

- Lemaire, B., Dessus, P., & Zampa, V. (2001, 18 janvier). *LSA (Latent Semantic Analysis) : Représentations sémantiques, modélisations cognitives de l'apprentissage et applications aux environnements d'apprentissage avec ordinateur*. Grenoble : Univ. P.-Mendès-France, LSE, journée d'étude.

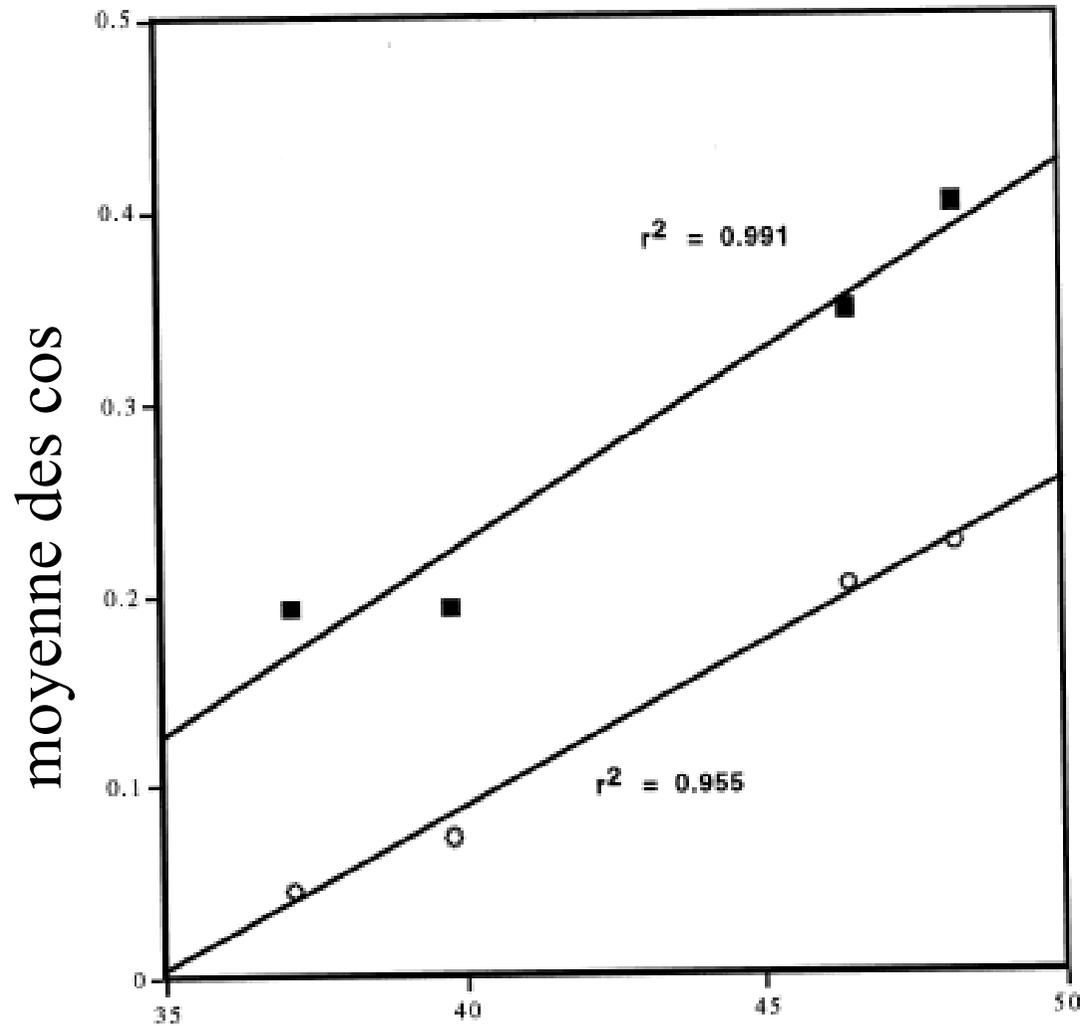
## II. Représenter et acquérir des connaissances à partir de textes

- Outils basés sur la représ. des connaissances
  - LSA prédit la cohérence de textes
  - LSA évalue des copies d'étudiants
  - LSA analyse du contenu
  - LSA prédit l'ordonnance de cours
- Modélisations cognitives
  - LSA simule l'apprentissage d'étudiants
  - LSA passe des QCM
  - LSA simule l'acquisition de vocabulaire
  - LSA simule la compréhension de textes

## II.1.a LSA prédit la cohérence de textes (Foltz *et al.* 98)

- Le degré de cohérence entre 2 phrases adjacentes d'un texte est lié à leur proximité sémantique, mesurée par LSA
- Retraitement avec LSA d'études avec humains, étudiant le lien entre compréhension de textes et leur cohérence
- La moyenne des proximités des phrases adjacentes croît avec la cohérence des textes
- La cohérence mesurée croît avec la compréhension des sujets (mesurée dans une étude précédente)

## II.1.b LSA prédit la cohérence de textes



score test de compréhension

■ LSA Ency preds  
○ Word overlap

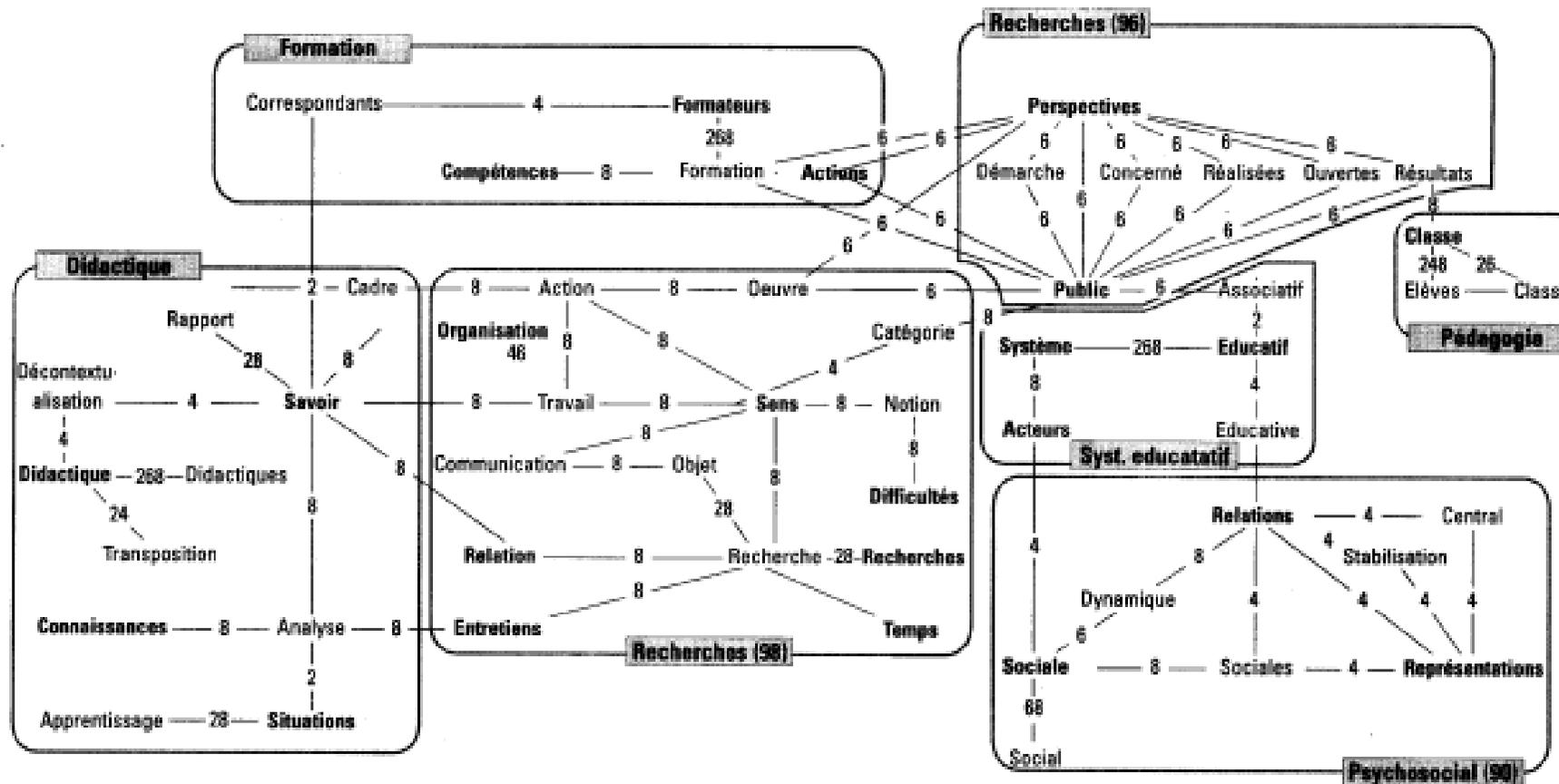
## II.3 LSA évalue des copies d'étudiants

- On utilise la capacité de LSA à représenter les connaissances pour évaluer celle contenue dans des textes d'étudiants, en la comparant avec des textes-cibles :
  - cours, textes de références
  - copies standard notées par des humains

## II.4.a LSA analyse du contenu (Dessus, 1999a)

- 4 corpus de résumés de communications aux Biennales de l'éduc. et de la form. (92 à 00)
- Examen des fluctuations de centralité des mots les plus centraux
  - Permet de détecter des années à fortes fluctuations
  - Permet de dresser des cartes de mots selon leur voisinage

# II.4.b LSA analyse des corpus



## II.2.a LSA simule l'apprentissage d'étudiants (Wolfe *et al.* 98)

- Simuler l'apprentissage à partir de l'adéquation entre la difficulté de textes et les connaissances préalables des étudiants

Pré-Essai

Pré-Test

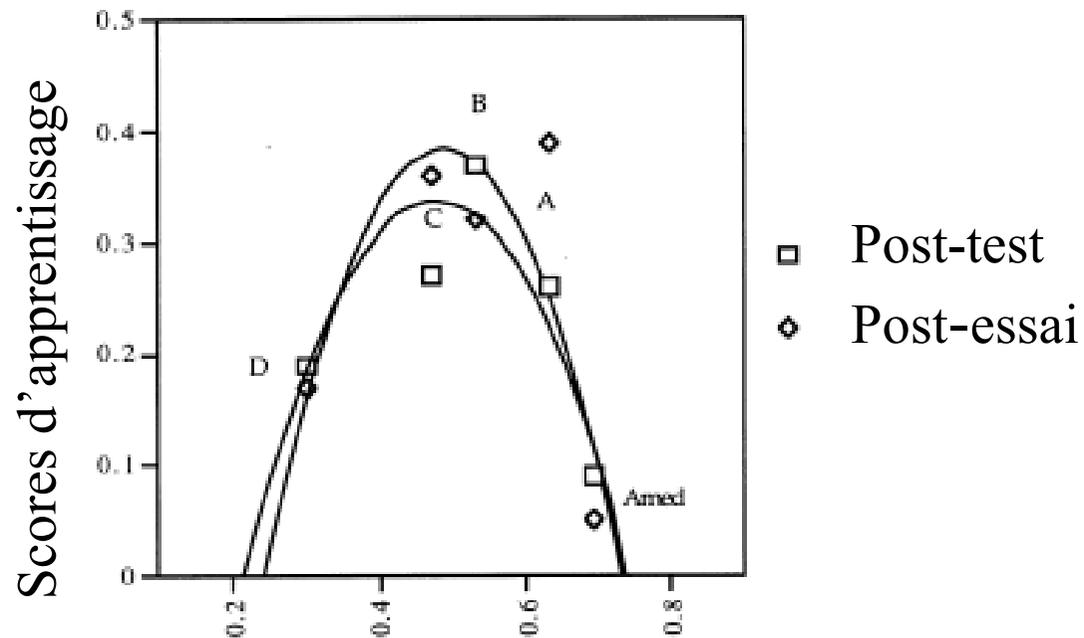
Lecture texte  
A,B,C,D

Post-Test

Post-Essai

1. LSA peut évaluer la difficulté des textes ABCDE
2. Lien entre la note LSA du pré-essai et des pré-tests ( $r = .63$  à  $.74$ )
3. LSA peut évaluer le gain pré-post essai et proposer des textes en conséquence

## II.2.b LSA simule l'apprentissage d'étudiants



cos. entre pré-essai et texte lu  
(A, B, C ou D)

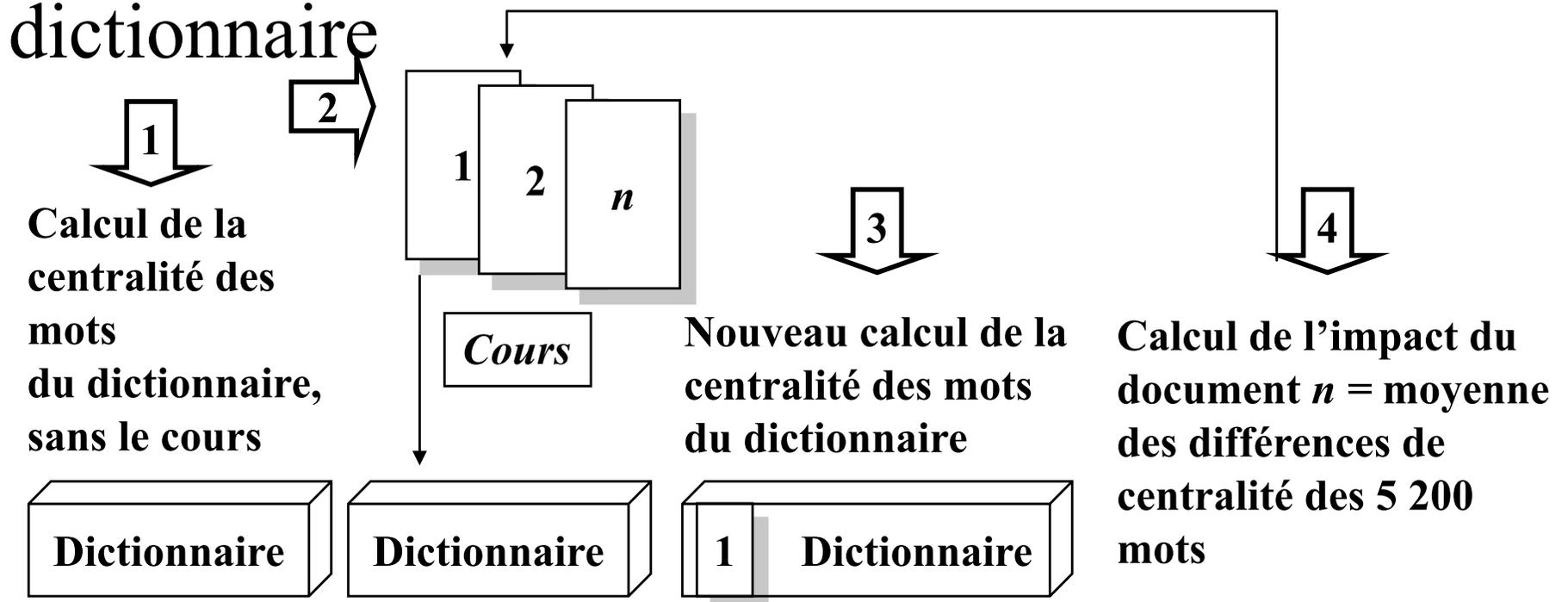
# LSA passe des QCM

(Dessus, à paraître)

- LSA peut-il répondre à un QCM lié à un cours, avec les seules connaissances de ce cours ?
  - LSA traite le cours, et on lui fait répondre au QCM. Note obtenue : 12/27 ou 44,4 % de bonnes réponses ; moy. des étudiants : 18,8/27
  - Les bonnes réponses de LSA sont assez liées à celles des étudiants ( $r = 0,50$ ) ; ses erreurs, en revanche, sont indépendantes de celles des étudiants

## II.5. LSA ordonne un cours

- Cours d'astronomie (43 docs) et dictionnaire
- Liaison moyenne ( $\rho = .4$  ;  $p < .005$ ) entre le rang du document et son « impact » dans le dictionnaire



# LSA passe le TOEFL

- BL ?

### III. Tuteurs intelligents basés sur LSA

- Tuteurs analysant les interactions
  - Autotutor (Wiemer-Hastings et *al.*, 99)
  - ELSA (Epistemological LSA, Teplovs & van Aalst, 98)
- Évaluation de copies
  - *Intelligent Essay Assessor* (Foltz et *al.*, 99)
  - *Summary Street* (E. Kintsch, à paraître)
  - *Select-a-Kibitzer* (Wiemer-Hastings & Graesser)
  - *Apex*

# III. Évaluer automatiquement des productions écrites

- Par QCM
- Par l'analyse de critères de surface
- Par une analyse sémantique : p. ex. avec LSA

# Autotutor (Wiemer-Hastings *et al.* 99)

- Tuteur intelligent pour l'apprentissage de l'informatique dont un module utilise LSA
  - Dialogue avec les élèves
  - Modèle de l'élève : qualité de ses réponses, niveau dans le domaine, niveau dans le sujet couvert
- Problème principal : longueur des phrases traitées

# Select-a-Kibitzer

(Wiemer-Hastings & Graesser, à paraître)



# Structure type d'un cours dans Apex

- Un cours est un texte d'un domaine ( $\pm 30$  p.), écrit par un enseignant, structuré en 2 niveaux
  - Sujets ( $\pm 3$  pages)
  - Notions ( $\pm 1$  paragraphe)
- Les étudiants choisissent un Sujet, et composent librement sur ce sujet en se guidant avec les Notions
- Les enseignants peuvent composer des sujets d'examen, en spécifiant des Notions appartenant à des Sujets différents.

# Fonctionnalités d'Apex (Ens./Étudiants)

## Enseignant

- 
- Écrire des cours et des questions
  - Maintenir la base de données
    - cours, enseignants, étudiants
  - Visualiser l'activité des étudiants
    - copies, notes

## Étudiant

- Choix d'un cours, d'un Sujet
  - (Lecture du cours)
  - Écriture (ou correction) sur ce sujet
  - Demander une évaluation
    - du contenu
    - du plan
    - de la cohérence
-

Navigation icons: Home, Back, Forward, Stop, Reload, Print, Mail, Chat, Help, etc.

Go To: <http://shs-doiss7/EcrireDissertEtu2.php>

AltaVista FAST Search Education-line IUFM Grenoble ABSYS WEB Designer Student

# Apex Assessor

## Notions of Topic 9 : Le climat de la classe

Définition de la notion de climat

L'écologie

Le milieu

Le système social

La culture

**Please write your essay below :**

EDMONDS dans ses synthèses des travaux sur l'efficacité des écoles a mis à jour cinq facteurs associés à de meilleures performances des écoles. Ce "modèle des cinq facteurs" d'EDMONDS aura un gros retentissement et sera à l'origine de la mise en place de plusieurs programmes d'amélioration des écoles (projets RISE, SIP et LSDP notamment). Ces cinq facteurs sont les suivants :

- une forte direction ("leadership") ;
- des attentes élevées concernant les performances des élèves ;
- un climat discipliné, sans toutefois être rigide;
- un fort accent mis sur l'enseignement des savoirs de base (lecture, écriture, mathématiques) ;
- des évaluations et des contrôles fréquents des progrès des élèves.

L'école est conçue comme un système social qui, en tant que tel, a un fonctionnement spécifique, développe un système particulier de

GENERAL GRADE: 8.8/20

The following Notions were covered very poorly:

- l'écologie (0.03)

The following Notions were covered poorly:

- le système social (0.34)

The following Notions were covered well:

- définition de la notion de climat (0.63)
- le milieu (0.58)
- la culture (0.61)

# Procédure des évaluations

- *Contenu* : Similarité entre la copie entière et chaque Notion du Sujet-cible. La moyenne générale est affichée, ainsi que la note Notion par Notion
- *Plan* : Similarité entre chaque § de la copie et chaque Notion. Le titre de la Notion la plus proche est affichée comme item de plan
- *Cohérence* : La similarité entre 2 phrases adjacentes est liée à leur cohérence. Affichage des sauts de cohérence

# Deux tests d'Apex

(1) 31 copies d'examen ont été notées par LSA, et comparées avec leurs notes "humaines" :  $r = 0.59$

(2) 3 groupes d'étudiants écrivent une copie :  
a) *on-line* b) à la demande c) trait. textes

Aucune différence significative intergroupes

## IV. Discussion : questions posées

- Taille du corpus (Perfetti 98) ?
- Nature et taille de l'éventuel corpus annexe ?
- Que mesure vraiment la centralité d'un mot ?  
Sa typicalité ou la quantité de connaissance contenue dans le corpus ?
- Problème de directionalité (Rehder *et al.* 98)
  - un vecteur représente de la connaissance, il peut être comparé à un autre *via* cosinus, qui n'indique pas quelle connaissance est supérieure en qualité

## IV. Discussion

- Pour quelles applications LSA est-il suffisant ?
- Où sont les manques ?
- Dépendance au langage ? LSA peut-il modéliser d'autres formes d'apprentissage ?