
Préface

Comme il est de coutume, cette préface de numéro non thématique donne des nouvelles de notre comité de rédaction, mentionne les progrès dans la gestion de notre revue et fournit des statistiques sur les articles soumis et publiés, avant de présenter brièvement les articles du numéro.

Dans la rubrique vie du comité de rédaction, Éric de la Clergerie nous quitte après plusieurs années de bons et loyaux services en tant que rédacteur en chef. Nous le remercions pour l'équanimité avec laquelle il a toujours su diriger nos débats, pour son souci à mettre en évidence les aspects positifs des articles et pour son acuité dans la synthèse des opinions et l'énoncé des décisions. Le premier auteur de ces lignes quitte aussi ses fonctions de rédacteur en chef, avec essentiellement la gestion des numéros *Varia*, tout en restant au comité. Il le fait avec une certaine nostalgie, mais aussi soulagement, étant donnée la lourdeur de cette charge annuelle récurrente. Isabelle Tellier et Emmanuel Morin ont accepté de prendre la charge de rédacteur en chef en remplacement.

Certains événements récents nous ont poussés à revoir ou préciser certaines de nos procédures. En particulier, une longueur minimale des articles soumis avait été adoptée dès 2012 mais avait été oubliée depuis. Nos conseils aux auteurs mentionnent donc maintenant explicitement une longueur comprise entre 20 et 25 pages.

Un très grand pas en avant a été accompli dans la production automatique des numéros sous forme d'un seul fichier PDF incluant couverture et table des matières. Il faut ici en particulier remercier Jean-Luc Minel pour ses efforts continus, avec l'aide de Thierry Hamon et de Patrick Paroubek. Ainsi, ont été complètement produits sous cette forme tous les numéros depuis le premier numéro après le passage de la revue en version électronique, c'est-à-dire le numéro 47-1 de 2006, jusqu'aux numéros récents. Les prochains numéros seront produits par Sophie Rosset et Maxime Amblard.

Un travail est en cours afin d'augmenter la visibilité de la revue et de permettre le moissonnage automatique de nos métadonnées par les sites de référencement. Il passera sans doute par la mise en place d'un dépôt de nos métadonnées visible par les moteurs d'indexation.

Pour ce qui est de la régularité de publication, nous pouvons là aussi nous montrer satisfaits de l'établissement d'un calendrier prévisionnel qui nous permet de caler les différents numéros d'un même volume et de tenir de façon plus 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 aux cours desquelles nous décidons collégalement, à l'appui des relectures reçues, de l'acceptation ou du rejet des articles soumis.

Passons maintenant à nos statistiques. Elles considèrent toujours les dix derniers numéros sur les trois dernières années, en l'occurrence donc, du début de 2013 jusqu'à ce numéro *Varia* de 2016 inclus. Le tableau 1 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 32,6 % en moyenne, c'est-à-dire que, sur trois articles soumis, un est accepté. Ce taux est stable dans le temps, entre 31 et 35 %, d'après les chiffres donnés depuis le numéro 51-1. Notre comité de rédaction est très attaché à sélectionner les articles selon leur qualité, indépendamment du nombre d'articles soumis. Or l'on peut observer que ce nombre fluctue. L'un de nos soucis actuels est de nous assurer d'un nombre stable de soumissions. Cela semble le cas pour les numéros *Varia*, mais moins pour les numéros thématiques. Peut-être faudrait-il élargir les appels à des articles d'état de l'art ou de prise de position ? Nos plus récents appels thématiques offrent cette possibilité.

Les statistiques que nous donnons sur l'origine des articles considèrent le pays du premier auteur, hors de France ou pas, ainsi que la langue de la soumission, français en principe ou anglais si l'un des co-auteurs n'est pas francophone. Les chiffres sont donnés dans le tableau 2 pour la même période de temps que le tableau 1. En comparant aux chiffres donnés dans les derniers numéros *Varia*, on constatera une diminution récente d'articles acceptés en anglais alors que le nombre de soumissions dans cette langue ne baisse pas vraiment (chiffres non communiqués ici). Il en va de même pour le nombre de premiers auteurs hors de France, qui est tombé à 0 durant tout 2015, pour remonter à un article sur quatre dans le numéro présent.

Le numéro présent contient les articles retenus lors de l'appel non thématique lancé début août et clos à la mi-novembre 2015. Cet appel portait comme d'habitude sur tous les aspects du traitement automatique des langues. Dix-neuf articles ont été soumis dont quatre en anglais, dans la ligne des trois dernières années avec vingt-et-une soumissions dont quatre en anglais en 2013, dix-neuf soumissions dont trois en anglais en 2014, et vingt soumissions dont quatre en anglais en 2015 (le nombre d'articles soumis en anglais n'apparaît pas dans les tableaux). À l'issue du processus de sélection habituel à deux tours, cinq articles, soit un article sur quatre (26,3 %), ont été retenus pour publication, dont un en anglais.

Les tâches abordées par les articles de ce numéro sont extrêmement diverses : reproduction de prononciation, paraphrase, constitution de ressource, analyse en constituants et traduction automatique. Elles couvrent donc tout le spectre de la linguistique de la phonétique (phénomène de liaison) à la sémantique (paraphrase) en passant par la morphologie (pour la paraphrase) et la syntaxe (Verbønet, analyse en constituants). Les techniques utilisées sont aussi très variées : constitution manuelle de ressource (Verbønet), apprentissage automatique par l'exemple (pour la liaison)

| Intitulé | Vol. | N° | Année | Soumis | Acceptés | % acceptés |
|------------------------------|-----------|----|------------------------------------|------------|-----------|---------------|
| Varia | 54 | 1 | 2013 | 21 | 7 | 33,3 % |
| Entités nommées | 54 | 2 | 2013 | 13 | 4 | 30,1 % |
| Réseaux sociaux et TAL | 54 | 3 | 2013 | 8 | 2 | 25,0 % |
| Sous-total | 54 | | 2013 | 42 | 13 | 30,1 % |
| Varia | 55 | 1 | 2014 | 19 | 5 | 26,3 % |
| Trait. auto. langage parlé | 55 | 2 | 2014 | 9 | 6 | 66,6 % |
| TAL et sciences cognitives | 55 | 3 | 2014 | 7 | 4 | 57,1 % |
| Sous-total | 55 | | 2014 | 35 | 15 | 42,9 % |
| Varia | 56 | 1 | 2015 | 20 | 4 | 20,0 % |
| Sémantique distributionnelle | 56 | 2 | 2015 | 7 | 4 | 57,1 % |
| Recherche d'information | 56 | 3 | 2015 | 12 | 3 | 25,0 % |
| Sous-total | 56 | | 2015 | 39 | 11 | 28,9 % |
| Varia | 57 | 1 | 2016 | 19 | 5 | 26,3 % |
| Total | | | Dix derniers n^{os} | 135 | 44 | 32,6 % |

Tableau 1. Taux de sélection aux appels de la revue TAL sur les dix derniers numéros de la période 2013-2016

| Intitulé | Vol. | N° | Année | % 1 ^{er} auteur hors France | % en anglais |
|--------------------------------|-----------|----|------------------------------------|---|---------------|
| Varia | 54 | 1 | 2013 | 28,6 % | 14,3 % |
| Entités nommées | 54 | 2 | 2013 | 0,0 % | 0,0 % |
| Réseaux sociaux et TAL | 54 | 3 | 2013 | 100,0 % | 50,0 % |
| Pourcentages par volume | 54 | | 2013 | 30,1 % | 38,5 % |
| Varia | 55 | 1 | 2014 | 0,0 % | 0,0 % |
| Trait. auto. langage parlé | 55 | 2 | 2014 | 16,6 % | 0,0 % |
| TAL et sciences cognitives | 55 | 3 | 2014 | 25,0 % | 50,0 % |
| Pourcentages par volume | 55 | | 2014 | 13,3 % | 13,3 % |
| Varia | 56 | 1 | 2015 | 0,0 % | 0,0 % |
| Sémantique distributionnelle | 56 | 2 | 2015 | 0,0 % | 0,0 % |
| Recherche d'information | 56 | 3 | 2015 | 0,0 % | 0,0 % |
| Pourcentages par volume | 56 | | 2015 | 0,0 % | 0,0 % |
| Varia | 57 | 1 | 2016 | 20,0 % | 20,0 % |
| Pourcentages totaux | | | Dix derniers n^{os} | 15,9 % | 11,4 % |

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 2013-2016. Attention, les pourcentages totaux ne sont pas de simples moyennes des chiffres donnés plus haut, car les dénominateurs changent.

ou réseaux de neurones (pour l'analyse ou pour la traduction). Mentionnons à ce sujet que ni la conférence TALN ni notre revue n'échappent à la vague des réseaux de neurones puisque les deux articles du présent numéro qui en font usage sont des extensions des deux meilleurs articles de la conférence TALN 2015 (voir : <https://taln2015.greyc.fr/>). L'ordre dans lequel nous avons rangé les articles est un peu acrobatique. Il essaie, sans grand succès, de suivre l'ordre classique de la phonétique à la sémantique.

1) « Predicting Liaison: an Example-Based Approach », Alexander Greefhorst et Antal van den Bosch ;

2) « Un Verbønet du français », Laurence Danlos, Quentin Pradet, Lucie Barque, Takuya Nakamura et Matthieu Constant ;

3) « Prédiction structurée pour l'analyse syntaxique en constituants par transitions : modèles denses et modèles creux », Maximin Coavoux et Benoît Crabbé ;

4) « Exploitation de la morphologie pour l'extraction automatique de paraphrases grand public des termes médicaux », Natalia Grabar et Thierry Hamon ;

5) « Apprentissage discriminant de modèles neuronaux pour la traduction automatique », Quoc-Khanh Do, Alexandre Allauzen et François Yvon.

On trouvera en suite des articles une note de lecture consacrée à un livre sur les mesures de similarité sémantique par approches à base de connaissances ou à base de corpus. Nous encourageons nos lecteurs à se faire mutuellement profiter de leurs lectures et à se mettre en contact avec Denis Maurel (denis.maurel@univ-tours.fr) pour les publier ici. Suit une liste de résumés de thèses ou d'habilitations à diriger les recherches en traitement automatique des langues. Merci à Sylvain Pogodalla pour son travail de veille et de collecte.

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

Yves Lepage
IPS, université Waseda, Japon
Yves.Lepage@waseda.jp

Jean-Luc Minel
MoDyCo, université Paris-Ouest,
Nanterre La Défense
Jean-Luc.Minel@u-paris10.fr

Emmanuel Morin
LINA, université de Nantes
Emmanuel.Morin@univ-nantes.fr

Pascale Sébillot
IRISA, INSA de Rennes
Pascale.Sebillot@irisa.fr

Isabelle Tellier
LaTTiCe, université Paris 3
Isabelle.Tellier@univ-paris3.fr

Merci aux relecteurs spécifiques de ce numéro :

Loïc Barrault, LIUM, université du Mans
Delphine Bernhard, LiLPa, université de Strasbourg
Hervé Blanchon, IMAG, université Grenoble-Alpes
Florian Boudin, LINA, université de Nantes
Philippe Boula de Mareüil, LIMSI
Basilio Calderone, université de Toulouse–Le Mirail
Marco Dinarelli, LaTTICe, ENS
Jean Eychenne, université de Hankuk, Corée
Karën Fort, université Paris-Sorbonne
Eric Gaussier, IMAG, université Grenoble-Alpes
Nicolas Hernandez, LINA, université de Nantes
Marie-Claude L’Homme, université de Montréal, Canada
Guy Lapalme, IRO, université de Montréal, Canada
Pierre-Francois Marteau, université de Bretagne-Sud
Sylvain Meignier, LIUM, université du Mans
Sebastián Peña Saldarriaga, société Dictanova
Fatih Sadat, université du Québec à Montréal, Canada
Leo Wanner, universitat Pompeu Fabra, Espagne

ainsi qu’aux membres du comité de rédaction de la revue (voir sa composition sur notre site).