
Préface

Les préfaces des numéros non thématiques de la revue *TAL* permettent de faire chaque année le point sur la vie de la revue et de commenter les statistiques que nous présentons traditionnellement pour les numéros des trois dernières années.

La revue a connu un renouvellement important de son comité de rédaction : après le départ en 2023 de Pascale Sébillot, Sophie Rosset a terminé son mandat avec le numéro 64-2, Emmanuel Morin avec le numéro 65-1, et Cécile Fabre avec ce numéro 66-1. Le quatuor s'est reconstitué autour de Maxime Amblard et Benoit Favre, rejoints en janvier 2025 par Marie Candito et Aurélie Névéol. Les rédacteurs en chef poursuivent la réflexion sur les évolutions requises pour assurer une meilleure visibilité à la revue. Après le passage au fil de l'eau des *Varia* et l'attribution d'un DOI à chaque article qui va prendre effet à partir du volume 64, la question des modalités de diffusion de la revue est en cours d'examen. Nous explorons plusieurs pistes permettant d'héberger la revue sur une plateforme d'édition en libre accès (type OpenEdition ou Mersenne).

Les statistiques du tableau 1 considèrent les dix derniers numéros sur les trois dernières années, du début de 2022 jusqu'à ce numéro *Varia* de 2025 inclus. Ce tableau donne les taux de sélection par numéro et par volume. La ligne du total synthétise ces chiffres sur l'ensemble des dix numéros considérés. L'attractivité des numéros (entre 2 et 8 propositions) reste assez faible, avec un nombre moyen d'articles soumis un peu plus élevé pour les numéros thématiques que pour les *Varia* (6 vs 4,5). Le comité de rédaction de la revue est très attaché à sélectionner les articles sur le seul critère de leur qualité, indépendamment du nombre d'articles soumis, et n'hésite pas, comme cela a été le cas pour deux des derniers *Varia*, à préserver cette exigence et à ne retenir qu'un article. Le taux de sélection est de fait très variable, allant de 25 à 75 %. Le choix des deux thématiques annuelles est fait chaque année en comité de rédaction. Un appel à proposition de thématiques est lancé sur la liste *In* en début d'année.

Les statistiques que nous donnons, dans le tableau 2, sur l'origine des articles, prennent en compte le pays du premier auteur, France ou hors de France. Nous considérons également la langue de la soumission, le français ou l'anglais. Les chiffres sont fournis pour la même période de temps que pour le tableau 1. Globalement, un cinquième des premiers auteurs sont des chercheurs hors de France, et près d'un tiers des articles sont en anglais. Néanmoins, ces chiffres cachent des disparités fortes selon les

Intitulé	Vol.	N°	Année	Soumis	Acceptés	% acceptés
- <i>Varia</i>	63	1	2022	5	3	60,0 %
- Intermodalité et multimodalité en TAL	63	2	2022	4	3	75,0 %
- États de l'art en TAL	63	3	2022	8	3	37,5 %
Sous-total	63		2022	17	9	52,9 %
- <i>Varia</i>	64	1	2023	4	1	25,0 %
- Robustesse et limites des modèles de TAL	64	2	2023	4	2	50,0 %
- Explicabilité des modèles de TAL	64	3	2023	5	5	100,0 %
Sous-total	64		2023	13	8	61,5 %
- <i>Varia</i>	65	1	2024	1	1	100,0 %
- Scholarly Document Processing	65	2	2024	5	1	20,0 %
- Discours de haine	65	3	2024	8	3	37,5 %
Sous-total	65		2024	14	5	35,7 %
<i>Varia</i>	66	1	2025	7	2	28,5 %
Total	Dix derniers n°s			51	24	47,0 %

Tableau 1. Taux de sélection aux appels de la revue TAL sur les dix derniers numéros de la période 2017-2020

Intitulé	Vol.	N°	Année	% 1 ^{er} auteur hors France	% en anglais
- <i>Varia</i>	63	1	2022	0,0 %	33,3 %
- Intermodalité et multimodalité en TAL	63	2	2022	0,0 %	0,0 %
- États de l'art en TAL	63	3	2022	66,6 %	66,6 %
Pourcentages par volume	63		2022	22,2 %	33,3 %
- <i>Varia</i>	64	1	2023	0,0 %	0,0 %
- Robustesse et limites des modèles de TAL	64	2	2023	0,0 %	0,0 %
Explicabilité des modèles de TAL	64	3	2023	20,0 %	20,0 %
Pourcentages par volume	64		2023	12,5 %	12,5 %
<i>Varia</i>	65	1	2024	0,0 %	0,0 %
Traitement automatique de documents scientifiques	65	2	2024	0 %	0 %
Discours de haine	65	3	2024	66 %	66 %
Pourcentages par volume	65		2024	40 %	40 %
<i>Varia</i>	66	1	2025	0 %	50 %
Pourcentages totaux	Dix derniers n°s			20,8 %	29,1 %

Tableau 2. Proportion des articles publiés d'un premier auteur hors de France et proportion des articles publiés rédigés en anglais sur les dix derniers numéros de la période 2022-2025

numéros : si certains numéros thématiques ont attiré plus de contributions internationales (c'est le cas des numéros sur les états de l'art et sur les discours de haine), les *Varia* ont intégré exclusivement des articles rédigés par des chercheurs français. Rappelons que, depuis 2024, les articles peuvent être rédigés indifféremment en anglais ou en français, alors que jusqu'à cette date les chercheurs francophones devaient opter pour le français. La dimension internationale de la revue se manifeste également à travers la composition du comité de rédaction de la revue, la constitution d'un comité scientifique international pour les numéros thématiques, la présence d'un chercheur étranger parmi les rédacteurs invités en charge de ces numéros.

Le présent numéro contient deux articles retenus lors de l'appel non thématique lancé en janvier 2024. Cet appel portait, comme c'est l'usage, sur tous les aspects du traitement automatique des langues. Sept articles ont été soumis entre septembre et décembre 2024. Deux articles ont été retenus à l'issue du processus de sélection habituel à deux tours.

Automatic characterization of French language registers: Illustration on tweets, Jade Mekki (U. Rennes, CNRS, IRISA) — Delphine Battistelli (U. Paris Nanterre, CNRS, MoDyCo), Nicolas Béchet (U. Rennes, CNRS, IRISA), Gwénolé Lecorvé (Orange Labs / U. de Bretagne Sud, CNRS, IRISA)

L'article propose une méthode pour détecter et pour analyser les variations de registre (familier, standard, formel) dans un large corpus de tweets en français. L'annotation d'un échantillon du corpus permet d'entraîner un classifieur pour généraliser l'annotation, afin d'explorer les caractéristiques linguistiques qui distinguent chaque paire de registre. La méthode fait appel à une technique d'extraction de motifs qui rend possible la caractérisation des registres de langue à différents niveaux d'analyse linguistique.

De l'annotation intégrale à l'analyse des réseaux de personnages : un modèle pour la REN dans les textes littéraires en français, Motasem Alrahabi (ObTIC, Sorbonne U.) — Arthur Amalvy (LIA, Avignon U.) — Vincent Labatut (LIA, Avignon U.) — Perrine Maurel (ObTIC, Sorbonne U.)

Cet article présente une contribution du TAL au champ des humanités numériques, et plus particulièrement à l'analyse littéraire. L'étude porte sur la reconnaissance d'entités nommées (EN) dans des textes littéraires en français. Elle se fonde sur l'annotation en EN d'un corpus de sept oeuvres du XIX^e siècle, librement disponible, et sur l'utilisation d'un système de reconnaissance d'EN à base d'un modèle Camembert préentraîné puis fine-tuné sur ce corpus littéraire. L'article montre comment ces résultats peuvent produire des réseaux de personnages favorisant l'analyse des dynamiques littéraires.

Merci aux membres du comité de rédaction de la revue qui ont participé aux différentes étapes d'élaboration de ce numéro et à la décision collégiale, à l'appui des relectures reçues, qui a abouti à l'acceptation ou au rejet des articles soumis. Merci en

particulier à celles et ceux qui ont pris en charge des relectures (voir la composition du comité sur le site de la revue : <http://www.atala.org/content/comité-de-rédaction-0>).
Merci aux relecteurs spécifiques de ce numéro : Timothée Bernard (LLF), Mathieu Lafourcade (LIRMM), Philippe Langlais (U. de Montréal), Gaël Lejeune (STIH, CERES), Filip Miletic (U. de Stuttgart), Thierry Poibeau (LATTICE), Céline Poudat (BCL).

Nous adressons nos remerciements à la Délégation générale à la langue française et aux langues de France (DGLFLF) pour son soutien financier à la revue.

Maxime Amblard
LORIA, Université de Lorraine

Cécile Fabre
CLLE, Université Toulouse - Jean Jaurès

Benoît Favre
LIS, Aix-Marseille Université

Sophie Rosset
Université Paris-Saclay, CNRS, LISN