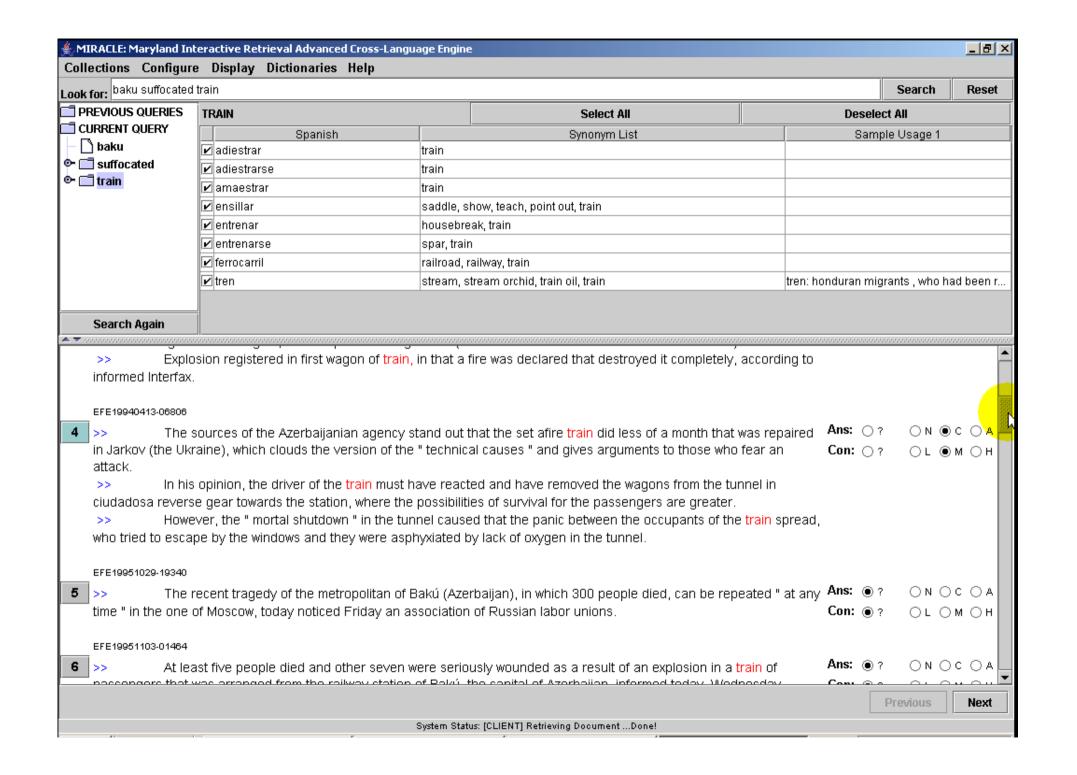# iCLEF 2004 at Maryland: Summarization Design for Interactive Cross-Language QA

Daqing He, Jianqiang Wang, Jun Luo and
Douglas W. Oard
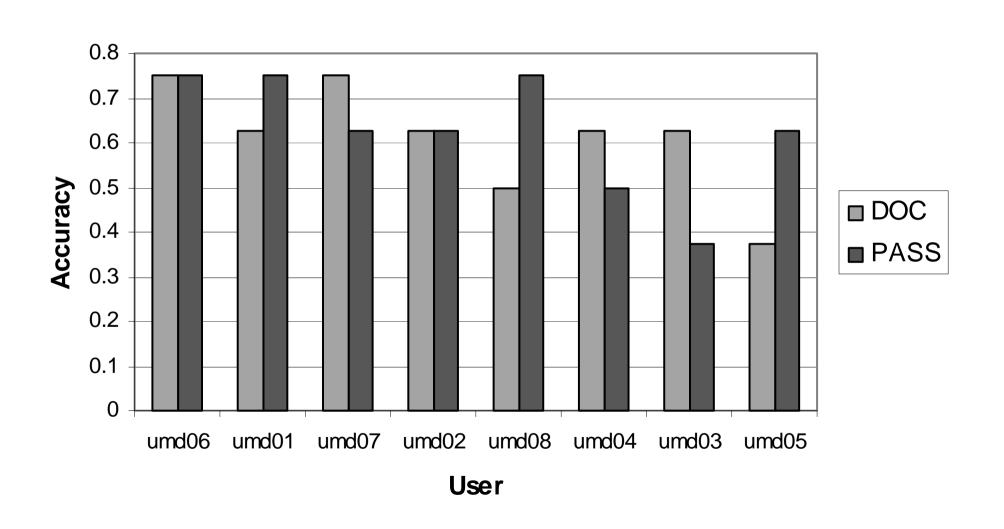
University of Maryland, College Park, MD USA

Collections Configure Display Dictionaries Help

Look for: baku suffocated train   Search   Reset

PREVIOUS QUERIES

CURRENT QUERY
- baku
- suffocated
- train

Search Again

TRAIN	Select All	Deselect All
Spanish	Synonym List	Sample Usage 1
✔ adiestrar	train	
✔ adiestrarse	train	
✔ amaestrar	train	
✔ ensillar	saddle, show, teach, point out, train	
✔ entrenar	housebreak, train	
✔ entrenarse	spar, train	
✔ ferrocarril	railroad, railway, train	
✔ tren	stream, stream orchid, train oil, train	tren: honduran migrants , who had been r...

>>   Explosion registered in first wagon of train, in that a fire was declared that destroyed it completely, according to informed Interfax.

EFE19940413-06806

4 >>   The sources of the Azerbaijanian agency stand out that the set afire train did less of a month that was repaired in Jarkov (the Ukraine), which clouds the version of the " technical causes " and gives arguments to those who fear an attack.
>>   In his opinion, the driver of the train must have reacted and have removed the wagons from the tunnel in ciudadosa reverse gear towards the station, where the possibilities of survival for the passengers are greater.
>>   However, the " mortal shutdown " in the tunnel caused that the panic between the occupants of the train spread, who tried to escape by the windows and they were asphyxiated by lack of oxygen in the tunnel.

Ans: ○ ?   ○ N ◉ C ○ A
Con: ○ ?   ○ L ◉ M ○ H

EFE19951029-19340

5 >>   The recent tragedy of the metropolitan of Bakú (Azerbaijan), in which 300 people died, can be repeated " at any time " in the one of Moscow, today noticed Friday an association of Russian labor unions.

Ans: ◉ ?   ○ N ○ C ○ A
Con: ◉ ?   ○ L ○ M ○ H

EFE19951103-01464

6 >>   At least five people died and other seven were seriously wounded as a result of an explosion in a train of passengers that was arranged from the railway station of Bakú, the capital of Azerbaijan, informed today Wednesday

Ans: ◉ ?   ○ N ○ C ○ A
Con: ◉ ?   ○ L ○ M ○ H

Previous   Next

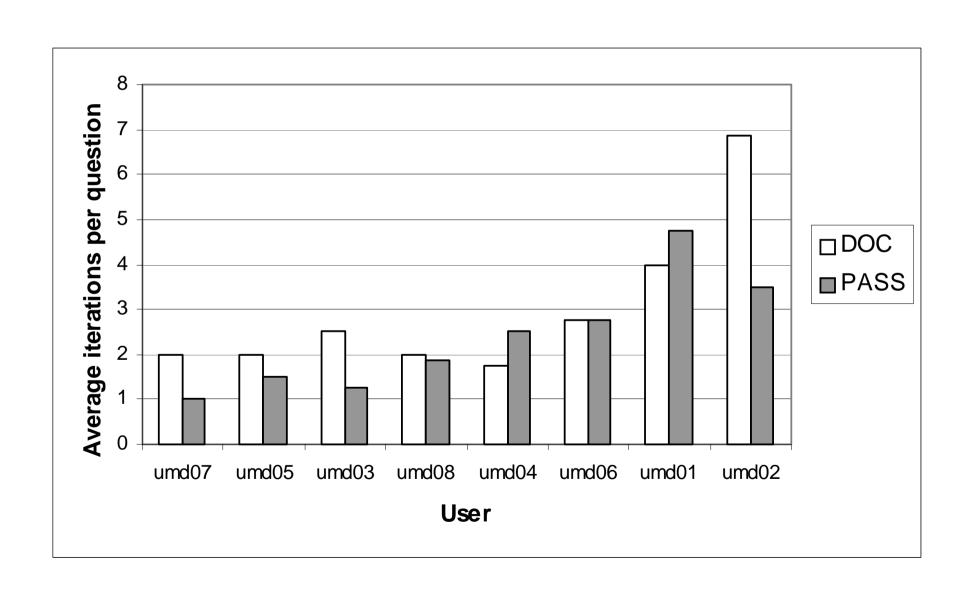System Status: [CLIENT] Retrieving Document ...Done!

iCLEF 2004 User Study Design

- Two systems
  - 3 Keyword in Context (KWIC) summaries
  - 1 single-passage summary

- 16 questions in English
- Spanish documents
- 8 skilled searchers, unskilled in Spanish

- 1-best answer, recorded on paper
- Supporting documents noted in interface
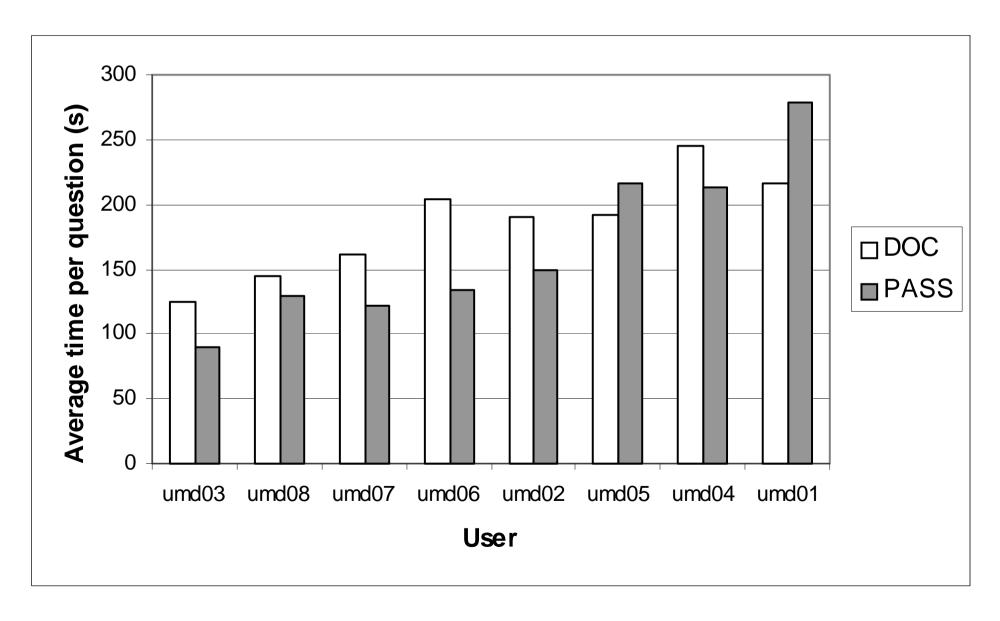 - 1 arbitrarily chosen for submission

Accuracy: No Difference

Iterations: No Difference

Search Time: Passages Help?

Question Difficulty

Questions, Grouped by Difficulty

8 Who is the managing director of the International Monetary Fund?

11 Who is the president of Burundi?

13 Of what team is Bobby Robson coach?

4 Who committed the terrorist attack in the Tokyo underground?

16 Who won the Nobel Prize for Literature in 1994?

6 When did Latvia gain independence?

14 When did the attack at the Saint-Michel underground station in Paris occur?

7 How many people were declared missing in the Philippines after the typhoon "Angela"?

2 How many human genes are there?

10 How many people died of asphyxia in the Baku underground?

15 How many people live in Bombay?

12 What is Charles Millon's political party?

1 What year was Thomas Mann awarded the Nobel Prize?

3 Who is the German Minister for Economic Affairs?

9 When did Lenin die?

5 How much did the Channel Tunnel cost?

Harder Questions Take Longer

More Iterations for Hard Questions

Subjective Reactions

- Both systems were "easy to use"

- Users liked question-term highlighting

- 5 of 8 users preferred single passages
  - Provided more context for answers

- 3 of 8 users preferred 3 KWIC sentences
  - Provided a good sense of document content
  - Required less time to read

Observed Search Behavior

- Consistently precision-oriented
  - Stopped after finding 1-2 confirming documents

- Initial use of exact question terms
  - Less use of synonyms than in topic searching

- Reformulated query with possible answers
  - Strikingly similar to "answer verification" in QA

- Sometimes the summary alone was enough

More Observations

- Bad MT sometimes hurt, but that was rare
  - Is a "Minister of Outer Subjects" the same as the "German Minister for Economic Affairs?"

- Prior knowledge did not always help
  - 5 of 8 found "When did Lenin die?" familiar
  - Only 2 found the right answer

- Desired answer granularity was not clear
  - July 25 was wrong (without the year)
  - September 1991 was right (without the day)

Conclusions

- CL-QA baselines are remarkably high
  - Makes it difficult to see significant differences

- Need to measure interannotator agreement
  - We're closer to the limit than automated systems

- We might learn from watching people
  - Perhaps design future evaluations around this?