un contenu dans carte
- ❖ Evaluer la production d'un El.
- ❖ Détection de plagiat
- ❖ Allouer des El. à un groupe
- ❖ Analyser les actions d'un Ens.
- ❖ Détecter les intentions d'un El., etc.

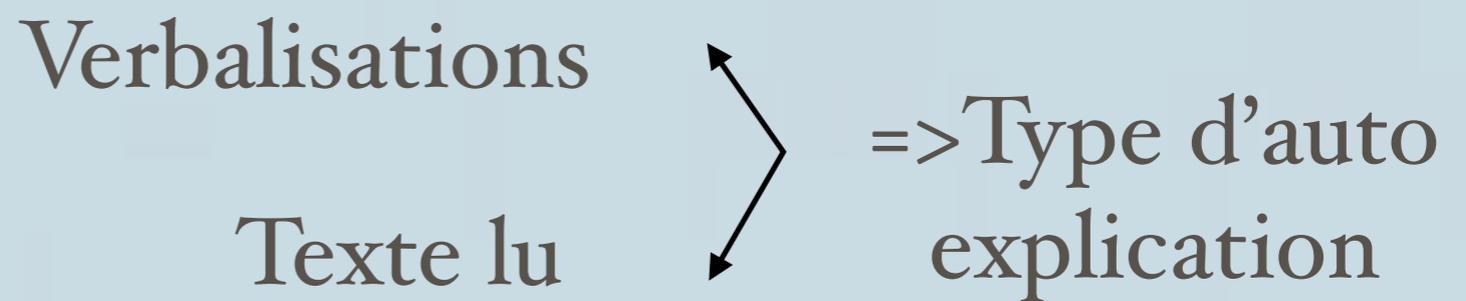
- Sélectionner un texte à lire
- Comprendre tâches et textes
- Résumer ou faire des plans
- Être évalué
- Chercher de l'aide des pairs
- Réfléchir sur sa production
- etc.

## 2. Type de Comparaisons

Compréhension  
de métaphores  
[Kintsch '00]



Auto-explications  
pendant la lecture  
avec *iStart*  
[McNamara *et al.* '04]



## 3. Output Rangé par complexité

- ❖ Sélection ou production de texte
  - ❖ Recherche et génération de textes (*extracts*)
- ❖ Evaluation de textes produits (e.g., essais, dissertations)
  - ❖ comparer des textes pour les évaluer (e.g., *gold standard*)
- ❖ Evaluation des connaissances ou compréhension
  - ❖ inférer ce qui est appris à partir d'une production de textes
- ❖ Evaluation de processus méta-cognitifs
  - ❖ coder et comparer des actions avec des verbalisations

# Applications possibles [Dessus & Trausan-Matu 10]

45

## ❖ “Qu’est-ce que j’ai vraiment à faire ?”

Verbalisation de la tâche

Texte de consigne

=> compréhension de la tâche

## ❖ Une analyse bakhtinienne des voix dans l’ens. à distance

Manuel

Propos de l’enseignant

Productions des élèves

Evaluation de l’enseignant

Discussions collaboratives



=> Construction de connaissances

# Merci de votre attention !

- ❖ *Résum'Web* a été partiellement financé par un projet «Ecole et Sciences cognitives» du MENRT (dir. G. Denhière);
- ❖ *Pensum* a été financé par le projet européen  (7e PCRD, ICT-STREP, dir. W. Greller)
- ❖ Les réfs de cette présentation sont disponibles à <http://www.citeulike.org/user/pdessus/tag/indlang> et elle sera prochainement sur le site du cycle de conférences
- ❖ Merci à Damien Dupré, Benoît Lemaire, Mathieu Loiseau, Sonia Mandin, Emmanuelle Villiot-Leclercq & Virginie Zampa qui ont collaboré aux travaux présentés ici.