

Sec 6:

Real-world Applications

Automatic generation of subtitles

Simultaneous translation

Sec 6.1

Automatic Generation of Subtitles

Automatic subtitling - Motivation



- Explosion of audio-visual content available (Cinema, OTT platforms, social media,...)
 - Need: offer high-quality subtitles into dozens of languages in a short time
 - Problem: human subtitling is slow and costly (1-15\$/min)
 - Goal: automatic solutions to reduce human workload and costs

What is special about Subtitling?

- Importance of time
- Text needs to satisfy spatial and temporal constraints

In and out times based on speech rhythm

Length:

max. 2 lines (of \approx length)

max. 42 characters/line

Reading speed:

max. 21 characters/second



Segmenting into proper subtitles

This kind of harassment keeps women <eob> from accessing the internet – <eol>
essentially, knowledge. <eob>

```
10
00:00:31,066 --> 00:00:34,390
This kind of harassment keeps women
11
00:00:34,414 --> 00:00:36,191
from accessing the internet --
essentially, knowledge.
```

Segmenting into proper subtitles

This kind of harassmt keeps women <eol> from accessing the internet – <eob>
essentially, knowledge. <eob>

```
10
00:00:31,066 --> 00:00:34,390
This kind of harassmt keeps women
11
00:00:34,414 --> 00:00:36,191
from accessing the internet --
essentially, knowledge.
```

```
10
00:00:31,066 --> 00:00:34,390
This kind of harassmt keeps women
from accessing the internet --
11
00:00:34,414 --> 00:00:36,191
essentially, knowledge.
```

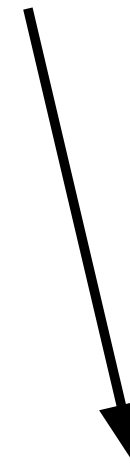
Segmentation approaches

Manual template

This kind of harassment keeps
women <eol> from accessing
the internet – <eob>



MT



Ce harcèlement empêche
les femmes <eol>
d'accéder à Internet,
<eob>

Segmentation approaches

Manual template

This kind of harassment keeps women <eol> from accessing the internet – <eob>

MT

Previous works focused only on length-matching given the template

(Matusov et al., 2019; Lakew et al., 2019)

Ce harcèlement empêche les femmes <eol> d'accéder à Internet, <eob>

Segmentation approaches

Manual template

This kind of harassment keeps
women <eol> from accessing
the internet – <eob>

MT

Cascade



ASR

this kind of harassment
keeps woman from
accessing internet

MT

Ce harcèlement empêche
les femmes <eol>
d'accéder à Internet,
<eob>

Segmentation approaches

Manual template

This kind of harassment keeps
women <eol> from accessing
the internet – <eob>

MT

Cascade



ASR

this kind of harassment
keeps woman from
accessing internet

MT

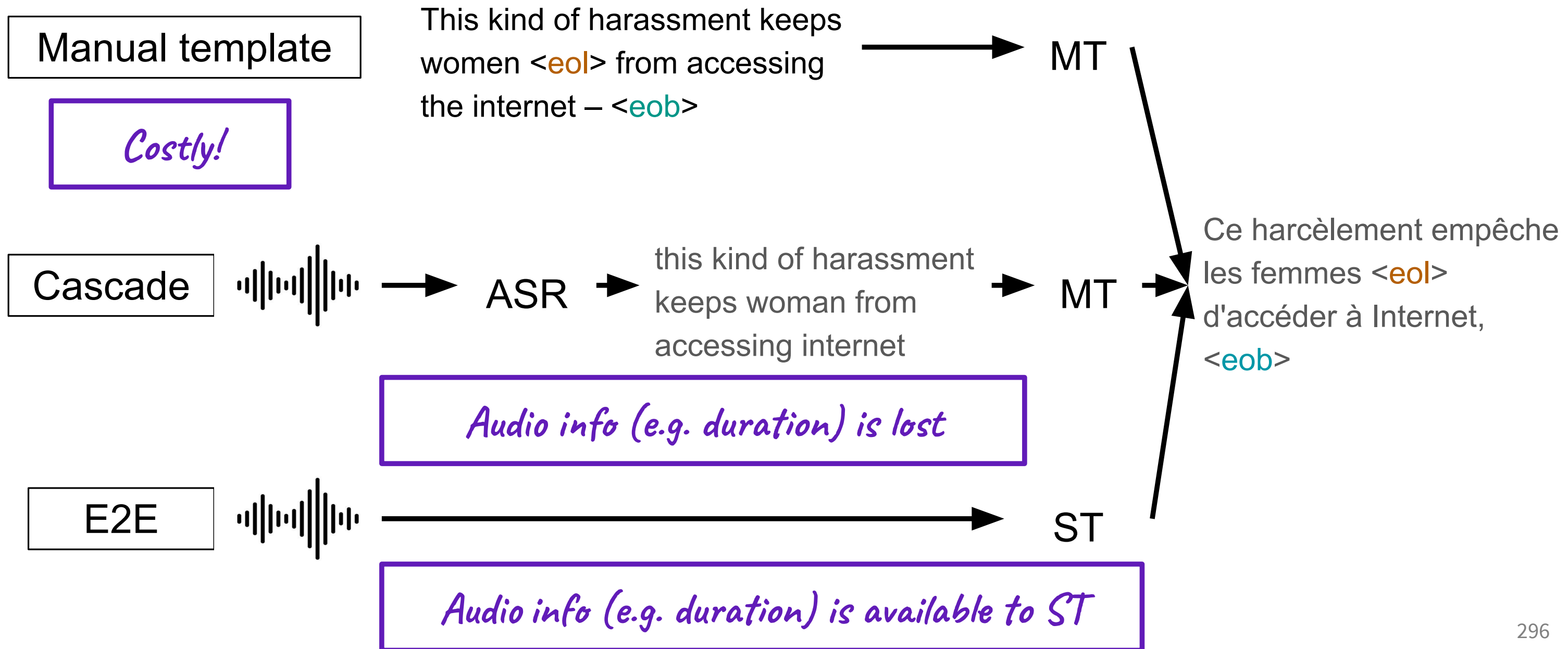
Ce harcèlement empêche
les femmes <eol>
d'accéder à Internet,
<eob>

E2E



ST

Segmentation approaches



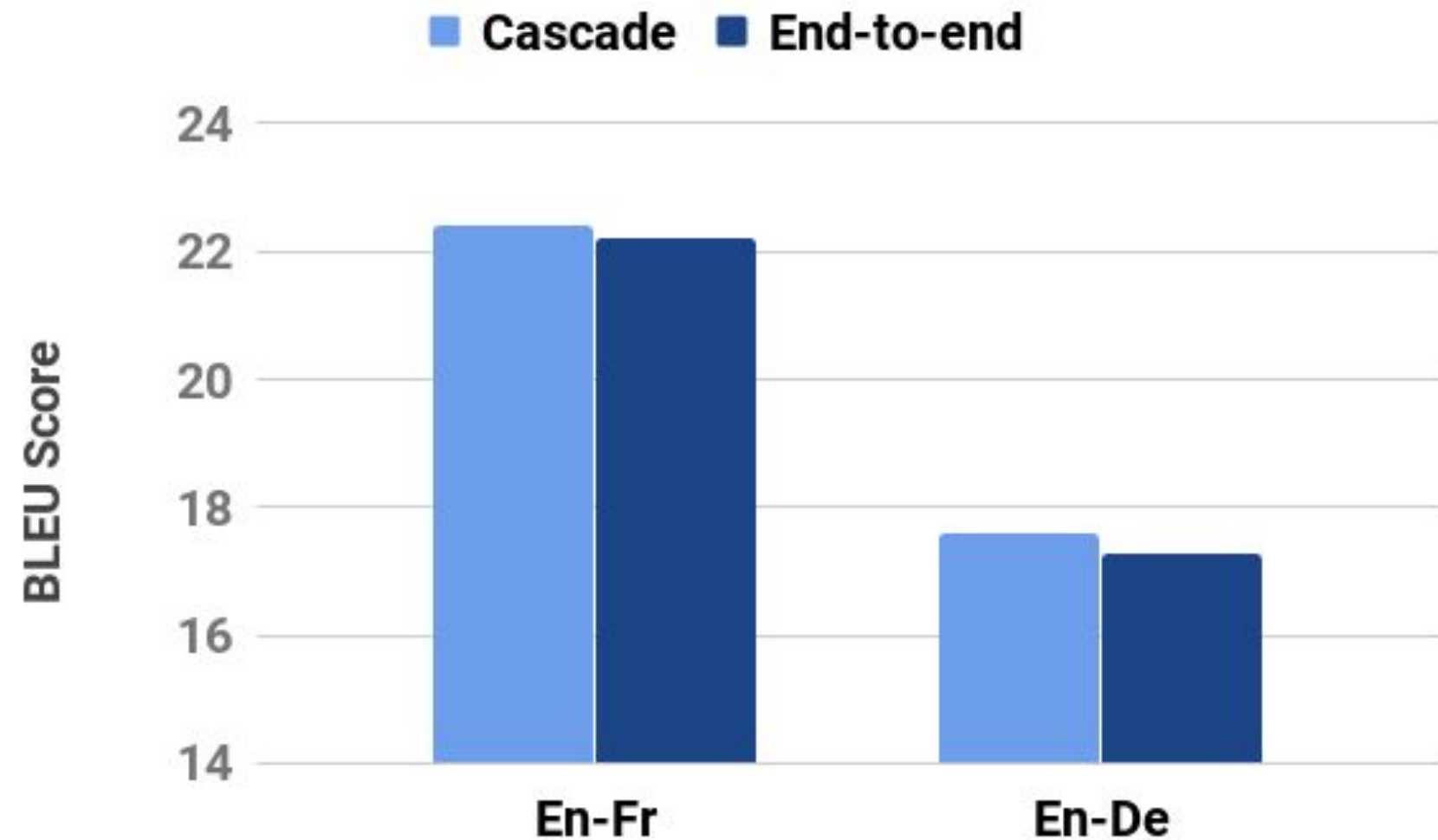
Automatic subtitling - Data

- **OpenSubtitles** (Lison and Tiedemann, 2016) -- 60 languages
 - Variable quality (professional/amateur subt., automatic sentence-level alignm.):
 - No information about subtitle breaks
 - No alignment with audio (mostly copyright-protected videos)
- **JESC** (Pryzant et al., 2018) -- Ja-En
 - Automatic alignments (caption level = only subtitles with matching timestamps)
 - No alignment with audio
- **Must-Cinema** (Karakanta et al., 2020) -- En→ 7 languages
 - Derived from MuST-C (TED talks)
 - Annotated with subtitle breaks
 - Audio-transcript-translation alignments

E2E subtitling: experiments on En-Fr/De

- **Doable?**

- Translation quality



E2E subtitling: experiments on En-Fr/De

- **Effective?**

- Segmentation (<eol> and <eob> insertion)

