

LEARNING TO SUMMARIZE & SUMMARIZING FOR LEARNING

SYMPOSIUM AROUND THE INAUGURAL ADDRESS
OF PETER SLOEP
HEERLEN, THE NETHERLANDS
14/11/2008

PHILIPPE DESSUS

LABORATOIRE DES SCIENCES DE L'ÉDUCATION, UPMF,
GRENOBLE, FRANCE

PHILIPPE.DESSUS@UPMF-GRENOBLE.FR



OUTLINE AND CREDITS

- I. First Example: Apex 1 & how it works
 - II. Some in-between Literature
 - III. Second Example: Apex 2 & how it works
 - IV. Third Example: Resum'Web
 - V. How do they all work in Educational Settings?
 - VI. Brainstorming Session: How Discover New Applications?
- ▶ Collaborative Work with Benoît Lemaire, Sonia Mandin & Virginie Zampa

I. A FIRST EXAMPLE

PRESENTATION OF APEX 1,
A SYSTEM THAT HELPS REVISE COURSES

I. Apex 1

II. Literature

III. Apex 2

IV. Resum'web

V. Do they work

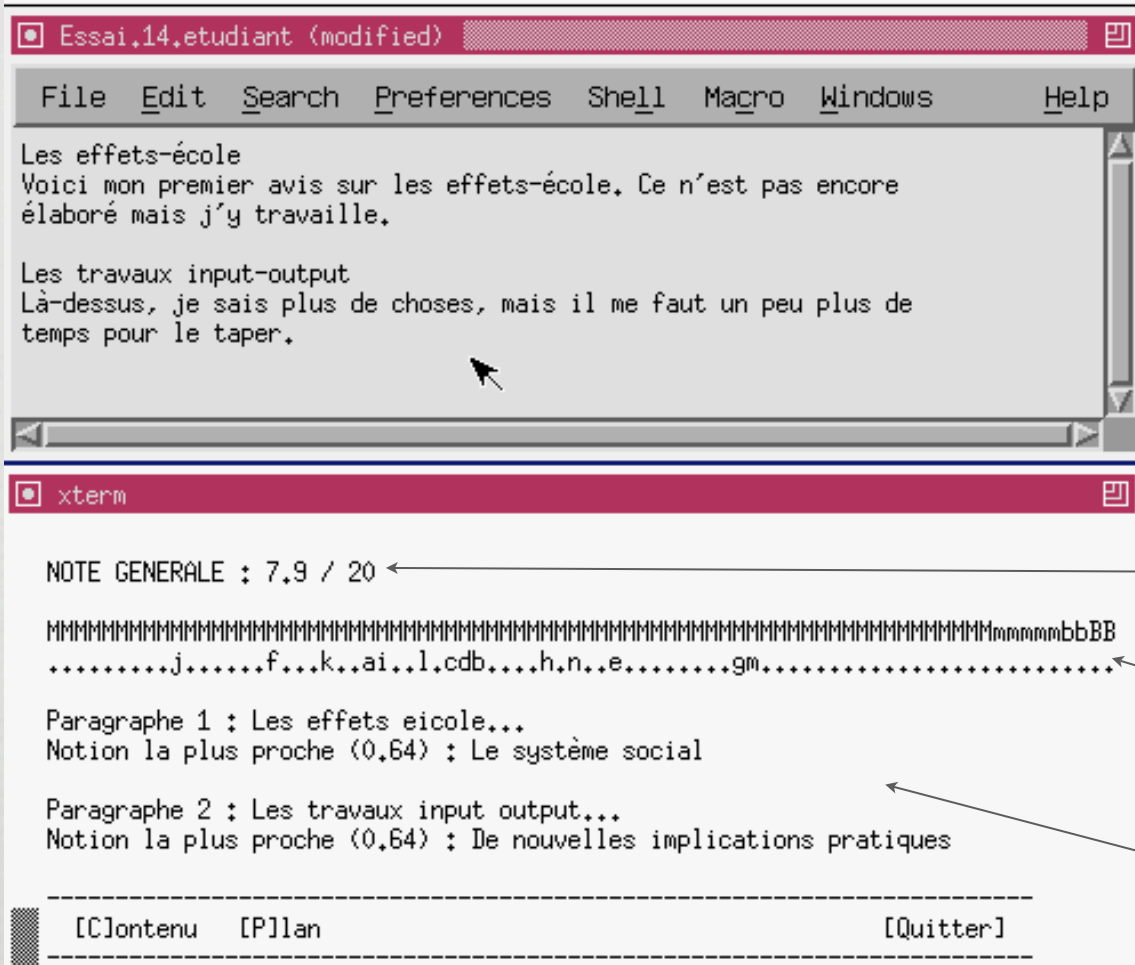
VI. Brainstorm

I. CONTEXT

- Students usually have to revise their courses...
 - a good way to do that is to summarize them...
 - ... but they seldom benefit from *content*-based assessments, cognitively demanding for the teacher
 - ... and current ICT only deliver assessment on shallow features
- Would ICT be an adequate way to *freely* allow students:
 - to write out course summaries for understanding course notions?
 - to be assessed on higher-level features?

- I. Apex 1
- II. Littérature
- III. Apex 2
- IV. Resum'web
- V. Do they work
- VI. Brainstorm

I. FIRST EXAMPLE: APEX 1



Summary Pane:
Written Course
Summary

Feedback Pane:
Overall Grade
How well each topic
is covered
Possible followed
outline

I. Apex 1

II. Literature

III. Apex 2

IV. Resum'web

V. Do they work

VI. Brainstorm

I. HOW DOES APEX 1 WORK?

- Beforehand
 - teacher processing : All the course is two-level tagged (topic/ notions)
 - *Latent Semantic Analysis* processing : course + “general language” corpus compiled
- On-line and as many times as required
 - course summary semantically compared with each of the course notions

I. Apex 1

II. Literature

III. Apex 2

IV. Resum'web

V. Do they work

VI. Brainstorm

LATENT SEMANTIC ANALYSIS : A PRIMER 1/2 (LEMAIRE, 2005)

- LSA determines the statistical context in which each word occurs; semantically compares words; serves as semantic memory
 - two words are similar if they occur in ~~same~~ paragraphs
 - two paragraphs are similar if they contain ~~common~~ words
 - two words are similar if they occur in **similar** paragraphs
 - two paragraphs are similar if they contain **similar** words

I. Apex 1

II. Literature

III. Apex 2

IV. Resum'web

V. Do they work

VI. Brainstorm

LATENT SEMANTIC ANALYSIS: A PRIMER 2/2 (LEMAIRE 2005)

- Given a corpus processed beforehand
 - splited in paragraphs
 - words are projected in a n -dimension space so that
 - words can be compared to each other by computing the cosine of their corresponding vectors
 - paragraphs can be compared to each other by computing the cosine of the sum vector of the words they are composed of
- Requires huge corpora to acquire “common ground knowledge”
- Doesn't capture all features (negations, style, etc.)

I. Apex 1

II. Literature

III. Apex 2

IV. Resum'web

V. Do they work

VI. Brainstorm

WHAT'S LACKING?

- Involving the student in a more complete workflow
 - Revise courses also consists in reading them
- Focussing on processes rather than products
 - metacognitive judgements on course comprehension
 - model the summarization process
- Stop! More literature is needed!

II. SOME IN-BETWEEN LITERATURE REVIEW

EXISTING SYSTEMS; SUMMARIZING <-> LEARNING;
METACOGNITIVE ACCOUNT

I. Apex 1

II. Literature

III. Apex 2

IV. Resum'web

V. Do they work

VI. Brainstorm

CURRENT FORMS OF TEXT ASSESSMENT

- Students already benefit from several computer-based ways of assessment
 - *very shallow*: text length, spelling, grammar, etc.
 - *mid-shallow*: readability
 - *multiple choice questions*: both plausible and inadequate??
- The semantic level is promising...

I. Apex 1

II. Literature

III. Apex 2

IV. Resum'web

V. Do they work

VI. Brainstorm

SEMANTIC ASSESSMENT LEVEL: SOME SYSTEMS

- *Intelligent Essay Assessor* (Foltz et al., 1999): LSA-based comparison with gold standard essays
- *Summary Street* (E. Kintsch et al., 2007): holistic grade, the score is this of the closest pre-graded essay (by a human)
- *Select-A-Kibitzer* (Wiemer-Hastings & Graesser, 2000): idem as previous, but with agents

I. Apex 1

II. Literature

III. Apex 2

IV. Resum'web

V. Do they work

VI. Brainstorm

IS SOMETHING MISSING?

- Prescriptive view of assessment: mostly depends on human judgment (comparison with pre-graded essays)
- Mostly feedback on products (e.g., number of words, cohesion, etc.) rather on processes. Necessity to have a closer look on activities
- Metacognition (how students judge themselves) seldom taken into account
- Notion of environment: students freely engaged in activity loops
 - *reading loop* (acquiring input, building knowledge)
 - *writing loop* (producing/reorganizing ideas)

I. Apex 1

II. Literature

III. Apex 2

IV. Resum'web

V. Do they work

VI. Brainstorm

LINKS BETWEEN SUMMARIZING–LEARNING

- Understanding necessary for summarizing
 - Text macrostructure recall similar to a summary of this text
- Summarizing fosters understanding (Bean & Steenwyk, 1984; Stevens, 1988)
 - strengthen connections between source text ideas and knowledge
 - helps memorizing the source text content
 - helps monitoring comprehension (Thiede & Anderson, 2003)

I. Apex 1

II. Literature



III. Apex 2

IV. Resum'web

V. Do they work

VI. Brainstorm

THE PROCESS OF SUMMARIZING

- Two main interleaved processes, to be simulated
 - selection of the main (most important) ideas of the source text
 - macrorule application on some sentences of the source text (ST) to compose the summary (Sum) (W. Kintsch & van Dijk, 1978)
 - *copy*: a Sum sentence is very close to a ST one
 - *deletion*: a ST sentence is very far to all Sum ones 
 - *generalization*: a Sum sentence is close to several ST sentences
 - *construction*: a Sum sentence is not very close to all ST one, but pretty close to some 

I. Apex 1

II. Literature

III. Apex 2

IV. Resum'web

V. Do they work

VI. Brainstorm

A METACOGNITIVE ACCOUNT

- Metacognitive processes can be triggered by feedback. Student's...
 - awareness of strategies and goals
 - I'll begin to read the text then I'll summarize it
 - ability to assess their own understanding
 - I think I grasped the notion X
 - ability to compare their assessment to that of the computer
 - you said you understood X but obviously you didn't
 - you said notion X is important but you deleted it

III. SECOND EXAMPLE: APEX 2

AN ILE THAT HELPS EXPLORING & LEARNING CONTENT

- I. Apex 1
 - II. Literature
 - III. Apex 2
 - IV. Resum'web
 - V. Do they work
 - VI. Brainstorm
-

III. APEX 2

- An Interactive Learning Environment
- 2 loops: Reading & Writing in which the student is freely engaged
- Written in PHP 5 with C routines, on top of LSA
- Demo

- I. Apex 1
 - II. Literature
 - III. Apex 2
 - IV. Resum'web
 - V. Do they work
 - VI. Brainstorm
-

WHAT'S LACKING YET?

- Assessment on products rather than processes
- Relying on a cognitive model of summarizing/reading

IV. THIRD APPLICATION: *RESUM'WEB*

A TUTOR THAT HELPS SUMMARIZING

DEVELOPED AS PART OF THE PHD THESIS OF SONIA MANDIN

- I. Apex 1
- II. Literature
- III. Apex 2
- IV. Resum'web
- V. Do they work
- VI. Brainstorm

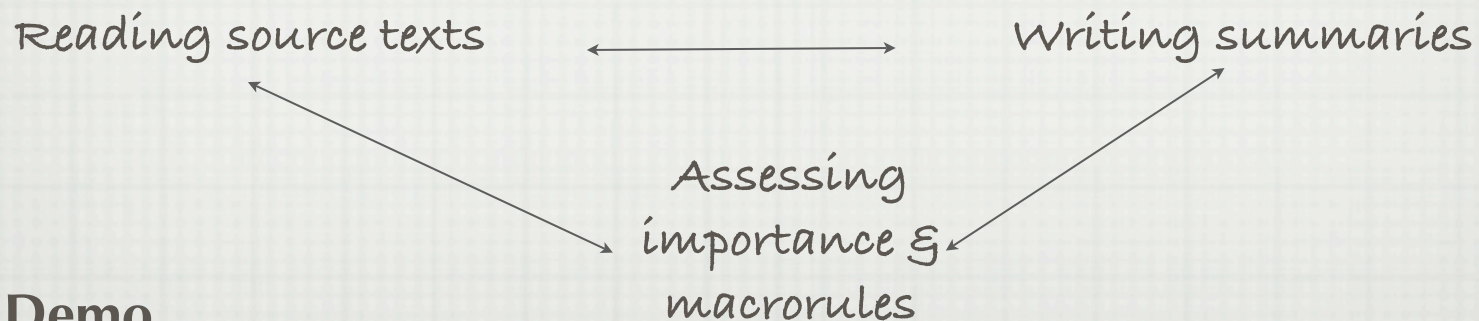
RESUM'WEB: HOW IT (VERY BRIEFLY) WORKS

- Main sentences identification process
 - an important sentence of a text is highly semantically connected with other sentences of this text (LSA comparison)
- Macrorule identification process
 - each sentence of the summary is semantically compared with all the source text sentences
 - 3 similarity thresholds empirically determined (too far, adequately close, too close)
 - macrorules determined accordingly (see slide # 15)

- I. Apex 1
- II. Literature
- III. Apex 2
- IV. Resum'web
- V. Do they work
- VI. Brainstorm

RESUM'WEB ARCHITECTURE

- A set of resources: courses on summarization
- A set of texts to be summarized
- A 2,5 (0,5: assessment activity) loops involvement




- Demo

V. DOES THIS STUFF WORK ANYMORE?

ASSESSMENT IN EDUCATIONAL SETTINGS

APEX / TEST IN CONTEXT

- I. Apex 1
- II. Literature
- III. Apex 2
- IV. Resum'web
- V. Do they work
- VI. Brainstorm

- 21 real-settings exam essays processed by *Apex 1*
- Pretty high correlation ($r = .59$) teacher vs. *Apex 1* grades
- *Apex 2* currently under tests ( project)

- I. Apex 1
- II. Literature
- III. Apex 2
- IV. Resum'web
- V. Do they work
- VI. Brainstorm

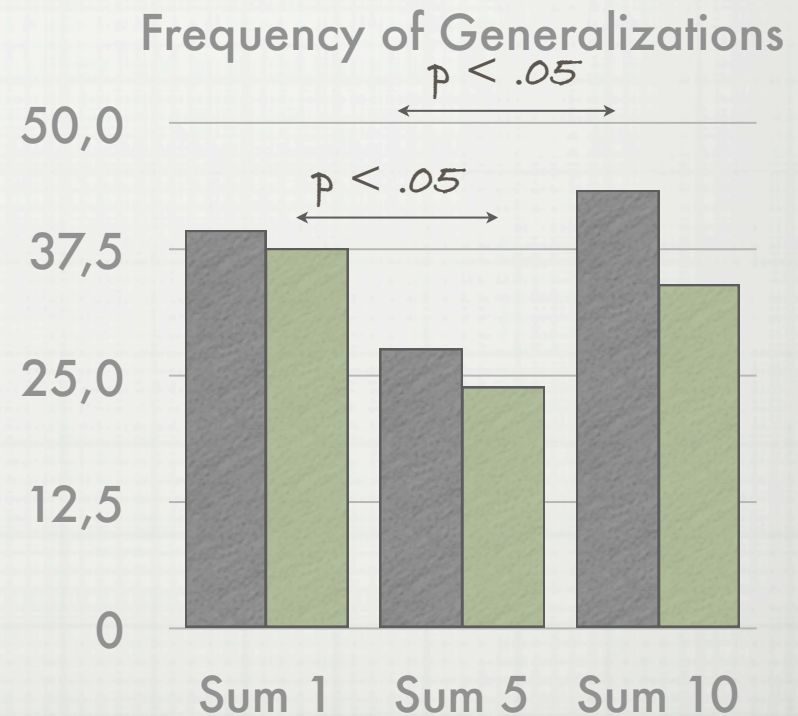
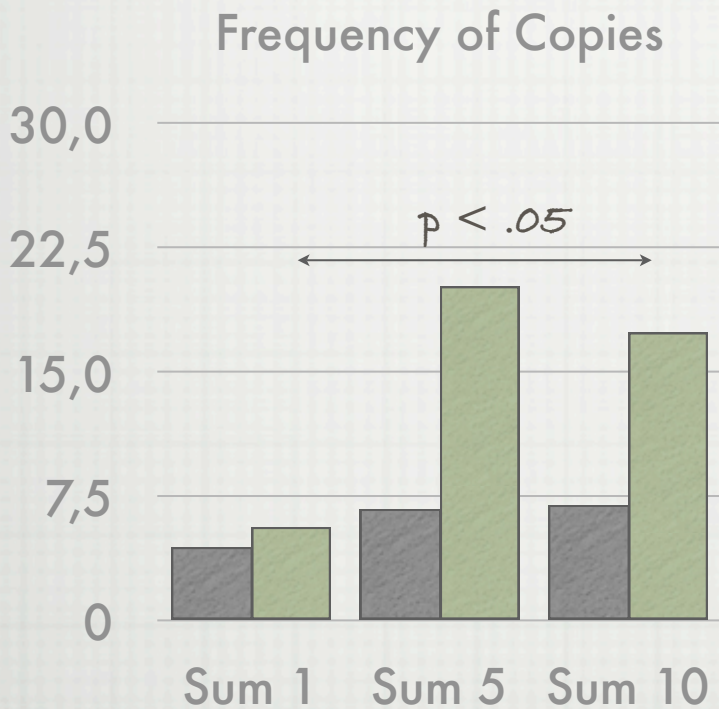
RESUM'WEB: RESULTS IN CONTEXT (MANDIN, IN PREP)

- 32 Highschool students are given 10 texts to summary (1 per session)
 - 18 with the help of the "full" *Resum'Web*
 - 14 in a control group, with main ideas identification but without macrorule (self-)assessment (placebo task)
- Focus on progress for summaries 1, 5 and 10 (human assessment)
 - macrorules used
 - overall grades

- I. Apex 1
- II. Literature
- III. Apex 2
- IV. Resum'web
- V. Do they work
- VI. Brainstorm

EFFECT OF RESUM'WEB ON STUDENT'S ACTIVITY

■ Resum'Web ■ Control



Frequency of "off-the-subject" and constructions were equal
 Summary grades by teachers were equal

I. Apex 1

II. Literature

III. Apex 2

IV. Resum'web

V. Do they work

VI. Brainstorm

FURTHER RESEARCH PROJECTS

- Focus on

- adding a chatting loop (interact with peers, discuss, argument)
- scrutinizing differences on summarizing narrative vs. expositive texts
- integrating more specific ways of modelling understanding (e.g., using the Construction-Integration model, Kintsch 1988; Lemaire et al., 2006)
- modelling students' ZPD (Zampa & Lemaire, 2002)

VI. BRAINSTORMING ZONE

WHAT TO DO NEXT? HOW TO DISCOVER NEW APPLICATIONS?

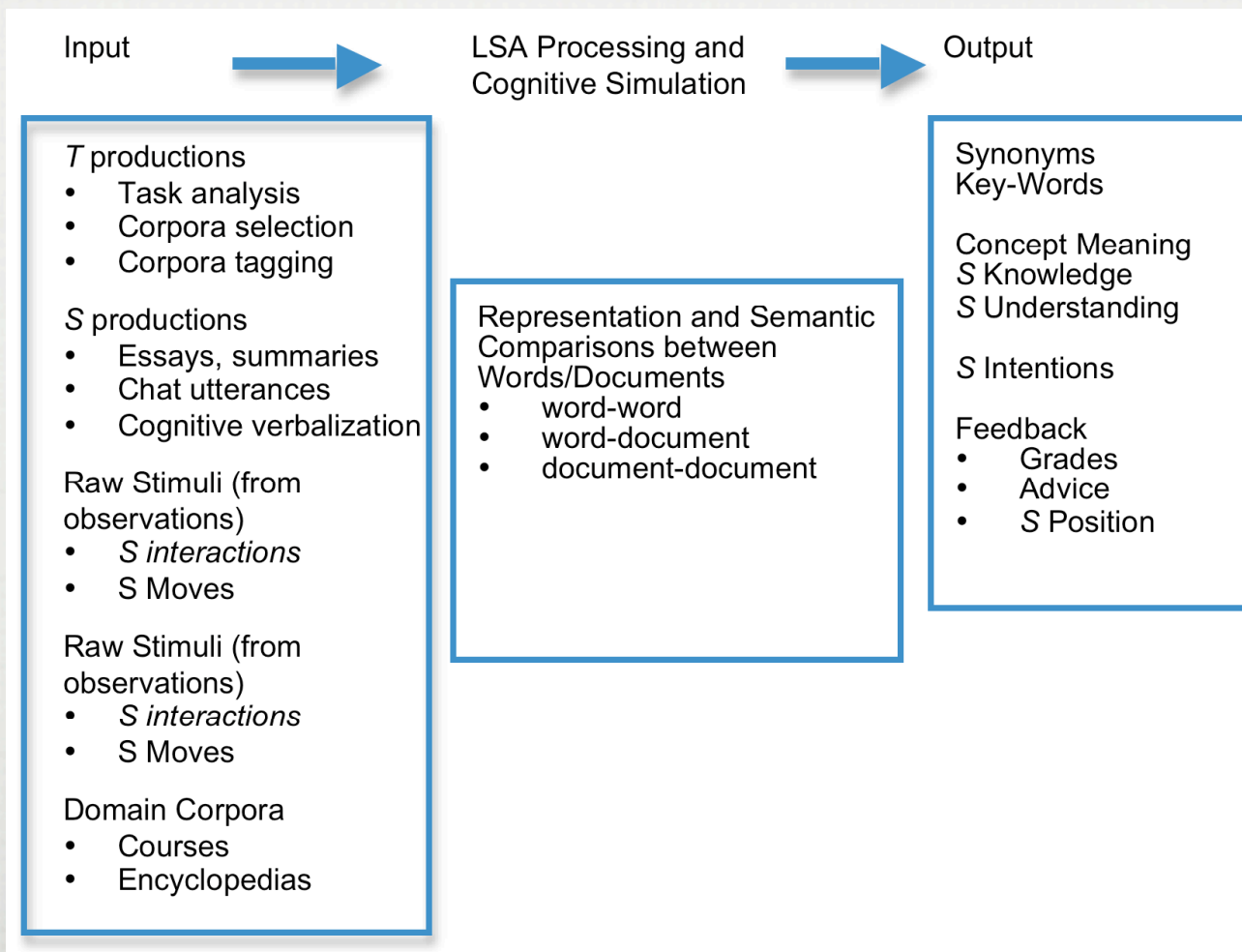
- I. Apex 1
- II. Literature
- III. Apex 2
- IV. Resum'web
- V. Do they work
- VI. **Brainstorm**

OVERVIEW OF THE EDUCATIONAL USES OF LSA

- *Text selection and production*: plagiarism detection, semantic-based text search, keyword selection
- *Essay assessment*: text readability, essay grading, outline analysis, macrorule analysis
- *Understanding assessment*: metaphor comprehension, matching texts to readers wrt. pre-knowledge, simulation of understanding, learner positioning wrt. knowledge
- *Self-regulation processes assessment*: self-regulated learning analysis, self-regulative vs. system assessment comparison, intentions uncovering

- I. Apex 1
- II. Literature
- III. Apex 2
- IV. Resum'web
- V. Do they work
- VI. Brainstorm

HOW TO FIND NEW LSA APPLICATIONS



- I. Apex 1
- II. Literature
- III. Apex 2
- IV. Resum'web
- V. Do they work
- VI. Brainstorm

THANKS FOR YOUR ATTENTION!



A
B
C
D

Source: <http://upslinks.net/humor/grading.htm>

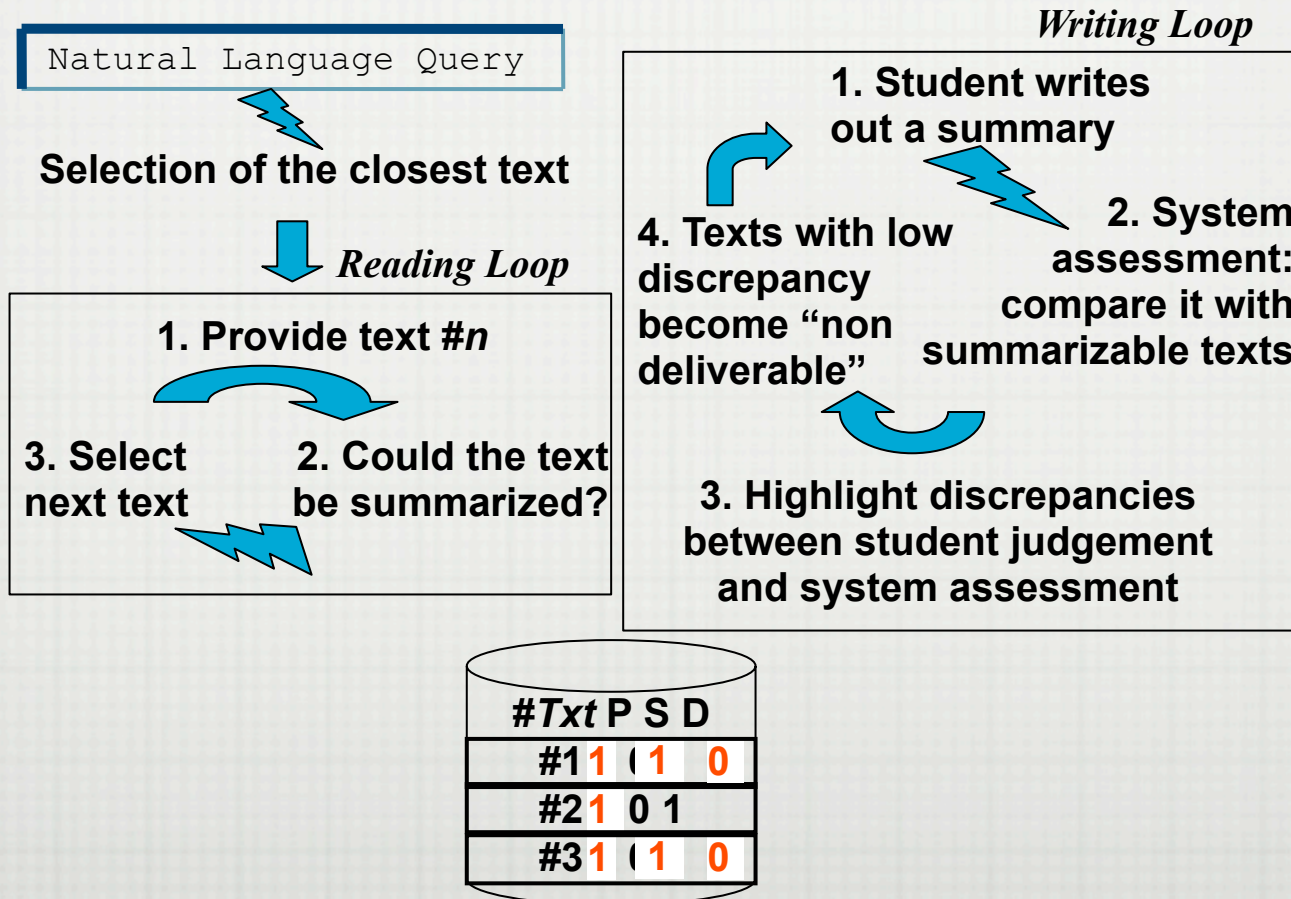
- I. Apex 1
- II. Literature
- III. Apex 2
- IV. Resum'web
- V. Do they work
- VI. Brainstorm

SOME REFERENCES

- Apex 1: Lemaire, B. & Dessus, P. (2001). A system to assess the semantic content of student essays. *Journal of Educational Computing Research*, 24(3), 305-320.
- Apex 2: Dessus, P. & Lemaire, B. (2002). Using production to assess learning: an ILE that fosters Self-Regulated Learning. In S. A. Cerri, G. Gouardères, & F. Paraguaçu (Eds.), *Intelligent Tutoring Systems (ITS 2002)* (pp. 772-781). Berlin: Springer.
- Resum'Web: Lemaire, B., Mandin, S., Dessus, P. & Denhière, G. (2005). Computational cognitive models of summarization assessment skills. In B. G. Bara, L. Barsalou & M. Bucciarelli (Eds.), *Proceedings of the 27th Annual Conference of the Cognitive Science Society (CogSci' 2005)* (pp. 1266-1271). Mahwah: Erlbaum
- All available on: <http://webu2.upmf-grenoble.fr/sciedu/pdessus/>

- I. Apex 1
- II. Literature
- III. Apex 2
- IV. Resum'web
- V. Do they work
- VI. Brainstorm

APPENDIX I: HOW DOES APEX 2 WORK?



- I. Apex 1
- II. Literature
- III. Apex 2
- IV. Resum'web
- V. Do they work
- VI. Brainstorm

APPENDIX 2: HOW TO FIND NEW COGNITIVE TOOLS

- Interactively specify (after Popper, 1998; Dessus, 2004, 2006)
 - what are the objects at hand or metaphors used (world 1)
 - what are the knowledge artifacts used (world 3)
 - what are the simulated cognitive processes (world 2)

