Interculturalité et localisation de l'humour et des jeux de mots

Contenu sous forme de paragraphes

Résumé

La communication interculturelle repose en grande partie sur la traduction. L'humour reste de loin l'un de ses aspects les plus difficiles. Pour comprendre l'humour, il faut souvent saisir les références culturelles implicites et/ou capter le double sens, ce qui soulève bien sûr la question de l'(in)traductibilité de l'humour. Les jeux de mots sont une source d'humour courante et sont utilisés par les romanciers et les poètes, ainsi que dans les titres et la publicité pour leur capacité à attirer l'attention ou leur caractère ludique et subversif. La traduction de l'humour et des jeux de mots est donc très demandée. La traduction moderne est fortement assistée par des outils technologiques. Cependant, peu de travaux ont étudié l'automatisation de la traduction de l'humour et des jeux de mots. À notre connaissance, il n'existe pas de corpus parallèle. Le projet JokeR aborde la question de l'identité européenne à travers l'étude de l'humour dans une perspective interculturelle. L'objectif principal du projet JokeR est d'étudier les stratégies de localisation de l'humour et des jeux de mots et de créer un corpus parallèle multilingue, annoté selon ces stratégies, ouvert et disponible librement. Le corpus parallèle multilingue issu du projet JokeR constituera une avancée dans l'automatisation de la localisation de l'humour afin d'entraîner et d'évaluer les modèles de traduction automatique. Ce corpus pourrait également être utile pour les étudiants en traduction.

Abstract

Intercultural communication relies heavily on translation. Humor remains by far one of its most difficult aspects; to understand humor, one often has to grasp implicit cultural references and/or capture double meanings, which of course raises the question of the (in)translatability of humor. Puns are a common source of humor and are used by novelists, poets, and playwrights, as well as in titles, headlines, toponyms, anthroponyms, organizations, and advertising for their attention-getting or mnemonic, playful, subversive, etc. values. The translation of humor and puns is therefore in high demand. Modern translation is heavily aided by technological tools, yet few works have studied the automation of humor and puns translation and the creation of humor corpora. To the best of our knowledge, there does not exist any parallel corpus. The JokeR project addresses the issue of European identity through the study of humor in a cross-cultural perspective. The main objective of the JokeR project is to study the strategies of localization of humor and puns and to create a multilingual parallel corpus, annotated according to these strategies, open and freely available. The multilingual parallel corpus resulting from the JokeR project is a step forward in the automation of humor localization in order to train and evaluate machine translation models. This corpus might also be useful for translation students.

Equipe

Élise Mathurin
linguistique anglaise
EA 4249 HCTI

Gaelle Le Corre
linguistique anglaise
EA 4451 CRBC

Radia Hannachi
linguistique française
EA 4249 HCTI

Sílvia Araújo
humanités numériques
Université du Minho
Département d'Études Romanes

Mohamed Saki
linguistique anglaise
EA 4249 HCTI

Publications

Search Results *--> Url version détaillée* , *Url version formatée*
Criteria : Id : "hal-04299292;hal-04608078;hal-04196754;hal-04161894;hal-03795272;hal-04161789"
Number of occurrences founded : 6.

titre
The JOKER Corpus: English-French Parallel Data for Multilingual Wordplay Recognition
auteur
Liana Ermakova, Anne-Gwenn Bosser, Adam Jatowt, Tristan Miller
article
SIGIR '23: The 46th International ACM SIGIR Conference on Research and Development in Information Retrieval, Jul 2023, Taipei Taiwan, France. pp.2796-2806, ⟨10.1145/3539618.3591885⟩

titre
Overview of JOKER – CLEF-2023 Track on Automatic Wordplay Analysis
auteur
Liana Ermakova, Tristan Miller, Anne-Gwenn Bosser, Victor Manuel Palma Preciado, Grigori Sidorov, Adam Jatowt
article
Experimental IR Meets Multilinguality, Multimodality, and Interaction, 14163, Springer Nature Switzerland, pp.397-415, 2023, Lecture Notes in Computer Science, 978-3-031-42447-2. ⟨10.1007/978-3-031-42448-9_26⟩

titre
Science for Fun: The CLEF 2023 JOKER Track on Automatic Wordplay Analysis
auteur
Liana Ermakova, Tristan Miller, Anne-Gwenn Bosser, Victor Manuel Palma Preciado, Grigori Sidorov, Adam Jatowt
article
Advances in Information Retrieval - 45th European Conference on Information Retrieval, ECIR 2023, Apr 2023, Dublin, Ireland. pp.546-556, ⟨10.1007/978-3-031-28241-6_63⟩

titre
Humorous Wordplay Generation in French
auteur
Loïc Glémarec, Anne-Gwenn Bosser, Julien Boccou, Liana Ermakova
article
13th Conference and Labs of the Evaluation Forum (CLEF 2022), Sep 2022, Bologna, Italy. pp.1719-1726

titre

Poetic or Humorous Text Generation: Jam Event at PFIA2022

auteur

Anne-Gwenn Bosser, Liana Ermakova, Florence Dupin de Saint-Cyr, Pierre de Loor, Victor Charpenay, Nicolas Pépin-Hermann, Benoît Alcaraz, Jean-Victor Autran, Alexandre Devillers, Juliette Grosset, Aymeric Hénard, Florian Marchal

article

13th Conference and Labs of the Evaluation Forum (CLEF 2022), Sep 2022, Bologna, Italy. pp.1719-1726

titre

Overview of JOKER@CLEF 2022: Automatic Wordplay and Humour Translation Workshop

auteur

Liana Ermakova, Tristan Miller, Fabio Regattin, Anne-Gwenn Bosser, Claudine Borg, Élise Mathurin, Gaëlle Le Corre, Sílvia Araújo, Radia Hannachi, Julien Boccou, Albin Digue, Aurianne Damoy, Benoît Jeanjean

article

Experimental IR Meets Multilinguality, Multimodality, and Interaction, 13390, Springer International Publishing, pp.447-469, 2022, Lecture Notes in Computer Science, ⟨10.1007/978-3-031-13643-6_27⟩

Axe de recherche

Identités et transitions européennes

Responsables

Liana Ermakova

liana.ermakova@univ-brest.fr

EA 4249 HCTI

Université de Bretagne Occidentale

Discipline(s)

linguistique de corpus

traduction

Humanités numériques

interculturalité

traitement automatique des langues naturelles

Mots clés thématiques

humour, jeux de mots, corpus, traduction automatique, identité européenne

Dates du projet

1 septembre 2021 - 31 août 2023