Statistical Machine Translation
LING-462/COSC-482
Week 13:
Integrating MT in other NLP applications
Speech-to-speech translation

Achim Ruopp

achim.ruopp@georgetown.edu

Agenda

- Language in ten minutes: Greek – Meng Han
- Integrating MT in other NLP applications

 - Break -

- Speech-to-speech translation

MT Use Case: Dissemination

- High-quality, locale appropriate translations
  - Not possible without humans in the loop
  - Covered in last lecture: process, post-editing, interactive translation prediction
- Less important, time-sensitive, cheaper "good enough" translations for
  - User-generated content
    - Reviews
    - Product listings on market places
    - …
  - Support documentation
  - Live subtitling
  - …

MT Use Case: Market Place
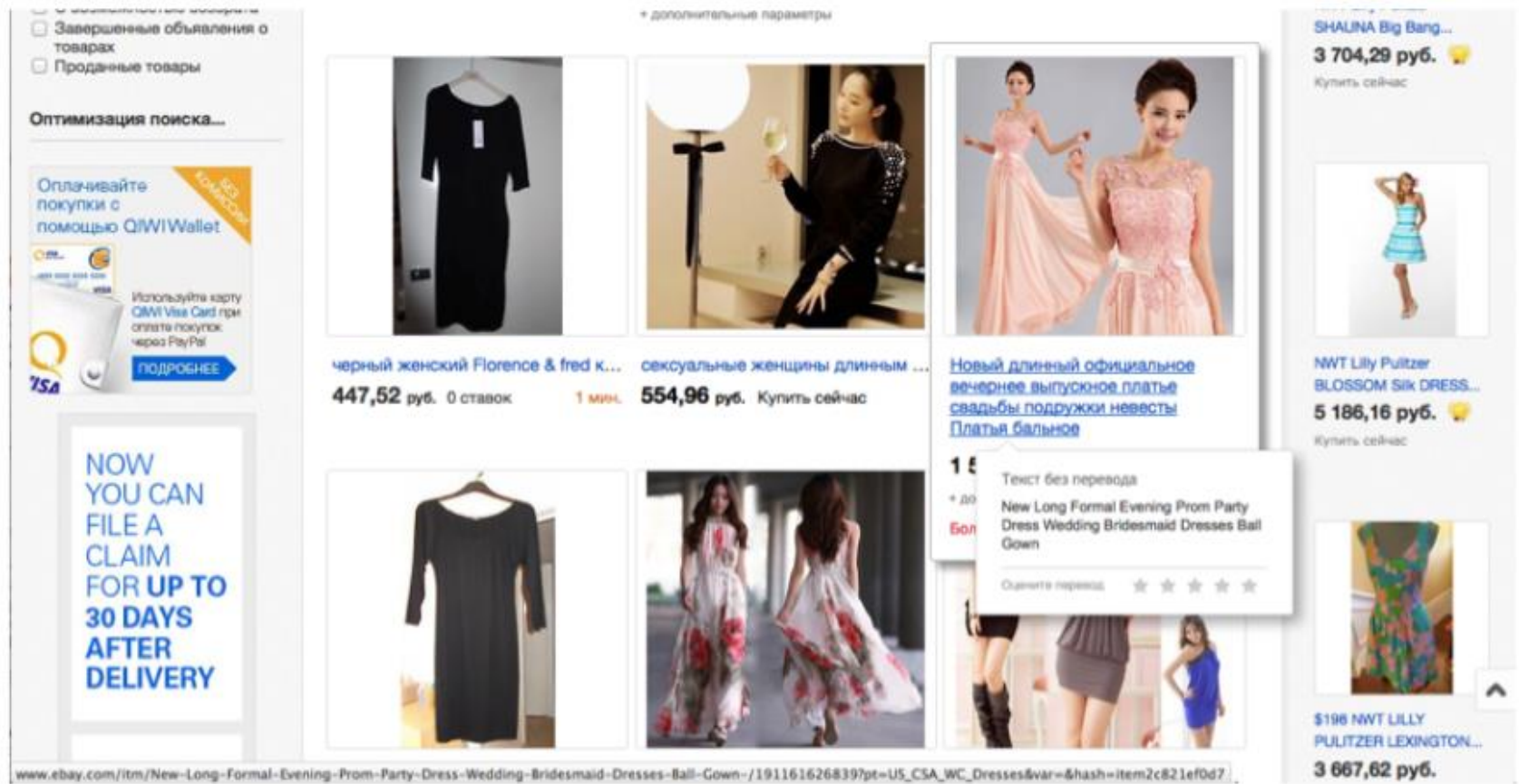
Figure 2: Search Result Page for the search term "платье", showing a Russian user's results with automatic translation into Russian. Original text can be seen, when hovering with the mouse over the translation. Also, the user is able to provide feedback to the MT system, allowing to incorporate this feedback to improve machine translation quality.

MT Use Case: Market Place

Jyoti Guha, Carmen Heger; Machine Translation for Global eCommerce on eBay
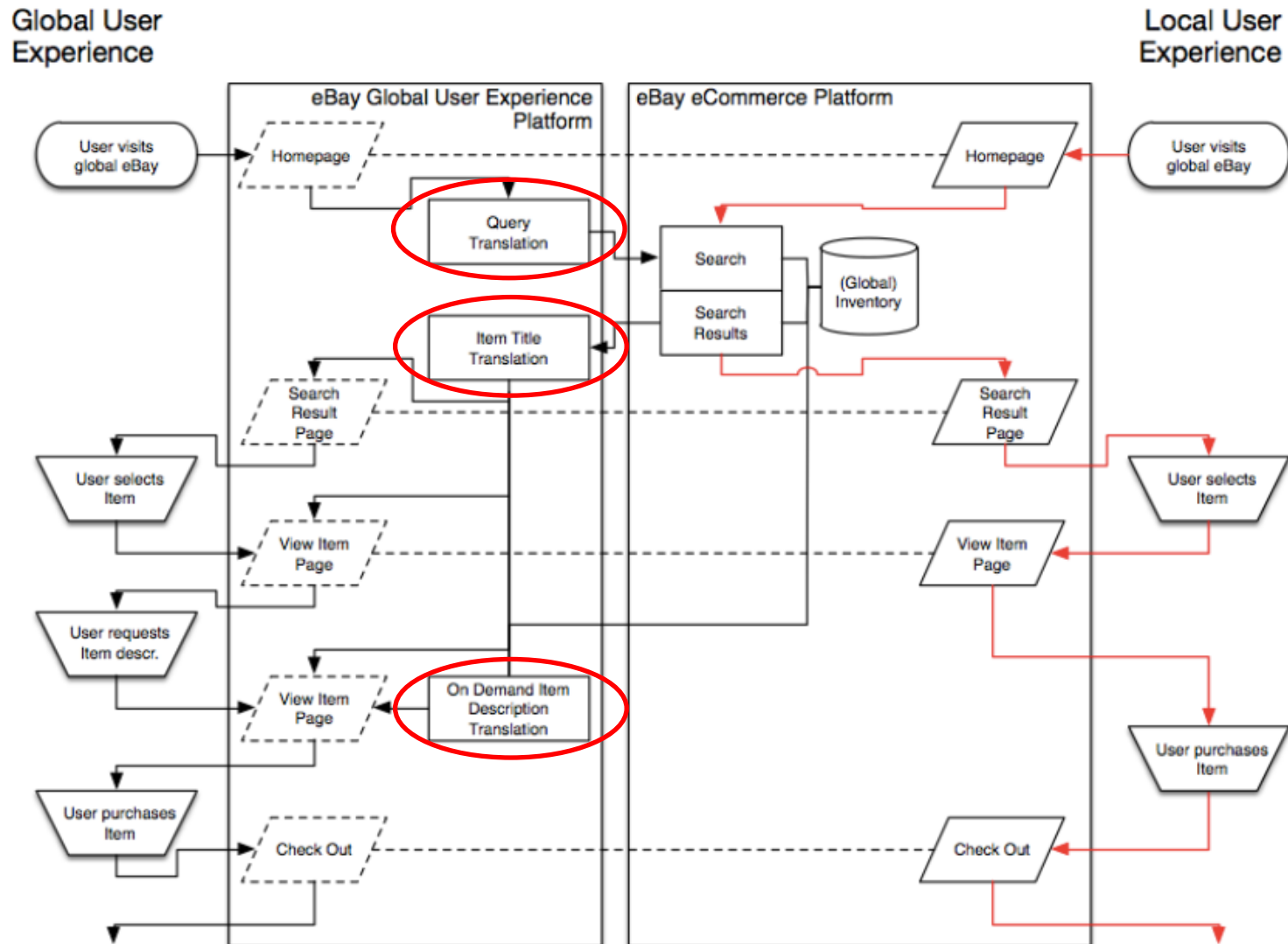
MT Use Case: Market Place

- Search query translation
  - Short phrases that need to be optimized for search engine
  - Russian query "сумки из натуральной кожи" can be translated to
    - "bags of genuine leather" – 160 search results
    - "genuine leather bags" – 47,000 search results
  - Brand names – do not translate
    - "GAP" – ambiguity in English
    - "REI" – means "king" in Portuguese

MT Use Case: Market Place

- Topic titles
  - Example: "new men's white jekyll & hyde jeans winston designer regular fit shirt size s-xxl"
  - Topic adaptation
- Topic descriptions
  - Also require domain-adapted engine, but different from topic translation engine

Mathur et. al.; Topic adaptation for machine translation of e-commerce content

MT Use Case: Assimilation

- Closely related to dissemination in some cases
  - E.g. Facebook shifted from on-demand translation of foreign-language posts ("Translate" link), i.e. assimilation to automatically translating posts to the native language of the user, i.e. dissemination
  - Pull (Assimilation) vs. Push (Dissemination)
  - Makes a big difference on needed machine resources

MT Use Case: Cross-lingual Information Retrieval (CLIR)

- Cross-language or cross-lingual IR
- "retrieving information written in a language different from the language of the user's query" – Wikipedia
- Machine translation
  - Translate the material to be searched
  - Translate the search query
    - Need to translate discovered data for gisting/other use afterwards

MT Use Case: Cross-lingual Information Retrieval (CLIR)

- MT not the only way
  - Parallel corpora-based methods
  - Dictionary-based methods
- CLEF Campaigns
  - Cross-Language Education and Function, now Conference and Labs of the Evaluation Forum
  - http://www.clef-initiative.eu/
  - Earlier: http://www.clef-campaign.org/
  - Until 1999 a track of the TREC conference (text retrieval and evaluation conference)
- Again different metrics! Precision/recall

https://en.wikipedia.org/wiki/Cross-language_information_retrieval

MT Use Case: Cross-lingual Information Retrieval (CLIR)

CLIR classification by use

- Information retrieval
- (Media) monitoring
- Information filtering and routing
- Sentiment analysis
- Information extraction
- Question answering

CLIR classification by material

- Ecommerce listings (see above)
- eDiscovery
  - Discovery of relevant documents for lawsuits or other legal cases
  - Iconic Translation Machines
- Intelligence
  - Web/News/secret stuff
  - Multiple ongoing DARPA programs: DEFT, LORELEI, BOLT

# MT Use Case: Internet is becoming more multilingual

TOP TEN LANGUAGES IN THE INTERNET	World Population for this Language (2018 Estimate)	Internet Users by Language	Internet Penetration (% Population)	Internet Users Growth (2000 - 2018)	Internet Users % of World Total (Participation)
Top Ten Languages Used in the Web - December 31, 2017 (Number of Internet Users by Language)					
English	1,462,008,909	1,052,764,386	72.0 %	647.9 %	25.3 %
Chinese	1,452,593,223	804,634,814	55.4 %	2,390.9 %	19.4 %
Spanish	515,759,912	337,892,295	65.5 %	1,758.5 %	8.1 %
Arabic	435,636,462	219,041,264	50.3 %	8,616.0 %	5.3 %
Portuguese	286,455,543	169,157,589	59.1 %	2,132.8 %	4.1 %
Indonesian / Malaysian	299,271,514	168,755,091	56.4 %	2,845.1 %	4.1 %
French	127,185,332	118,626,672	93.3 %	152.0 %	2.9 %
Japanese	143,964,709	109,552,842	76.1 %	3,434.0 %	2.7 %
Russian	405,644,599	108,014,564	26.6 %	800.2 %	2.8 %
German	94,943,848	84,700,419	89.2 %	207.8 %	2.2 %
TOP 10 LANGUAGES	5,135,270,101	3,206,613,856	62.4 %	1,091 %	77.1 %
Rest of the Languages	2,499,488,327	950,318,284	38.0 %	935 %	22.9 %
WORLD TOTAL	7,634,758,428	4,156,932,140	54.4 %	1,051 %	100.0 %

Neural Conversational Model
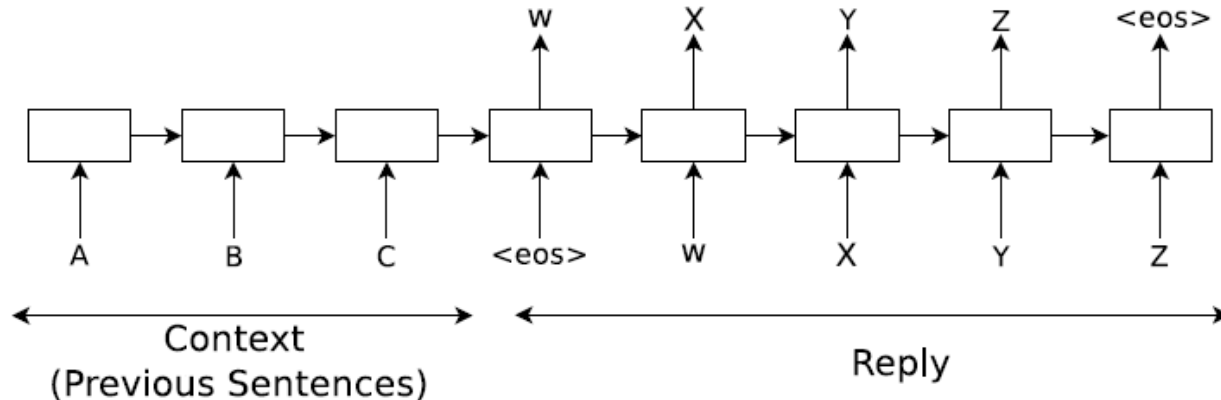
- **Machine:** *hi , this is <NAME> from helpdesk connect , what can i help you with today ?*
- **Human:** *hi , i forgot my password*
- **Machine:** *ok great*
- **Human:** *can you tell me how i can recover it ?*
- **Machine:** *i ' ll need to verify who it is at that*
- *machine*
- **Machine:** *can we do a hangout*
- *…*

Oriol Vinals and Quoc V. Le, A Neural Conversational Model, 2015, arXiv

Context-Aware Query Suggestion with Encoder-Decoder Model

Sordoni et. al.; A Hierarchical Recurrent Encoder-Decoder For Generative Context-Aware Query Suggestion
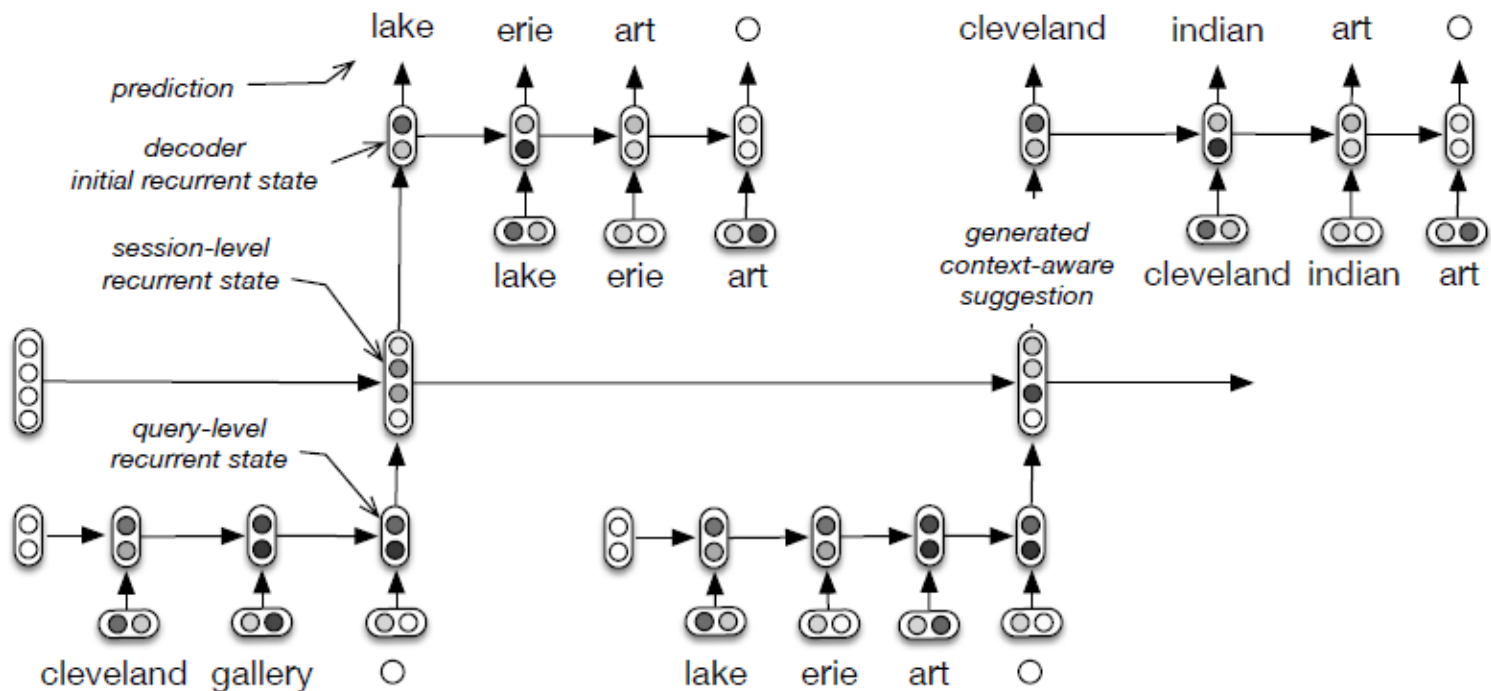
MT Use Case: Conversational MT

- Can be extended to include MT
- Unbabel is providing this commercially for support conversations

References

- Jyoti Guha, Carmen Heger; Machine Translation for Global eCommerce on eBay; 2014, Proceedings of the Conference of the Association for Machine Translation in the Americas (AMTA), Vancouver

- Prashant Mathur, Marcello Federico, Selçuk Köprü, Shahram Khadivi and Hassan Sawaf ; Topic adaptation for machine translation of e-commerce content; 2015, MT Summit XV, pp. 270-282

References

- Oriol Vinals and Quoc V. Le; A Neural Conversational Model; 2015, arXiv
- Alessandro Sordoni, Yoshua Bengio, Hossein Vahabi, Christina Lioma, Jakob G. Simonsen, Jian-Yun Nie; A Hierarchical Recurrent Encoder-Decoder For Generative Context-Aware Query Suggestion; 2015, arXiv
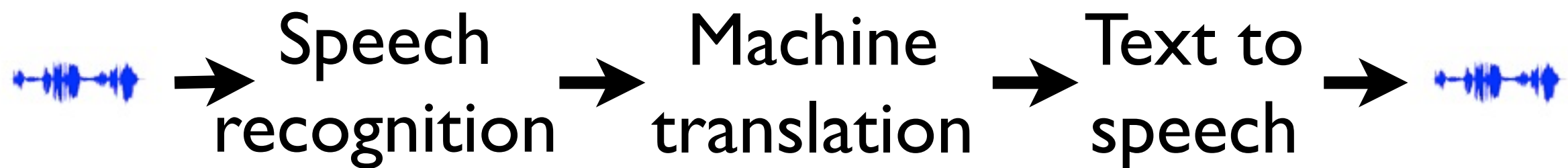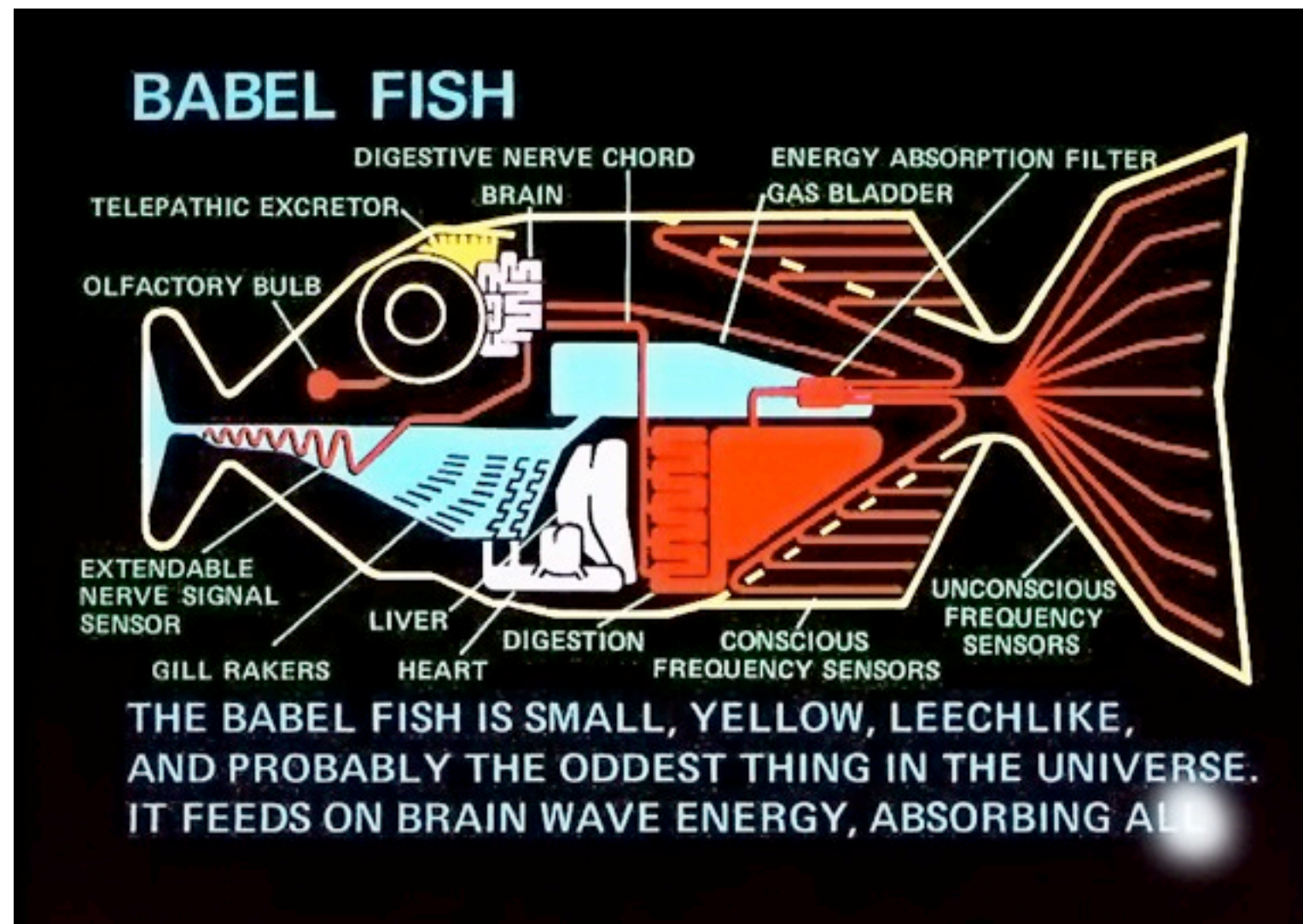
Spoken Language Translation

Alexandra Birch

School of Informatics
University of Edinburgh

Universal translator
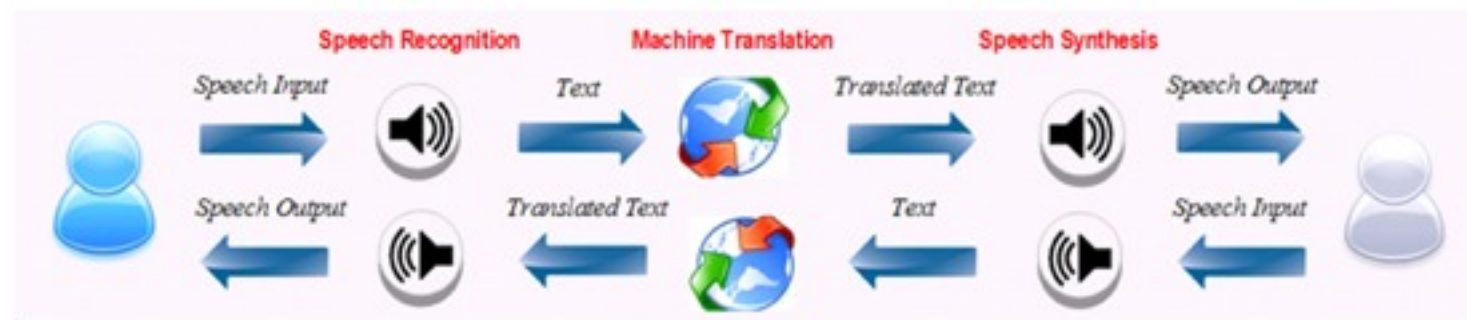

$\rightarrow$ **Speech recognition** $\rightarrow$ **Machine translation** $\rightarrow$ **Text to speech** $\rightarrow$

Science fiction

Industry adoption

2007 IBM's speech-speech translator for Iraq

limited domains

Industry adoption

open domain - more language pairs

Industry adoption

massive reach - uncertain quality

Research History

Project	Timespan	Description
C-STAR	1992-2004	One/Two-way Limited domain spontaneous speech
VERBMOBIL	1993-2000	One-way Limited domain conversational speech
DARPA BABYLON	2001-2003	Two-way Limited domain spontaneous speech; Laptop/Handheld
DARPA TRANSTAC	2005-2010	Two-way small domain spontaneous dialog; Mobile/Smartphone;
TC-STAR	2004-2007	One-way; Broadcast speech, political speech; Server-based
DARPA GALE	2006-2011	One-way Broadcast speech, conversational speech; Server-based
DARPA BOLT	2011-	One/two-way Conversational speech; Server-based

from Zhou and He ICASSP 2013

Research History

Example Based Interlingua $\longrightarrow$ Statistical Methods

Limited domain $\longrightarrow$ Broad domain

tourism, security   large vocabulary

Formal $\longrightarrow$ Conversational

Where is the train station please?   Catch you later!

Desktop $\longrightarrow$ Mobile/Cloud

Current Success
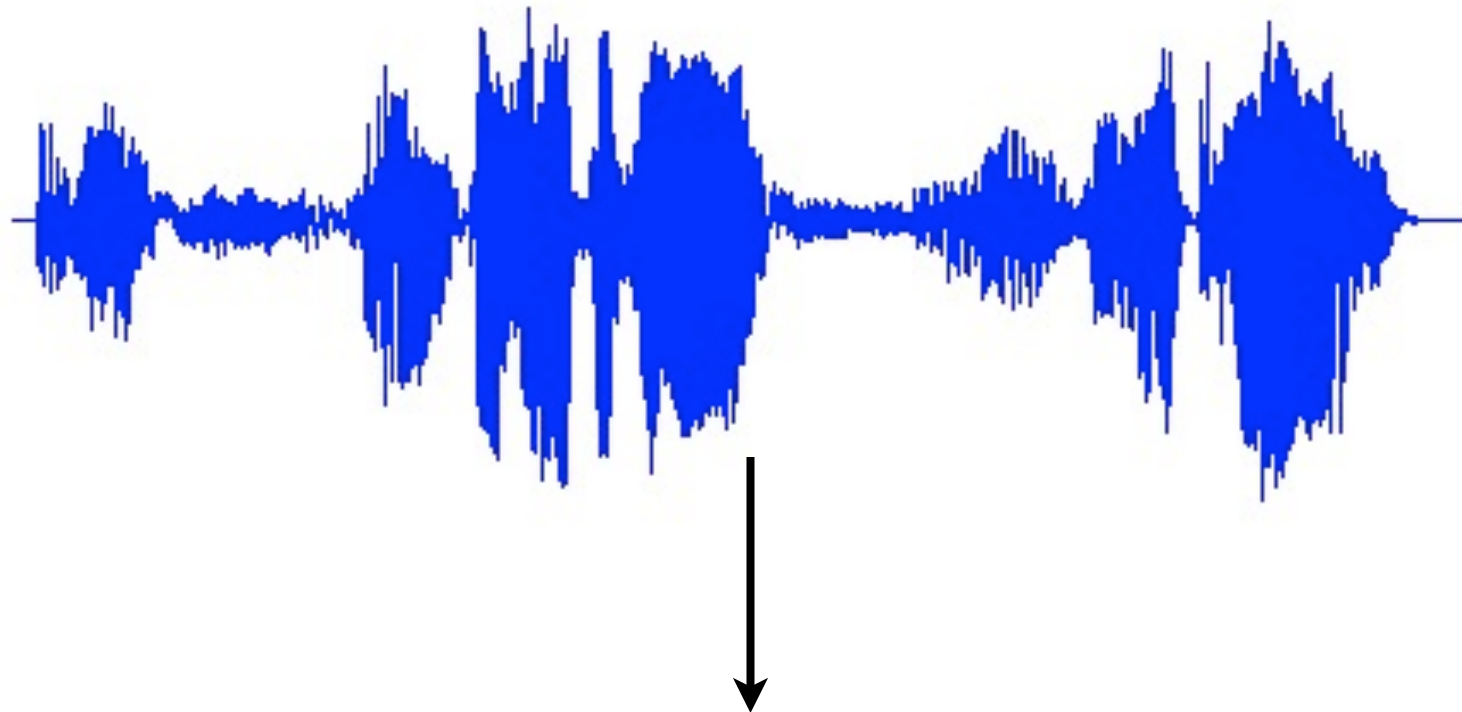
+ Statistical models

+ Large amounts of training data

+ Discriminative training

+ Better models (DNN for ASR)

Overview

- Spoken Language Translation

  ✦ Background

  ✦ Challenges and opportunities

  ✦ Demos

- Experiments in SLT

Automatic Speech Recognition

THE HEAVIEST TRAIN IS GOING TO FALL IN A MY MANCHESTER
AND THE SURROUNDING AREA

Challenges: accents, articulation, pronunciation, speed, recording quality, background noise

Automatic Speech Recognition

$$\mathbf{f}_{best} = argmax_f \, p(\mathbf{x}|\mathbf{f}) p(\mathbf{f})$$
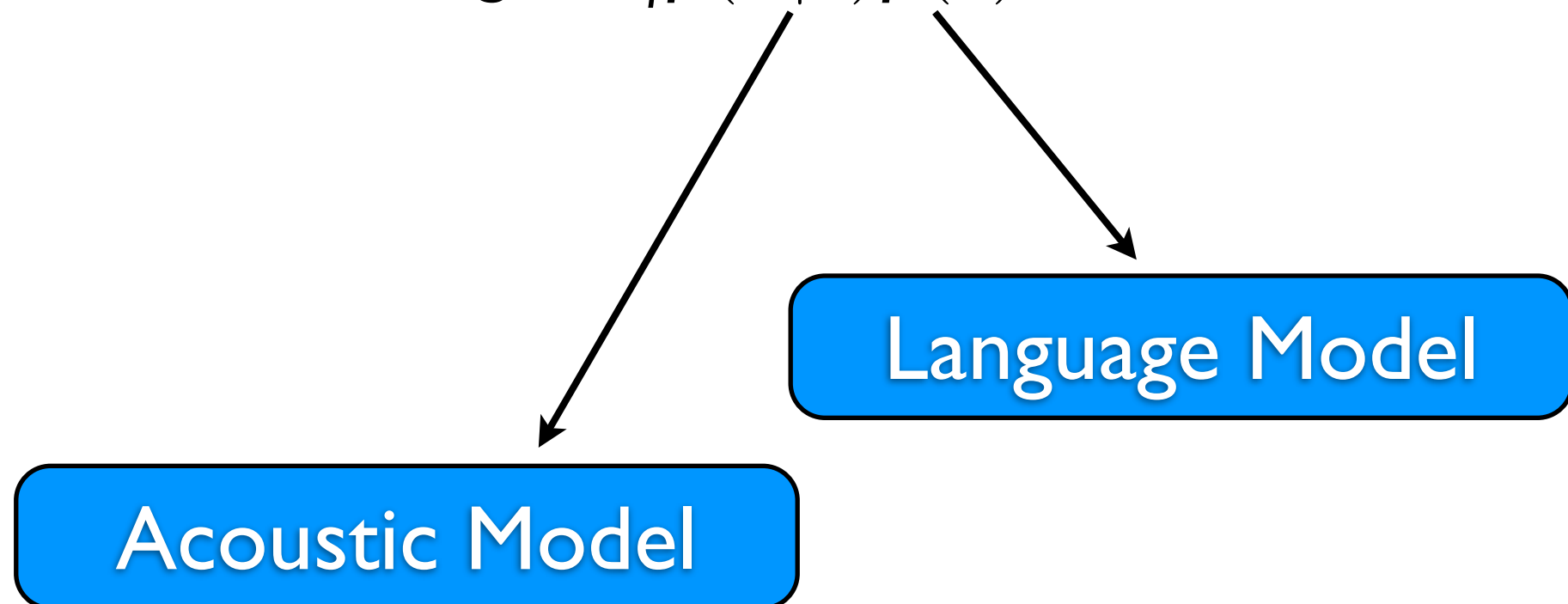
Automatic Speech Recognition

$$\mathbf{f}_{best} = argmax_f p(\mathbf{x}|\mathbf{f})p(\mathbf{f})$$
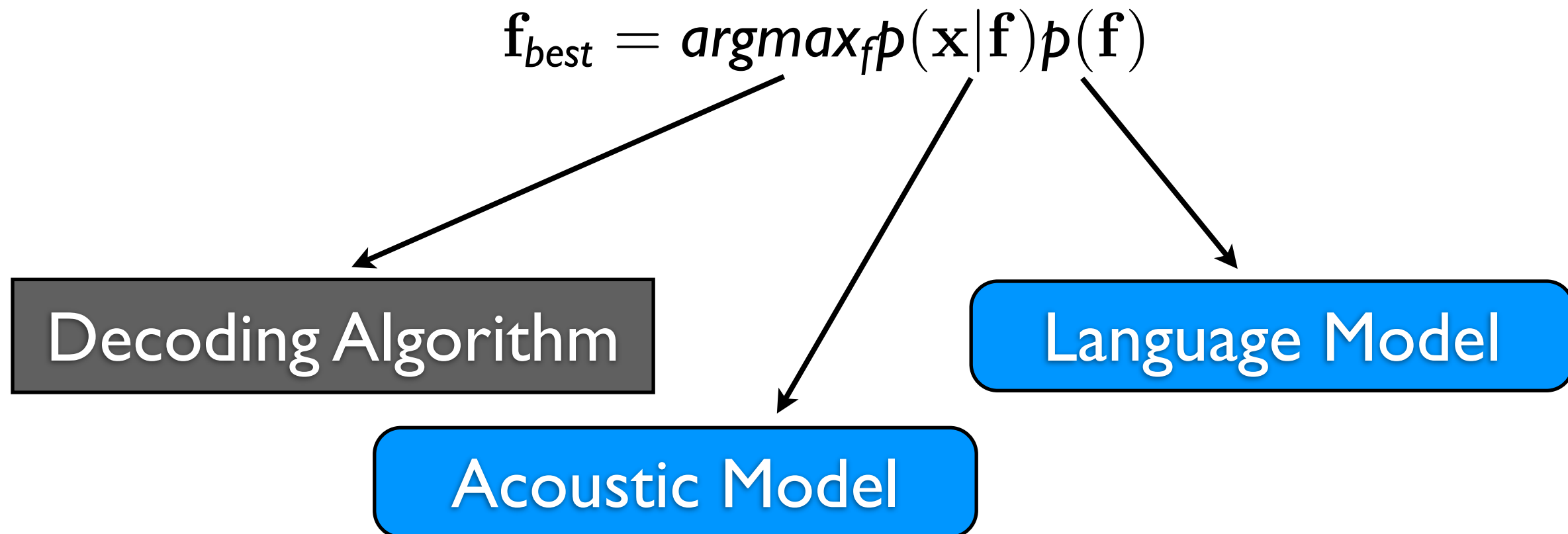
Language Model

Automatic Speech Recognition

$$\mathbf{f}_{best} = argmax_f\, p(\mathbf{x}|\mathbf{f})p(\mathbf{f})$$

Language Model

Acoustic Model

Automatic Speech Recognition

$$\mathbf{f}_{best} = argmax_f\, p(\mathbf{x}|\mathbf{f})p(\mathbf{f})$$

Decoding Algorithm

Acoustic Model

Language Model

ASR process

Word Sequence

CAN WE MEET

+ LM

Phoneme Sequence

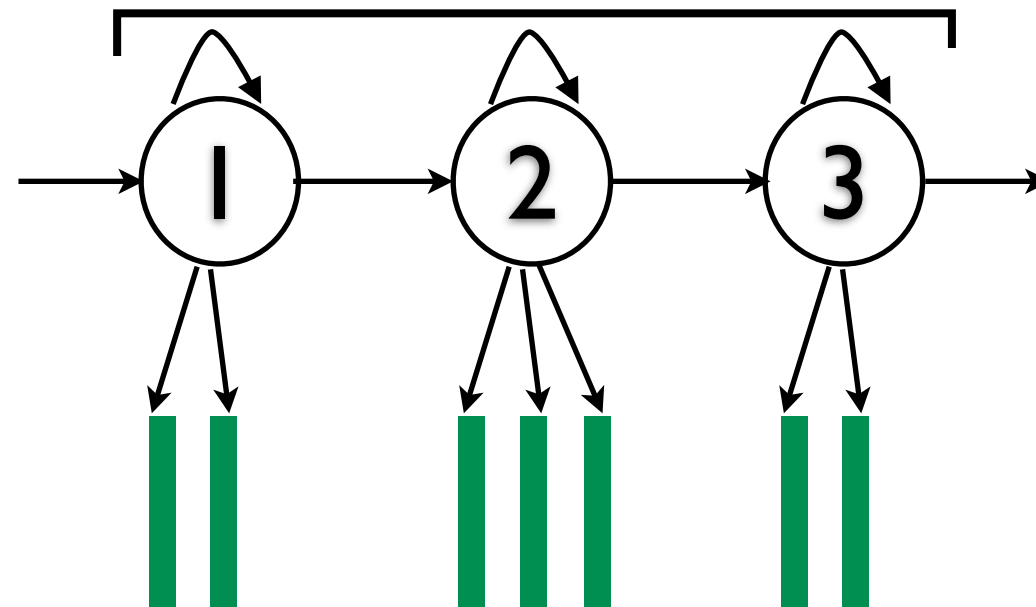**K AE N W IY M IY T**

+ pronunciation dictionary

Tri-phone Sequence

N-W+IY

HMM-based Tri-phone Model

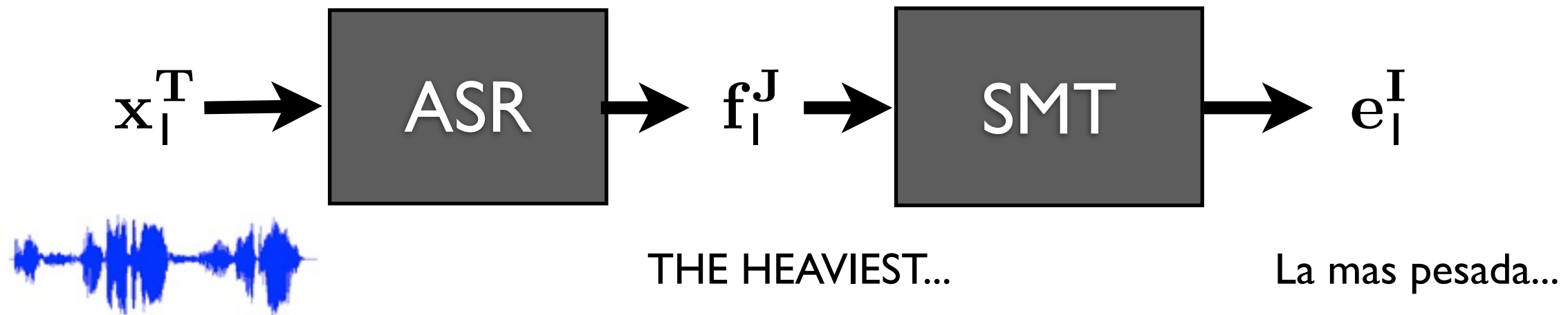Spectrum Observation Sequence

Speech Signal

Word Error Rate

$$WER = \frac{S + D + I}{N}$$

- S is the number of substitutions,
- D is the number of deletions,
- I is the number of insertions and
- N is the number of words in the reference

Much easier to evaluate ASR than MT!

Spoken Language Translation

$$x_1^T \rightarrow \boxed{\text{ASR}} \rightarrow f_1^J \rightarrow \boxed{\text{SMT}} \rightarrow e_1^I$$

THE HEAVIEST...   La mas pesada...

✦ Input: a source speech signal sequence

✦ ASR: recognises set of source words

✦ SMT: translates into target words

$$x_1^T = x_1, \ldots, x_T$$

$$f_1^J = f_1, \ldots, f_J$$

$$e_1^I = e_1, \ldots, e_I$$

ASR vs SMT

- Similar: sequential pattern recognition

ASR: $x_1^T \longrightarrow f_1^J$

MT: $f_1^J \longrightarrow e_1^I$

- Original IBM SMT models1990 based on ASR

- Different: ASR is **monotonic** but SMT is not

  ✦ Implications for modelling and decoding

SLT Challenges

ASR introduces errors

Verbatim Speech:

```
... RAIN IS GOING TO FALL IN ER MANCHESTER ...
```

ASR output:

```
... TRAIN IS GOING TO FALL IN A MY MANCHESTER ...
```

SLT Challenges

Lack of case leads to ambiguities

ASR output:

`GEORGE BUSH IS VISITING NEW ORLEANS`

Normal text:

George Bush is visiting New Orleans

SLT Challenges

Lack of punctuation

ASR:

`HE SAID YOU ARE NOT WELCOME`

Normal text:

He said "You are not welcome?"

Challenges

Disfluencies and Restarts and repairs

Verbatim Speech:

```
I beg your pardon, but at the beginning of
these uh human rights debates
I uh should have pointed out to im- uh
important human rights uh defenders from
Misses uh Estemirowa
from Chechnya and another from Russia are
here
they came along to visit uh group eh uh,
uh they weren't here at the time
at the beginning
uh uh when we began the debate,
but
they've uh come
they've just come in now .
```

Official transcript:

I beg your pardon, but at the beginning of these human rights debates, I should have pointed out that two important human rights defenders are here, Mrs Estemirowa, from Chechnya, and Mr Lokshina, from Russia.

They were not here at the beginning when we began the debate, but they have just come in now.

Challenges

Disfluencies and Restarts and repairs

Verbatim Speech:

```
I beg your pardon, but at the beginning of
these uh human rights debates
I uh should have pointed out to im- uh
important human rights uh defenders from
Misses uh Estemirowa
from Chechnya and another from Russia are
here
they came along to visit uh group eh uh,
uh they weren't here at the time
at the beginning
uh uh when we began the debate,
but
they've uh come
they've just come in now .
```

Official transcript:

I beg your pardon, but at the beginning of these human rights debates, I should have pointed out that two important human rights defenders are here, Mrs Estemirowa, from Chechnya, and Mr Lokshina, from Russia.

They were not here at the beginning when we began the debate, but they have just come in now.

Challenges

Disfluencies and Restarts and repairs

Verbatim Speech:

```
I beg your pardon, but at the beginning of
these uh human rights debates
I uh should have pointed out to im- uh
important human rights uh defenders from
Misses uh Estemirowa
from Chechnya and another from Russia are
here
they came along to visit uh group eh uh,
uh they weren't here at the time
at the beginning
uh uh when we began the debate,
but
they've uh come
they've just come in now .
```

Official transcript:

I beg your pardon, but at the beginning of these human rights debates, I should have pointed out that two important human rights defenders are here, Mrs Estemirowa, from Chechnya, and Mr Lokshina, from Russia.

They were not here at the beginning when we began the debate, but they have just come in now.

Challenges

Disfluencies and Restarts and repairs

Verbatim Speech:

```
I beg your pardon, but at the beginning of
these uh human rights debates
I uh should have pointed out to im- uh
important human rights uh defenders from
Misses uh Estemirowa
from Chechnya and another from Russia are
here
they came along to visit uh group eh uh,
uh they weren't here at the time
at the beginning
uh uh when we began the debate,
but
they've uh come
they've just come in now .
```

Official transcript:

I beg your pardon, but at the beginning of these human rights debates, I should have pointed out that two important human rights defenders are here, Mrs Estemirowa, from Chechnya, and Mr Lokshina, from Russia.

They were not here at the beginning when we began the debate, but they have just come in now.

Challenges

Segmentation

Human segmentation

- Je vais vous parler aujourd'hui de l'énergie et le climat.
- Et ça peut sembler un peu surprenant parce que mon travail à temps plein à la Fondation est principalement sur les vaccins et les graines sur les choses que nous avons besoin d'inventer et d'aider les plus pauvres, deux milliards de vivre une vie meilleure.
- Mais l'énergie et le climat sont extrêmement importantes pour ces gens, en fait, plus important que n'importe qui d'autre sur la planète.

23.87 BLEU

Automatic segmentation

- Je vais vous parler aujourd'hui.
- L'énergie et le climat.
- Ça peut sembler un peu surprenant, parce
- Travail à temps plein à la Fondation.
- Est principalement sur les vaccins et les graines sur les choses que nous avons besoin d'inventer et livrer.
- Aider les plus pauvres deux milliards de dollars.
- Vivre une vie meilleure.
- Mais l'énergie et le climat sont extrêmement importantes pour ces gens, en fait.
- Plus important.
- À quelqu'un d'autre sur la planète.

18.30 BLEU

Challenges

Lack of data

- Large quantities of parallel written text:

Europarl: 48M words

Commoncrawl: 53M

For German-English

- Small quantities of transcribed, parallel speech:

TED: 3M

OPUS subtitles: 31M - noisy

Challenges

Variety of spoken language

Europarl:

> I declare resumed the session of the European Parliament, adjourned on Friday 17 December 1999, and I would like once again to wish you a happy new year in the hope that you enjoyed a pleasant festive period.

TED:

> It can be a very complicated thing, the ocean.
> And it can be a very complicated thing, what human health is.

OPUS subtitles:

> Mewling quim!
> It's no sea serpent.
> Salamander?
> Yes?
> It's Fiks.
> What's up?
> I must return to Oslo.

Opportunities

- smaller vocabulary

- simpler, shorter sentences

- structured dialog

IWSLT

- Annual workshop on spoken language translation

- Running since 2003

- TED talks

- ASR, SLT, MT