Actes de CORIA-TALN 2026
Nantes, France, 29 juin au 3 juillet 2026

La traduction du point de vue de l'Information Bottleneck : une analyse d'efficience des prépositions spatiales dans des bitextes

Antoine Taroni Ludovic Moncla Frédérique Laforest
INSA Lyon, CNRS, Lyon 1 Université, LIRIS, UMR5205, 69621 Villeurbanne, France
prénom.nom@insa-lyon.fr

RÉSUMÉ

Une communication efficiente requiert un équilibre entre informativité et simplicité. L'*Information Bottleneck* (IB) formalise ce compromis en termes de théorie de l'information. Bien que ce cadre soit largement étayé dans des domaines visuels (couleurs, mouvement), l'étude de stimuli linguistiques comme des mots en contexte phrastique reste inexplorée. Nous proposons dans cet article une nouvelle méthodologie computationnelle comblant cette lacune, en dressant un parallèle entre le processus de traduction, et la question d'optimisation au sens de l'IB. Nous avons appliqué cette méthodologie aux prépositions spatiales de bitextes français, anglais, allemands et serbes. Une étude pilote de tri de cartes ($N = 35$) nous a permis de prédire l'informativité des traductions de prépositions. Les résultats suggèrent que les traductions authentiques sont plus efficientes que des variantes hypothétiques. Plus largement, nous proposons d'exploiter les données de traduction afin d'étudier les pressions cognitives qui façonnent les systèmes sémantiques naturels et artificiels.

ABSTRACT

Translation from the Information Bottleneck perspective : an efficiency analysis of spatial prepositions in bitexts

Efficient communication requires balancing informativity and simplicity when encoding meanings. The Information Bottleneck (IB) captures this trade-off from an information-theoretic perspective. While supported in visual domains (colours, motion), linguistic stimuli such as words in sentential context remain unexplored. We address this gap by framing translation as an IB optimisation problem, treating source sentences as stimuli and target sentences as compressed meanings. We applied this to spatial prepositions in French, English, German and Serbian bitexts. To estimate informativity, we conducted a pile-sorting pilot-study ($N = 35$). Results from our IB analysis suggest that attested translations of prepositions are more efficient than counterfactual alternatives. More broadly, this work suggests that translation can serve as a window into the cognitive efficiency pressures shaping natural and artificial semantic systems.

MOTS-CLÉS : Théorie de l'Information, Langage Spatial, Cognition Spatiale, Information Bottleneck, Efficience du Langage, Traduction.

KEYWORDS: Information Theory, Spatial Language, Spatial Cognition, Information Bottleneck, Linguistic Efficiency, Translation.

ARTICLE ACCEPTÉ À : CMCL@LREC2026 : Workshop on Cognitive Modeling and Computational Linguistics, Palma, Mallorca, Spain, May 11-16, 2026.
URL : http://lrec-conf.org/proceedings/lrec2026/workshops/cmcl/2026.cmcl-1.0.pdf. TRADUCTION EN FRANÇAIS : https://hal.science/hal-05619899

Actes des 33ème Conférence sur le Traitement Automatique des Langues Naturelles.
Volume 2 : articles déjà publiés, pages 26–26.
Cette œuvre est mise à disposition sous licence Attribution 4.0 International.