

Modèles neuronaux pour la simplification de parole : application au sous-titrage

François Buet (groupe TLP, ED STIC)

Journée des doctorants 2019

Supervision : François Yvon et Alexandre Allauzen
Financement : projet ROSETTA (BPI)



Partenaires : MFP, MOCAPLAB, LIMSI, LUTIN

Développer des solutions d'accessibilité pour des contenus audiovisuels.

- Représentation en Langue des Signes
- Sous-titrage
 - ↪ simplification de parole



Produire le résumé d'une phrase, en équilibrant la réduction de la complexité et la préservation du sens et de la grammaticalité.

Exemple

Complexe	Il est dans la nature de l'Homme de se tromper, mais persévérer est diabolique.
Simple	Il est humain de se tromper, mais persévérer dans l'erreur est diabolique.

Complexité : lexicale, syntaxique ...



- Simplification de phrase par un Encodeur-Décodeur avec attention [Rush et al., 2015]
- Simplification de phrase par un modèle Transformer [Zhao et al., 2018]
- Traduction de parole sans transcription [Duong et al., 2016]

Contrôle de la longueur et analyse des représentations internes

Sondage des informations présentes dans des plongements de phrases [Adi et al., 2016].

Ajout extérieur de l'information de longueur au début ou au cours du décodage :

- LenInit, LenEmb [Kikuchi et al., 2016]
- LDPE, LRPE [Takase and Okazaki, 2019]

Approche type "disentanglement", via réseaux adverses, à la manière de [Lample et al., 2017].



- Transcription verbatim d'émissions télévisées, puis alignement avec les sous-titres afin de générer un corpus parallèle de travail.
- Analyse de la représentation de la longueur dans les méthodes de transduction.

References I

-  Adi, Y., Kermayn, E., Belinkov, Y., Lavi, O., and Goldberg, Y. (2016). Fine-grained analysis of sentence embeddings using auxiliary prediction tasks.
CoRR, abs/1608.04207.
-  Duong, L., Anastasopoulos, A., Chiang, D., Bird, S., and Cohn, T. (2016).
An attentional model for speech translation without transcription.
In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics : Human Language Technologies*, pages 949–959. Association for Computational Linguistics.

References II



Kikuchi, Y., Neubig, G., Sasano, R., Takamura, H., and Okumura, M. (2016).

Controlling output length in neural encoder-decoders.

In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 1328–1338. Association for Computational Linguistics.



Lample, G., Zeghidour, N., Usunier, N., Bordes, A., Denoyer, L., and Ranzato, M. (2017).

Fader networks : Manipulating images by sliding attributes.

In Guyon, I., von Luxburg, U., Bengio, S., Wallach, H. M., Fergus, R., Vishwanathan, S. V. N., and Garnett, R., editors, *Advances in Neural Information Processing Systems 30 : Annual Conference on Neural Information Processing Systems 2017, 4-9 December 2017, Long Beach, CA, USA*, pages 5969–5978.



References III



Rush, A. M., Chopra, S., and Weston, J. (2015).

A neural attention model for abstractive sentence summarization.
In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 379–389. Association for Computational Linguistics.



Takase, S. and Okazaki, N. (2019).

Positional encoding to control output sequence length.
CoRR, abs/1904.07418.



Zhao, S., Meng, R., He, D., Saptono, A., and Parmanto, B. (2018).

Integrating transformer and paraphrase rules for sentence simplification.

In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3164–3173. Association for Computational Linguistics.

