
Préface

Ce numéro de la revue TAL inaugure un nouveau type d'appel à soumissions, consacré aux articles proposant un état de l'art dans le champ du traitement des langues.

La possibilité de soumettre un article de ce type était théoriquement ouverte par l'appel Varia (au même titre que les *position papers*), mais peu d'auteurs se saisissaient jusque-là de ce format. L'idée de proposer ce nouveau type de numéro, qui est appelé à être proposé régulièrement par la revue désormais, est née de plusieurs constats. Tout d'abord, la communauté scientifique concernée par la revue est très diverse, confirmant le caractère fortement interdisciplinaire du champ : un questionnaire diffusé cette année sur la liste de diffusion LN a montré que les répondants se répartissaient de la manière suivante : 38 % en informatique, 39,5 % en linguistique informatique, 14 % en linguistique, 8,5 % autres (humanités numériques, éducation, etc.). Les articles publiés par la revue et les intitulés des numéros thématiques sont également, et naturellement, le reflet de cette diversité : citons par exemple, sur les dernières années, les thématiques « diversité linguistique en TAL », « apprentissage profond pour le TAL », « dialogue et systèmes de dialogue », « TAL et humanités numériques ». Il est donc important que ces sous-communautés continuent à partager leurs questions de recherche malgré la diversité de leurs intérêts et de leur expertise. Par ailleurs, le champ du TAL interagit avec des communautés connexes (parole, image, recherche d'information, mathématiques, linguistique, cognition, etc.), qui sont intéressées par les résultats et les méthodes que le TAL produit. Enfin, le champ du traitement automatique des langues a évolué ces dernières années à un rythme très soutenu, nécessitant une mise à jour constante de la part des chercheurs. Tous ces constats ont amené le comité de rédaction de la revue à proposer un numéro dédié aux états de l'art, pour faciliter la circulation et l'appropriation des connaissances en TAL au sein de sa communauté et à destination d'autres communautés scientifiques.

L'appel à soumissions a sollicité des articles rendant compte des travaux les plus importants dans le domaine choisi par les auteurs, et présentant l'évolution de ce domaine pour aboutir aux avancées les plus récentes. La synthèse devait être rigoureuse, claire et accessible pour des lecteurs de la revue non spécialistes de la thématique de l'article. Elle devait proposer une mise en perspective des travaux présentés, permettant de comprendre leur articulation. Les thématiques couvertes étaient celles habituellement ciblées par les numéros Varia de la revue, soit tous les aspects du traitement

automatique des langues écrites, parlées et signées et de la linguistique computationnelle.

Le présent numéro contient les trois articles retenus lors de l'appel non thématique lancé en mars 2022 et clos à la mi-juillet 2022. Huit articles avaient été soumis, soit un taux de sélection de 37,5 %, comparable au taux moyen de la revue sur les dernières années. La moitié des premiers auteurs de ces soumissions initiales étaient issus de laboratoires hors de France. Signalons que l'appel demandait aux auteurs d'informer préalablement de leur intention à soumettre, de façon à permettre aux éditeurs d'anticiper la recherche de relecteurs experts. Dix-neuf résumés ont été envoyés dans cette phase initiale, ce qui, à la fois, indique un intérêt important pour ce type de numéro, et montre la nécessité de mieux calibrer la procédure pour réduire ce décalage entre le nombre d'intentions et le nombre de soumissions effectives. Chaque article, comme il est d'usage dans la revue, a été évalué en double aveugle et expertisé par trois relecteurs, dont deux relecteurs externes spécialistes du champ étudié, et un relecteur interne au comité de rédaction de la revue.

Les trois articles retenus sont les suivants :

– *Discourse annotation - Towards a dialogue system for pair programming*, Cecilia Domingo, Paul Piwek, Svetlana Stoyanchev et Michel Wermelinger (The Open University, Toshiba Europe Limited, UK) : cet article offre une présentation des systèmes de dialogue visant à assister la programmation (*programming-oriented dialog systems*);

– *Fillers in Spoken Language Understanding : Computational and Psycholinguistic Perspectives*, Tanvi Dinkar, Chloé Clavel et Ioana Vasilescu (Heriot Watt University, Télécom Paris, Université Paris-Saclay) : cet article propose une synthèse des recherches menées sur les disfluences et leur impact sur la compréhension de la langue parlée, en focalisant l'étude sur le cas des disfluences à un seul token (*filler*);

– *Les corpus arborés avant et après le numérique*, Sylvain Kahane et Nicolas Mazziotta (Modyco, Université Paris Nanterre & CNRS, U.R. Traverses, Université de Liège) : il s'agit d'une présentation du rôle et de l'usage des corpus arborés au fil des siècles, depuis l'époque pré-numérique jusqu'à nos jours.

On trouvera à la suite de ces articles des notes de lecture ainsi que des résumés de thèses et habilitations à diriger des recherches récemment soutenues en TAL. Un grand merci à Denis Maurel pour le soin accordé à la préparation de la rubrique *Notes de lecture* au fil des années. Nous encourageons désormais nos lecteurs à contacter Didier Schwab (didier.schwab@univ-grenoble-alpes.fr), qui prend le relais de Denis sur cette rubrique. Pour ce qui est des résumés de thèses ou d'HDR, nous vous invitons à les transmettre à Sylvain Pogodalla (sylvain.pogodalla@inria.fr) afin d'informer notre communauté. Merci à Sylvain pour son travail de veille et de collecte.

Nous remercions les membres du comité de rédaction de la revue qui ont participé aux différentes étapes d'élaboration de ce numéro, et en particulier ceux qui ont pris en charge des relectures (voir la composition du comité sur le site de la revue : <https://www.atala.org/content/comit%C3%A9-de-r%C3%A9daction-0>), ainsi que les

relecteurs spécifiques de ce numéro : Alexandre Allauzen, Jean-Yves Antoine, Frédéric Béchet, Béatrice Daille, Nuria Gala, Nabil Hathout, Nicolas Hernandez, Cyril Labbé, Frédéric Landragin, Gaël Lejeune, Aleksandra Miletic, Ludovic Tanguy, Guillaume Wisniewski.

Nous remercions enfin l'Institut des sciences humaines et sociales (INSHS) du CNRS pour le soutien financier apporté à la revue.

Cécile Fabre
CLLE, Université Toulouse 2
cecile.fabre@univ-tlse2.fr

Emmanuel Morin
LS2N, Université de Nantes
emmanuel.morin@univ-nantes.fr

Sophie Rosset
Université Paris-Saclay, CNRS, LISN
sophie.rosset@lisn.fr

Pascale Sébillot
IRISA, INSA Rennes
pascale.sebillot@irisa.fr