

Università della Svizzera italiana	Facoltà di scienze informatiche	
<h1>How Solid are the Foundation of Speech-Driven Information Retrieval?</h1> <p><i>Fabio Crestani USI, Lugano</i></p>		
The 3eme Cycle Romand in Informatics, October 2007		1

Università della Svizzera italiana	Facoltà di scienze informatiche	Outline
<ul style="list-style-type: none"> ▪ Information Retrieval is changing <ul style="list-style-type: none"> – From “traditional” to “modern” information access ▪ Mobile Information Retrieval ▪ Speech driven Mobile Information Retrieval ▪ Some research at the foundations of speech driven Mobile Information Retrieval: <ul style="list-style-type: none"> – Spoken versus written queries – User perception of relevance of spoken documents 		
The 3eme Cycle Romand in Informatics, October 2007		2

Università della Svizzera italiana	Facoltà di scienze informatiche	Traditional Information Access
<ul style="list-style-type: none"> ▪ The design of “traditional” information access (IA) systems was based on a number of assumptions that are changing ▪ These are, for example: <ul style="list-style-type: none"> – Typology of users: librarians, documentalists, trained users, computer literates – The user: completely devoted to the IA task, interested in high precision and high recall, knows which information source might held sough information – ... 		
The 3eme Cycle Romand in Informatics, October 2007		3

Università della Svizzera italiana	Facoltà di scienze informatiche	Traditional Information Access
<ul style="list-style-type: none"> – ... – The information source: access one source at a time, simple cost function, each source has different modality of interaction – The IA system: has keyboard, large colour display, graphics – Bandwidth: large, basically free (i.e. the user does not pay) ▪ These assumptions have been changing is recent times and are no longer valid for most IA applications 		
The 3eme Cycle Romand in Informatics, October 2007		4

Università della Svizzera italiana | Facoltà di scienze informatiche

"Modern" Information Access

- The design of "modern" IA systems should now be based on radically different assumptions:
 - **Typology of users:** casual users, computer illiterates, "men of the street", children, elderly users
 - **The user:** involved in many tasks at the same time, mobile, often interested only in high precision, submits short casual queries, does not know which information resource might hold sought information
 - ...

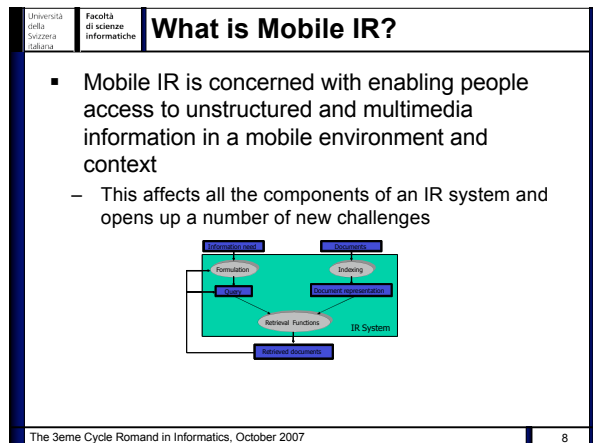
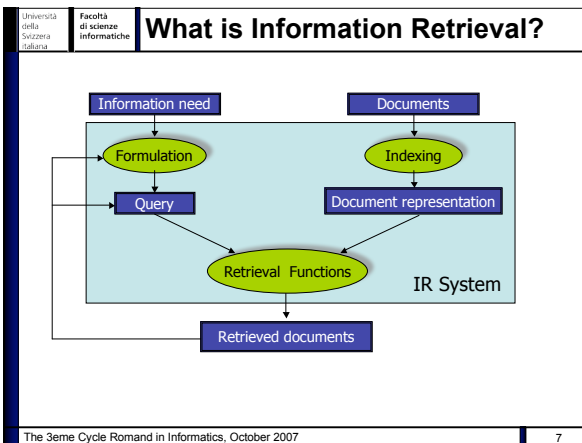
The 3eme Cycle Romand in Informatics, October 2007 | 5

Università della Svizzera italiana | Facoltà di scienze informatiche

"Modern" Information Access

- ...
- **The information source:** access to more than one source at the same time, complex cost function, toward simple and uniform modality of interaction
- **The IA system:** no keyboard (or very simple one), small screen, low quality graphics, voice recognition, handwriting recognition
- **Bandwidth:** low, costly (i.e. the user pays for it)
- This opens up several new areas of research
- One important area of research is Mobile Information Retrieval

The 3eme Cycle Romand in Informatics, October 2007 | 6



Università della Svizzera italiana	Facoltà di scienze informatiche	Research in Mobile IR
<ul style="list-style-type: none"> ▪ Research in Mobile IR is current and hot! ▪ Several research groups are involved in Mobile IR research <ul style="list-style-type: none"> – ICSI, Sheffield, Udine, ... ▪ There are many areas of research to do with Mobile IR ▪ One important area of research is concerned with speech-driven mobile IR research <ul style="list-style-type: none"> – Speech used for query formulation, user interaction, presentation of retrieval results, etc. 		
The 3eme Cycle Romand in Informatics, October 2007		9

Università della Svizzera italiana	Facoltà di scienze informatiche	IR research at USI
<ul style="list-style-type: none"> ▪ First at the University of Strathclyde and now at USI (University of Lugano) I have been engaged for several year in research in IR and, more recently, in mobile IR ▪ Currently working is on the foundations of speech driven mobile IR ▪ Trying to answer the following questions: <ul style="list-style-type: none"> – How is speech driven IR different from textual IR? – How should IR change to enable more effective speech driven mobile IR? 		
The 3eme Cycle Romand in Informatics, October 2007		10

Università della Svizzera italiana	Facoltà di scienze informatiche	Foundations
<ul style="list-style-type: none"> ▪ How is speech driven IR different from textual IR? <ul style="list-style-type: none"> – Is a spoken query different from a textual query? – Is there information in speech that should be used to make more effective IR of spoken documents? – Is the modality of presentation of a retrieved document affecting its perception of relevance to a user information need? – How should IR take advantage of spoken interaction with the user? ▪ These are just some research questions we are working on 		
The 3eme Cycle Romand in Informatics, October 2007		11

Università della Svizzera italiana	Facoltà di scienze informatiche	IR using Spoken Queries
<ul style="list-style-type: none"> ▪ Pros: <ul style="list-style-type: none"> – Speech is natural, rapid, more expressive – Speech has more information bearing cues than text (e.g. voice inflection, pitch, tone) ▪ Cons: <ul style="list-style-type: none"> – Automatic speech recognition (ASR) technology is imperfect – Speech is not always applicable in all situations – Cognitive load imposed by formulating spoken queries can't be ignored 		
The 3eme Cycle Romand in Informatics, October 2007		12

Università della Svizzera Italiana | Facoltà di scienze informatiche

Research Question

- Traditional IR has always been concerned with written queries and IR query processing has been designed for written queries
- But:
 - Are spoken queries the same as written queries?
 - Do they have the same characteristics of written queries?
 - Same length? Same words? Same "information content"?
 - Same level of effectiveness for information access?
- Should IR query processing change to deal with spoken queries?

The 3eme Cycle Romand in Informatics, October 2007 | 13

Università della Svizzera Italiana | Facoltà di scienze informatiche

Experimental Study

- Objective: compare written and spoken queries in the context of a search task
 - User has an information need
 - User express information need in a written or spoken query
 - User gives comments on complexity of task
- Analysis: compare the characteristics and the effectiveness of written and spoken queries
- Assumptions: perfect speech recognition

The 3eme Cycle Romand in Informatics, October 2007 | 14

Università della Svizzera Italiana | Facoltà di scienze informatiche

Experimental Setting

- Subjects: 12 native English speakers were recruited from the CIS Dept. of USG
- Test collection: 10 topics were extracted from the TREC topics collection
- Experimental procedural: two sessions per participant (written, spoken)
- Data capture: written queries were logged in text format, spoken queries were recorded and saved in wav format, along with the durations of query formulation

The 3eme Cycle Romand in Informatics, October 2007 | 15

Università della Svizzera Italiana | Facoltà di scienze informatiche

Query Length across Topics

Spoken queries were longer than written ones

Topic ID	Written queries (words)	Spoken queries (words)
1	~10	~15
2	~10	~15
3	~10	~15
4	~10	~15
5	~10	~15
6	~10	~15
7	~10	~15
8	~10	~15
9	~10	~15
10	~10	~15

- Spoken queries were longer than written ones for every topic
- The ease of speech encouraged participants to use more words
- More stopwords in spoken queries
- Average length of spoken queries (~14 words) almost doubled the average length of written queries (~7 words)

The 3eme Cycle Romand in Informatics, October 2007 | 16

Università della Svizzera Italiana Facoltà di scienze informatiche

Query Length across Participants

In general, more function words and more content words in spoken queries

- Spoken queries were longer than written ones for almost every user
- Length variation for written queries was small, but much larger for spoken queries
- Length variation related to user's personality and querying style
- Experts' spoken queries are very similar to written ones

The 3eme Cycle Romand in Informatics, October 2007 17

Università della Svizzera Italiana Facoltà di scienze informatiche

Query Formulation across Topics

Initially more time taken to formulate written queries than spoken queries

- On average, more time needed to formulate written queries than spoken ones
- Different situation when experience has been acquired
- Higher cognitive load and lacking of experience in issuing spoken queries put pressure on the formulation process

The 3eme Cycle Romand in Informatics, October 2007 18

Università della Svizzera Italiana Facoltà di scienze informatiche

Query Formulation across Part.

High variation in query formulation time across participants

- Relatively high variation of durations across participants
- Duration is related to the complexity of the topic
- Two thirds of the participants spent less time on spoken queries than written ones

The 3eme Cycle Romand in Informatics, October 2007 19

Università della Svizzera Italiana Facoltà di scienze informatiche

Evaluating Retrieval Effectiveness

- Test collection: articles from TREC, the Wall Street Journal collection (1990-1992)
- Queries: sets of written and spoken queries with document relevance assessment
- IR system: different systems built using Lemur IR toolkit (from CMU)
- Measures of IR system effectiveness: classical recall and precision

The 3eme Cycle Romand in Informatics, October 2007 20

Università della Svizzera Italiana Facoltà di scienze informatiche

TF-IDF Vector Space Model

RSVs calculated using classical tf-idf formula for vector space model

- Written and spoken query sets have very similar performance
- Common words set with a slightly better performance indicates important words are present in both query sets

The 3eme Cycle Romand in Informatics, October 2007 21

Università della Svizzera Italiana Facoltà di scienze informatiche

KL-JM Language Model

RSVs calculated using Kullback-Leibler similarity measure and Jelinek-Mercer smoothing for the IR language model

- All three sets have very similar performance and differences are not statistically significant
- This retrieval model does not make good use of important common words

The 3eme Cycle Romand in Informatics, October 2007 22

Università della Svizzera Italiana Facoltà di scienze informatiche

BM25 Probabilistic Model

RSVs calculated using the BM25 (a.k.a. Okapi) formula for the probabilistic model

- Written set outperforms spoken set and the differences are statistically significant
- Common words set obtains very similar but slightly better performance than written set
- Too much importance assigned by the term weighting model to words only appearing in the spoken query set

The 3eme Cycle Romand in Informatics, October 2007 23

Università della Svizzera Italiana Facoltà di scienze informatiche

Analysis of Results

- With TF-IDF and KL&JM models both query sets are equally effective
- Common words set with similar performance indicates useful words are present in both sets
 - But 71% of words in written queries are in the common words set whereas only 40.9% of spoken query words are in the set
- No clear explanation for the strange performance of the probabilistic model

The 3eme Cycle Romand in Informatics, October 2007 24

Università della Svizzera italiana | Facoltà di scienze informatiche | **Part of Speech (POS)**

Analysis of POS types in spoken and written queries

POS Type	Written Queries (%)	Spoken Queries (%)
n.	60	40
adj.	15	10
vb.	10	15
conj.	5	10
prep.	5	10
adv.	5	10
art.	5	10
pron.	5	10

Nouns have largest type shares in both queries and higher percentage in written ones
 Verbs are 2nd largest POS in spoken queries and 3rd largest in written ones, while adjectives are more common in written queries
 Prepositions and conjunctives are heavily used in spoken queries
 Experts' spoken queries are very similar to written ones

The 3eme Cycle Romand in Informatics, October 2007 | 25

Università della Svizzera italiana | Facoltà di scienze informatiche | **Conclusions from Exp. Results**

- Using speech to formulate an information needs not only provides a way to express it in a more natural way, but also encourages longer queries, but ...
 ... longer spoken queries do not produce better retrieval performance, as one would be inclined to assume!
- Query processing techniques designed for written queries do not work well for spoken queries. New techniques are needed!

The 3eme Cycle Romand in Informatics, October 2007 | 26

Università della Svizzera italiana | Facoltà di scienze informatiche | **User's Perception of Relevance**

- The absence of a screen compel Mobile IR application to present retrieval results in some non traditional way
- Presenting retrieval results using spoken document surrogates seems an obvious choice
- But:
 - Do users make correct relevance assessments when presented with document surrogates?
 - Do users make correct relevance assessments when presented with *spoken* document surrogates?
 - Do users make faster or slower assessments?

The 3eme Cycle Romand in Informatics, October 2007 | 27

Università della Svizzera italiana | Facoltà di scienze informatiche | **Perception of Relevance and Speech**

- We decided to experiment the effectiveness of retrieval results presentation using spoken surrogates
- In a previous study we have shown that document summaries are more effective than other surrogates to present retrieval results
 - Study carried out on PCs, PDAs and WAP mobile phones
 - Study involved different summary lengths and different sentence extraction based summarization techniques
- We now want to study spoken summaries

The 3eme Cycle Romand in Informatics, October 2007 | 28

Università della Svizzera italiana | Facoltà di scienze informatiche

Query-biased Summarisation

- Best technique of automatic text summarisation by sentence extraction for retrieval results presentation (better than full text!)
- Sentences are weighted using a combination of evidence from:
 - heuristics (e.g. ~news title and headings)
 - term weights based on collection and document statistics (tf-idf)
 - query score (consider terms in query)

The 3eme Cycle Romand in Informatics, October 2007 | 29

Università della Svizzera italiana | Facoltà di scienze informatiche

Experimental Settings

- Subjects: 10 native English speakers
- Text collection: summaries of documents results of 50 topics from TREC collection
- Experimental procedure:
 - subject assumed to submit a query
 - a list of relevant document surrogates is produced
 - document surrogates are presented in different modalities (on screen, human voice, human voice on telephone, synthesised voice on telephone)
- Take note of effectiveness, speed and ask to fill in a questionnaire at the end

The 3eme Cycle Romand in Informatics, October 2007 | 30

Università della Svizzera italiana | Facoltà di scienze informatiche

Experimental Results

Average Precision, Recall, and Time for different modalities of document delivery

Modality	Avg. Precision (%)	Avg. Recall (%)	Avg. Time (secs)
Computer - on screen	~45	~65	~15
Voice	~40	~55	~20
Telephone	~45	~50	~20
Computer - voice synthesiser	~45	~50	~25

The 3eme Cycle Romand in Informatics, October 2007 | 31

Università della Svizzera italiana | Facoltà di scienze informatiche

Analysis of Results

- User perception of relevance is highly influenced by the modality of results presentation
 - P remains almost constant across modalities
 - R and S decreases with increasing modality complexity (R and S are related)
 - Large across-subjects variations in P, R, S
 - Small across-topics variations in P, R, S
 - Most participants not happy with the quality of the synthesized voice
 - Most did not like spoken results presentation: too tiring (fatigue effect on data), difficult memorization of results

The 3eme Cycle Romand in Informatics, October 2007 | 32

Università della Svizzera italiana	Facoltà di scienze informatiche	Conclusions from Exp. Results
<ul style="list-style-type: none"> ▪ Presentation of retrieval results using speech is as almost as effective as results presentation on screen, but more can be done ▪ New ways of presenting retrieval results using speech are needed: <ul style="list-style-type: none"> – “Highlighting” search terms – Better speech synthesis systems – How do we aid results memorization? – What do we do with multimedia documents? 		
The 3eme Cycle Romand in Informatics, October 2007		33

Università della Svizzera italiana	Facoltà di scienze informatiche	Conclusions
<ul style="list-style-type: none"> ▪ Mobile and Ubiquitous Information Access is a very hot and interesting area of research ▪ Speech driven mobile IR is at the forefront of this research ▪ But: <ul style="list-style-type: none"> – Research is needed at the foundations, before building complex applications – Results from 50 years of IR research cannot just be applied to speech driven mobile IR as they are – Many IR techniques need to be modified to the new context! 		
The 3eme Cycle Romand in Informatics, October 2007		34

Università della Svizzera italiana	Facoltà di scienze informatiche	Thank you
<ul style="list-style-type: none"> ▪ Questions? ▪ References: <ul style="list-style-type: none"> – F.Crestani, H.Du. Written versus Spoken Queries: a Qualitative and Quantitative Comparative Analysis. <i>Journal of the American Society for Information Science and Technology</i>, 57(7):881-890, 2006. – A.Tombros, F. Crestani. Users' Perception of Relevance of Spoken Documents. <i>Journal of the American Society for Information Science</i>, 51(9):929-939, 2000. 		
The 3eme Cycle Romand in Informatics, October 2007		35