Approaches to Machine Translation:
Rule-based, Statistical and Hybrid

*Motivation, History and Applications
of MT*

Do we need to motivate MT?

MT is commercially and academically interesting

Commercially	Academically
US has invested in MT for intelligence purposes	MT requires or benefits from many other NLP technologies: Parsing, word sense disambiguation, named entity recognition, transliteration…
MT is popular on the web.	MT involves machine learning challenges: working with large data sets; noise; optimization; search…
EU spends more than 1,000,000,000 euros on translation costs each year	MT is not going to be solved in the next years

MT- related commercial successes

Example of a state-of-the-art statistical machine translation

ENGLISH-TO-PORTUGUESE

SOURCE TEXT

We are starting a course in statistical machine translation. Is it really interesting to learn about the subject?

MT OUTPUT

Estamos começando um curso de estatística tradução automática. É realmente interessante para aprender sobre o assunto?

Georgetown-IBM Experiment

New York, January 7, 1954: Russian was translated into English by an electronic "brain" today for the first time. [...]

UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

MT, one of the first applications evisioned for computers

- 1947-1954   Information Theory Foundations

- 1954-1966   Large bilingual dictionaries +rules

- 1966-1980   ALPAC Report, research continued in Europe and Canada

SOURCE: **The history of machine translation in a nutshell** *John Hutchins*
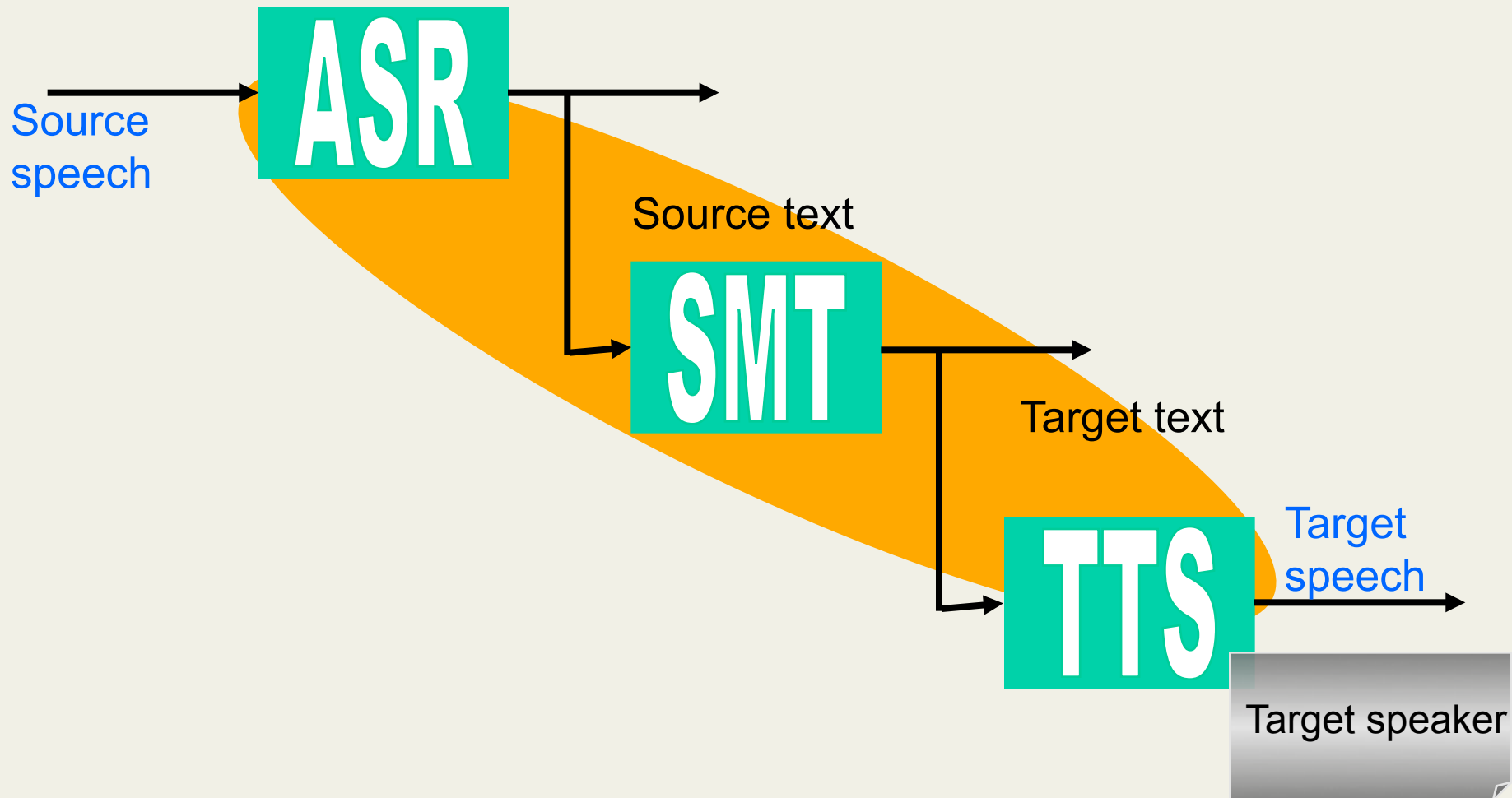
MT progressing from different perspectives

- 1980s   Variety of systems
- 1990s   Statistical Machine Translation
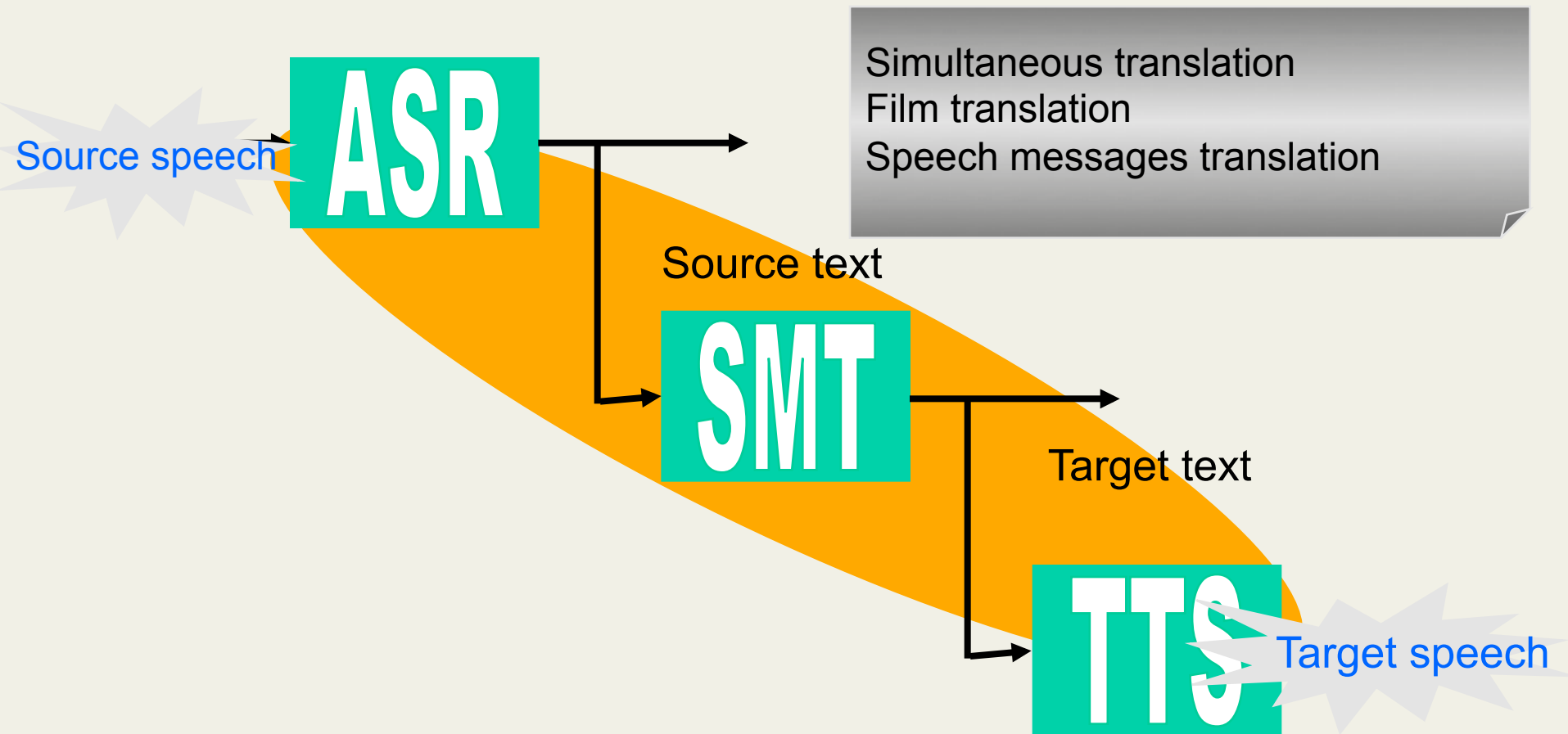- 2000s   MT software

What is going on now in MT?

- MT is consistently improving with resources and computational power

- Reasonable quality when resources available

- Moving to Hybrid Architectures with Statistical and Linguistic knowledge to cover what data does not cover
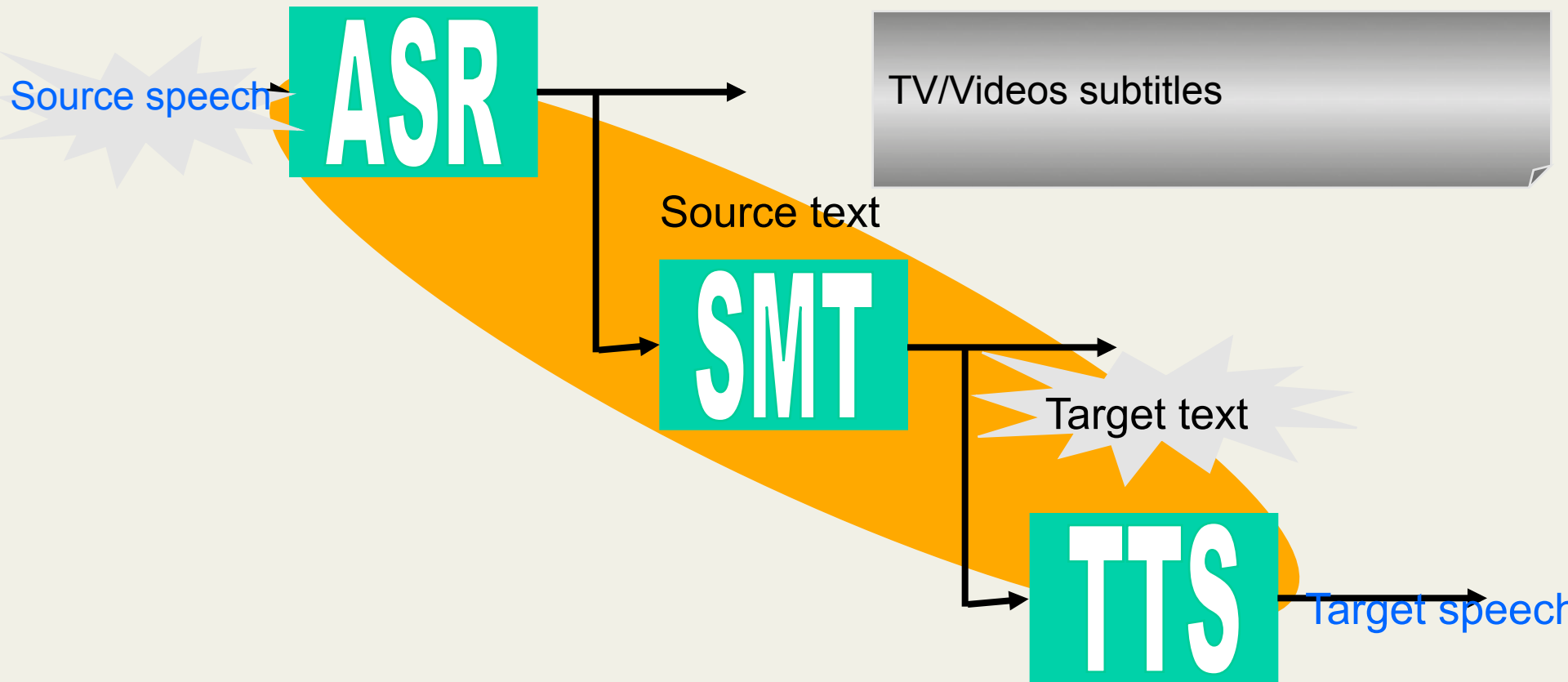
Related MT challenges
speech-to-speech translation

Source speech → **ASR** → Source text → **SMT** → Target text → **TTS** → Target speech

Target speaker

Applications (I)

Applications (II)

MT performance already quite acceptable in …

CLIR allows to search for [hotels in Paris](#) on French-language pages or [bars in Moscow](#) from Russian sources.

Computer-aided human translation

Communication via email, chat

Translation on hand-held devices

What next?

MT APPROACHES