

---

## 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.

Différentes actions visant l'amélioration de la visibilité de la revue ont été entreprises ou développées au cours de l'année 2021. Les pages de la revue *TAL* sur le site de l'ATALA ont ainsi été traduites intégralement en anglais, et l'accès aux anciens articles, mis à mal par un changement des URL lors de la refonte du site en 2019, a été consolidé. Concernant l'indexation, des métadonnées manquantes ont été complétées et un convertisseur vers des formats attendus par des sites de référencement est en cours de mise au point ; par ailleurs une indexation *via* ACL Ontology et un dépôt automatisé des articles sur HAL sont à l'étude. Nous remercions Gaël Guibon, Damien Nouvel et Philippe Muller pour leurs contributions respectives sur ces aspects.

La vie du comité a été marquée, lors du renouvellement annuel des membres, par l'accueil de quatre femmes, qui nous a permis d'atteindre pour la première fois une parité en genre. Parmi les autres sujets dont nous pouvons nous réjouir, la régularité de la publication de la revue a été maintenue, en nous fondant sur notre calendrier prévisionnel qui nous permet de caler les différents numéros d'un même volume et de tenir de façon régulière les réunions du comité de rédaction. Rappelons que l'une des caractéristiques de notre revue, à laquelle nous sommes foncièrement attachés, est la tenue des réunions du comité de rédaction – en visioconférence – au cours desquelles, entre autres, nous décidons collégalement, à l'appui des relectures reçues, de l'acceptation ou du rejet des articles soumis.

L'attractivité de la revue a été, au cours de cette année, le sujet de plusieurs discussions lors de ces réunions du comité de rédaction. Dès 2020, la diminution du nombre de soumissions aux derniers numéros *Varia* avait soulevé la question de la pertinence du maintien d'un tel numéro annuel. La décision avait été prise de se donner le temps de vérifier la persistance de cette tendance, que les chiffres du numéro 62:1 semblent confirmer (sept soumissions ; *cf.* tableau 1 *infra*). Les discussions de 2021 ont plus spécifiquement concerné les numéros thématiques, dont l'attractivité varie fortement d'un numéro à l'autre (de quatre à vingt-trois soumissions si l'on considère les trois

derniers volumes). La décision a été prise de proposer en 2021 deux numéros à thématiques assez larges, afin de tester si les appels pouvaient attirer un public plus large. Il semblerait que ce ne soit pas le cas, puisque le numéro portant sur les nouvelles applications du TAL n'a abouti qu'à deux soumissions et celui sur la diversité linguistique à trois. Le comité de rédaction de la revue doit donc continuer à se pencher sur cette question de l'attractivité, tout en gardant à l'esprit que les chiffres de ces deux dernières années peuvent être liés à la crise sanitaire, aux confinements successifs et à la tenue des conférences à distance.

<b>Intitulé</b>	<b>Vol.</b>	<b>N°</b>	<b>Année</b>	<b>Soumis</b>	<b>Acceptés</b>	<b>% acceptés</b>
<i>Varia</i>	59	1	2018	10	3	30,0 %
Apprentissage profond pour le TAL	59	2	2018	8	3	37,5 %
Trait. auto. des langues peu dotées	59	3	2018	23	4	17,4 %
<b>Sous-total</b>	<b>59</b>		<b>2018</b>	<b>41</b>	<b>10</b>	<b>24,4 %</b>
<i>Varia</i>	60	1	2019	8	1	12,5 %
Corpus annotés	60	2	2019	6	3	50,0 %
TAL et humanités numériques	60	3	2019	13	5	38,5 %
<b>Sous-total</b>	<b>60</b>		<b>2019</b>	<b>27</b>	<b>9</b>	<b>33,3 %</b>
<i>Varia</i>	61	1	2020	8	1	12,5 %
TAL et santé	61	2	2020	4	3	75,0 %
Dialogue et systèmes de dialogue	61	3	2020	5	3	60,0 %
<b>Sous-total</b>	<b>61</b>		<b>2020</b>	<b>17</b>	<b>7</b>	<b>41,2 %</b>
<i>Varia</i>	62	1	2021	7	2	28,6 %
<b>Total</b>			<b>Dix derniers n<sup>os</sup></b>	<b>92</b>	<b>28</b>	<b>30,4 %</b>

**Tableau 1.** Taux de sélection aux appels de la revue TAL sur les dix derniers numéros de la période 2018-2021

Les statistiques que nous présentons dans le tableau 1 considèrent les dix derniers numéros sur les trois dernières années, en l'occurrence donc du début de 2018 jusqu'à ce numéro *Varia* de 2021 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. Le taux de sélection sur l'ensemble de ces numéros s'élève à 30,4 % en moyenne, avec d'assez grandes variations selon les numéros (de 12,5 % pour les deux *Varia* de 2019 et 2020, dont les chiffres sont rigoureusement identiques, à 75 % pour le numéro dédié à la thématique TAL et santé).

Rappelons qu'un numéro ne peut pas excéder cinq articles, pour des raisons liées au coût du processus d'édition. Ce seuil a été atteint une fois seulement dans la période, pour le numéro consacré aux humanités numériques.

Intitulé	Vol.	N°	Année	% 1 <sup>er</sup> auteur hors France	% en anglais
Varia	59	1	2018	0,0 %	0,0 %
Apprentissage profond pour le TAL	59	2	2018	33,3 %	66,6 %
Trait. auto. des langues peu dotées	59	3	2018	25,0 %	0,0 %
<b>Pourcentages par volume</b>	<b>59</b>		<b>2018</b>	<b>20,0 %</b>	<b>20,0 %</b>
Varia	60	1	2019	0,0 %	0,0 %
Corpus annotés	60	2	2019	0,0 %	0,0 %
TAL et humanités numériques	60	3	2019	40,0 %	40,0 %
<b>Pourcentages par volume</b>	<b>60</b>		<b>2019</b>	<b>22,2 %</b>	<b>22,2 %</b>
Varia	61	1	2020	0,0%	0,0%
TAL et santé	61	2	2020	33,3 %	33,3 %
Dialogue et systèmes de dialogue	61	3	2020	66,7 %	66,7 %
<b>Pourcentages par volume</b>	<b>61</b>		<b>2020</b>	<b>42,9 %</b>	<b>42,9 %</b>
Varia	62	1	2021	0,0%	0,0%
<b>Pourcentages totaux</b>	<b>Dix derniers n<sup>os</sup></b>			<b>25,0 %</b>	<b>25,0 %</b>

**Tableau 2.** Proportion des articles publiés d'un premier auteur d'un laboratoire hors de France et proportion des articles publiés rédigés en anglais sur les dix derniers numéros de la période 2018-2021. Attention, les pourcentages totaux ne sont pas de simples moyennes des chiffres donnés plus haut, car les dénominateurs changent.

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 dans le cas des trois derniers *Varia*, à préserver cette exigence malgré un nombre de soumissions limité.

Les statistiques que nous donnons dans le tableau 2 sur l'origine des articles considèrent le pays du laboratoire du premier auteur, hors de France ou pas, ainsi que la langue de la soumission, le français ou l'anglais, l'anglais n'étant possible que si l'un des coauteurs n'est pas francophone. Les chiffres sont fournis pour la même période de temps que le tableau 1. Globalement, un quart des premiers auteurs sont des chercheurs de laboratoires hors de France, et un quart des articles sont en anglais. Néanmoins, ces chiffres cachent des disparités fortes selon les numéros, avec, pour chacune des trois années considérées, un numéro qui se singularise et attire plus nettement les chercheurs étrangers (apprentissage profond, TAL et humanités numériques, dialogue et systèmes de dialogue). Contrairement à des années plus anciennes (2016 ou 2017 par exemple), les numéros *Varia* n'attirent ou ne retiennent pas d'articles ayant ce caractère international. Ce dernier point soulève deux remarques fortement liées : d'une part, il pourrait être intéressant d'introduire dans nos statistiques le décompte des soumissions à caractère international – et non juste celles retenues pour

publication –, afin de mieux refléter l’attractivité hors de nos frontières de la revue TAL, d’autre part, à l’aune de ces nouveaux nombres spécifiquement pour les numéros *Varia*, il serait possible d’établir si ces numéros non thématiques répondent ou non à une attente à l’international.

Le présent numéro contient les articles retenus suite à l’appel non thématique lancé en juillet 2020. Cet appel portait comme d’habitude sur tous les aspects du traitement automatique des langues. Sept articles ont été soumis. À l’issue du processus de sélection, deux articles ont été acceptés pour publication :

– *Génération automatique de texte en langage naturel pour les systèmes de questions-réponses*, Imen Akermi, Johannes Heinecke, Frédéric Herledan (Orange Innovation) : cet article propose une méthode non supervisée permettant de générer des réponses concises mais complètes à des questions en langue naturelle ;

– *Étude des dépendances syntaxiques non projectives en français*, Guy Perrier (LORIA, Université de Lorraine) : en se fondant sur deux corpus annotés selon deux formats distincts, cet article présente une étude des constructions du français dont la structure syntaxique fait l’objet d’une représentation en dépendances non projective (caractère local du croisement de dépendances, etc.).

On trouvera à la suite de ces articles des notes de lecture rassemblées par Denis Maurel. Nous encourageons nos lecteurs à utiliser ce moyen pour faire profiter de leurs lectures la communauté. Suit une liste de résumés de thèses ou d’habilitations à diriger les recherches en TAL préparée par Sylvain Pogodalla. Merci à Denis et Sylvain pour leur travail de veille et de collecte.

Merci aux membres du comité de rédaction de la revue qui ont participé aux différentes étapes d’élaboration de ce numéro, et en particulier à ceux qui ont pris en charge des relectures (voir la composition du comité sur le site de la revue : <https://www.atala.org/content/comité-de-rédaction-0>).

Merci aux relecteurs spécifiques de ce numéro :

- Delphine Battistelli (MoDyCo, Université Paris-Nanterre) ;
- Denis Béchet (LS2N, Université de Nantes) ;
- Anne-Laure Ligozat (LISN, ENSIIE) ;
- Thierry Poibeau (CNRS, ENS/PSL et Université Sorbonne Nouvelle) ;
- Jean-Philippe Prost (LPL, Aix-Marseille Université) ;
- Agata Savary (LIFAT, Université de Tours) ;
- Laure Soulier (LIP6, Sorbonne Université).

La revue TAL reçoit un soutien financier de l'Institut des sciences humaines et sociales (INSHS) du CNRS et de la délégation générale à la langue française et aux langues de France (DGLFLF). Nous adressons nos remerciements à ces organismes.

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

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

Sophie Rosset  
LISN, CNRS  
*sophie.rosset@lisn.fr*

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