Sec 3.2.2 **Transfer Learning** & Pretraining

Pre-training SLT components

Pre-training components of the SLT systems on different tasks

- Encoder pre-training (Bansal et al., 2018) <--> Automatic Speech Recognition
- Decoder pre-training (Bérard et al., 2018) <--> Machine Translation

Encoder Pre-training

Spanish Audio



English text

What a wonderful tutorial!

Encoder Pre-training



Training an ASR using the same SLT architecture

Spanish text

¡Qué maravilloso tutorial!

Encoder Pre-training



Training an ASR using the same SLT architecture

Training the SLT system initializing the encoder with the trained ASR encoder

English text

What a wonderful tutorial!



Decoder Pre-training

Spanish Audio



English text

What a wonderful tutorial!

Decoder Pre-training

Spanish text

¡Qué maravilloso tutorial!



What a wonderful tutorial!

Training an MT system using the same SLT architecture

English text



Decoder Pre-training



Training an MT system using the same SLT architecture

Training the SLT system initialising the decoder with the trained MT decoder

English text

What a wonderful tutorial!



Encoder-Decoder Pre-training



Training the SLT system initializing:

- the encoder with the trained ASR encoder
- the decoder with the trained MT decoder

English text

What a wonderful tutorial!

Exploiting unlabelled data

Following the trends in MT and text generation, exploiting unlabelled data

Exploiting unlabelled data

Following the trends in MT and text generation, exploiting unlabelled data

Integration of:

- Encoder pre-training based on a general-purpose acoustic models: wav2vect (Ly et al., 2020)
- Decoder pre-training based on general-purpose language models: BERT or mBART (Wu et al., 2020)

Exploiting unlabelled data

Following the trends in MT and text generation, exploiting unlabelled data

Integration of:

- Encoder pre-training based on a general-purpose acoustic models: wav2vect (Ly et al., 2020)
- Decoder pre-training based on general-purpose language models: BERT or mBART (Wu et al., 2020)

Useful in low-resourced and zero-shot conditions