

Problématique de la traduction de bulletins anglais vers l'inuktitut

Philippe Langlais

RALI / DIRO / Université de Montréal

Atelier MITACS sur la Diffusion multi-format d'informations environnementales
Vendredi 14 janvier 2011

Réalisations : aligneur de mots

- ▶ Inuktitut = langue agglutinante

En fait, il ne viendra probablement pas tôt aujourd'hui
qaisaaliniacquunngikkaluaqpuq

- ▶ NUKTI (Gotti et al. 2005)

in regards to |^{c1} elders |^{c2} and |^{c3} youth
piijutigillugu innatuqait amma makkuttu

Réalisations : analyseur morphologique

- ▶ Développé dans le cadre de Morpho Challenge 2009 (Langlais 2009, Lavallée et Langlais, 2010)

(finnois)

In : kahvinjuojallekin

Out : kahvi + n + juo + ja + lle + kin

coffee + of + drink + er + for + also

also for [the] coffee drinker

- ▶ Pas encore testé sur l'inuktitut (finnois, turc, anglais, allemand)

Constat sur la traduction vers l'inuktitut de bulletins

- ▶ Pas de données spécifiques au domaine de la météo (quoi que ...)

- ▶ Difficile de développer une approche telle que celle déployée pour le système anglais ↔ français

Pistes : approche

- ▶ Mettre nos briques traductionnelles bout à bout sur les Hansards
(*alignement, sélection de données, entraînement de modèles*)
 - ▶ premier prototype

Pistes : approche

- ▶ Mettre nos briques traductionnelles bout à bout sur les Hansards
(*alignement, sélection de données, entraînement de modèles*)
 - ▶ premier prototype

- ▶ Étudier les analyseurs morphologiques de l'inuktitut

Pistes : approche

- ▶ Mettre nos briques traductionnelles bout à bout sur les Hansards (*alignement, sélection de données, entraînement de modèles*)
 - ▶ premier prototype

- ▶ Étudier les analyseurs morphologiques de l'inuktitut

- ▶ Étudier l'influence de la morphologie sur les performances
 - ▶ second prototype

Pistes : approche

- ▶ Mettre nos briques traductionnelles bout à bout sur les Hansards (*alignement, sélection de données, entraînement de modèles*)
 - ▶ premier prototype
- ▶ Étudier les analyseurs morphologiques de l'inuktitut
- ▶ Étudier l'influence de la morphologie sur les performances
 - ▶ second prototype
- ▶ Ajouter des données spécifiques au domaine de la météo
 - ▶ troisième prototype
 - ▶ quelles en sont les limites ?

Merci

ᑭᑦᑯᑦ ᑭᑦᑯᑦ ᑭᑦᑯᑦ