

DESCRIPTIF

Candidature à l'inscription sur les listes de qualification aux fonctions de
MAÎTRE DE CONFÉRENCES Section 7 (Sciences du Langage)

MARC ALLASSONNIÈRE-TANG¹

RÉSUMÉ DU DOSSIER

- **Soutenance de thèse:** mars 2019
- **Enseignement:** 78h en Sciences du Langage, 35h en Programmation Informatique.
- **Direction:** Soutien en direction de deux doctorants en Sciences du Langage et direction de stage d'un étudiant master en Sciences Cognitives.
- **Recherche:** 12 articles de revue avec comité de lecture, 4 chapitres d'ouvrages, 32 communications dans des colloques internationaux à comité de sélection. 1 projet ANR JCJC (Jeunes Chercheuses et Jeunes Chercheurs) financé.

PRINCIPALES ÉTAPES DE LA CARRIÈRE ANTÉRIEURE

J'ai eu une formation linguistique et j'ai acquis par la suite des compétences en traitement automatique des langues et en phylogénétique.

Après avoir obtenu mes licences en science politique et littérature arabe à l'Université NCCU (Université Nationale de ChengChi) de Taïwan, j'ai travaillé pendant 1 an et demi dans l'entreprise d'ordinateurs ASUS en tant que chef de produit pour le développement des tablettes et ordinateurs portables. Cette expérience m'a permis de réfléchir en connaissance de cause aux différentes options de carrière qui s'offraient à moi et m'a amené à faire le choix de la recherche. Ayant grandi entre la France et Taïwan et ayant les deux langues (français, mandarin) comme langues maternelles, j'ai souvent été confronté à des environnements où le français, le mandarin et l'anglais étaient employés simultanément. Ceci a attisé ma curiosité envers les langues et l'étude du langage.

Concernant mon parcours académique, mes Master 1 et 2 en linguistique ont été réalisés à l'Université Nationale ChengChi (NCCU) de Taïwan. Les thèmes abordés étaient la phonologie et la syntaxe théorique, p.ex., la théorie d'optimalité et la syntaxe générative. Ensuite, j'ai réalisé un doctorat en linguistique à l'Université d'Uppsala en Suède. Mes thèmes de recherche appartenaient majoritairement aux domaines de la morphosyntaxe et de la linguistique de terrain, et utilisaient des méthodes computationnelles et phylogénétiques pour analyser les données recueillies. Ma thèse intitulée *A typology of*

¹ Chercheur postdoctoral au Laboratoire Dynamique du Langage (DDL), CNRS/ Université Lumière Lyon 2, France

classifiers and gender: From description to computation a été dirigée par Michael Dunn (Université d'Uppsala en Suède) et Christine Lamarre (CRLAO, INALCO). J'y analyse les fonctions lexicales et discursives des systèmes de classification nominale (tel le genre en français) et déduis leurs tendances évolutives en utilisant des méthodes de phylogénétique. Après un contrat postdoctoral de six mois à l'Université d'Uppsala durant lequel j'ai continué à développer mon sujet de thèse sur les systèmes de classificateurs (comme en mandarin), je suis actuellement chercheur postdoctoral au laboratoire Dynamique Du Langage (UMR 5596) dans le projet *Variation, change and complexity and linguistic and health-related behaviours* dirigé par Dan Dediu (chercheur IDEX). Ce projet vise à expliquer la variation phonétique dans les langues du monde en employant des synthétiseurs phonétiques et des méthodes de typologie quantitative.

En ce qui concerne l'enseignement, j'ai obtenu la qualification suédoise d'enseignement supérieur (☞) et de direction (☞) en suivant le cursus de l'Unité de Pédagogie et Didactique de l'Université d'Uppsala. En termes d'expérience, j'ai enseigné deux cours de licence en linguistique et un cours de master en programmation. En outre, j'ai été enseignant pour trois modules d'ateliers et formations, durant lesquels j'ai enseigné les bases du travail de terrain et de l'analyse des données. Le contenu de l'un des ces ateliers a également été publié dans une revue pédagogique (Tang, 2020). Ci-après, une présentation succincte des cours et ateliers dispensés:

- **La recherche en linguistique:** Cours de licence (30h). Semestre printemps 2019 (Département de Linguistique et Philologie, Université d'Uppsala). Enseignant responsable du cours, traitant les méthodes qualitatives et quantitatives utilisées en linguistique. ☞
- **Linguistique cognitive:** Cours de licence (30h). Semestre printemps 2019 (Département de Linguistique et Philologie, Université d'Uppsala). Enseignant responsable du cours, traitant les bases théoriques en linguistique cognitive. ☞
- **Visualisation et statistiques:** Cours de master (30h). Semestre automne 2018 (Faculté des langues, Université d'Uppsala). Enseignant pour des sessions de laboratoire hebdomadaires de deux heures. L'enseignement de l'utilisation de la langue de programmation R pour visualiser et analyser des données quantitatives. ☞
- **Sessions de terrain avec locuteurs natifs:** module pour formation intensive (15h). Semestre automne 2019 (International School in Linguistic Fieldwork, Paris). Enseignant pour des sessions journalières de 2 heures pendant une semaine sur la planification et réalisation du terrain. ☞
- **Une introduction à la forêt aléatoire en R:** Module de master et doctorat (5h). Semestre printemps 2018 (Département de linguistique et philologie, Université d'Uppsala). Enseignant pour des sessions de 90 minutes sur l'utilisation de la méthode d'apprentissage automatique de forêt aléatoire. ☞

- **Une analyse statistique de la distinction morphologie-syntaxe:** Module de master et doctorat (3h). Semestre printemps 2020 (Dymanique du Language, UMR 5596). Enseignant pour des sessions de 3 heures sur l'utilisation des classificateurs d'arbres de décision et des modèles linéaires généralisés mixtes ↗

En termes d'encadrement, je suis responsable d'un stage de M1 en Sciences Cognitives à l'Université Lumière Lyon 2, dont le thème porte sur la modélisation bayésienne de l'évolution du lexique. Je suis aussi en soutien de direction pour deux doctorants en Sciences du Langage qui effectuent une étude phonétique des fricatives (↗) et des consonnes rhotiques (↗).

THÈMES DE RECHERCHE

Je suis principalement intéressé par l'évolution des systèmes linguistiques, plus particulièrement par l'évolution des systèmes de classification nominale.

Parmi tous les systèmes linguistiques existants, celui de la classification nominale (c'est-à-dire la classification qu'une langue opère vis à vis de ses référents nominaux) est motivé par le besoin de gérer de manière efficace les données et expériences sauvegardées dans le cerveau humain. Pour faciliter leur accès, ces informations ont besoin d'être catégorisées (Lakoff and Johnson, 2003; Kemmerer, 2014, 2017a,b). Cette nécessité se reflète fréquemment (mais pas obligatoirement) dans les langues à travers deux types de systèmes de classification nominale (Figure 1). Le premier système est le genre grammatical (Corbett, 1991), tel que la distinction entre le masculin et le féminin en français¹. Le deuxième système correspond aux classificateurs (Aikhenvald, 2000; Grinevald, 2015), qui, comme c'est le cas en mandarin, catégorisent les entités en distinguant la forme des objets, p.ex., *yi4 ke1 ping2guo3* (un CLF.ROUND pomme) 'une pomme'.

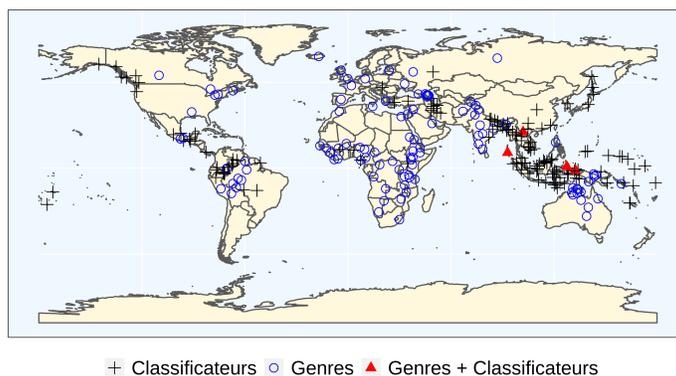


Figure 1: Une vue d'ensemble de la distribution des genres grammaticaux et classificateurs dans les langues du monde. (Corbett, 2013; Gil, 2013)

¹ Les travaux dans le domaine emploient les termes 'genres' et 'classes nominales' selon des traditions différentes. Dans le présent document, j'emploie la terminologie de Corbett (2013) et utilise le terme 'genre grammatical' pour couvrir les deux types de systèmes.

Les travaux scientifiques disponibles indiquant que les classificateurs ont tendance à se diffuser prioritairement de manière horizontale (géographique) plutôt que verticale (généalogique), j'ai tenté d'identifier les facteurs linguistiques qui pourraient faciliter cette diffusion (Tang, 2017; Her et al., 2019; Tang and Her, 2019; Her and Tang, 2020). J'ai par ailleurs effectué des recherches sur l'affectation des genres grammaticaux (Basirat and Tang, 2017, 2018, 2019). En parallèle de ces analyses typologiques quantitatives, je me suis intéressé aux langues indo-aryennes avec une perspective descriptive; ces langues représentent en effet le cas rarissime d'un groupe de langues dont certains membres ont développé des classificateurs, tandis que d'autres possèdent un système de genre grammatical et d'autres encore cumulent les deux types de systèmes (Eliasson and Tang, 2018; Saikia and Allasonnière-Tang, 2020; Kilarski and Tang, 2018; Allasonnière-Tang and Kilarski, 2020). J'ai aussi conduit des analyses phylogénétiques sur l'évolution du genre dans les langues indo-aryennes (Allasonnière-Tang and Dunn, 2020).

AUTRES RESPONSABILITÉS ASSUMÉES

En termes de projets, j'ai été co-responsable du projet *GIS for language study* financé par l'Université d'Uppsala entre 2017 et 2019. Ce projet avait pour but de promouvoir l'utilisation de systèmes d'information géographique (p.ex., QGIS) dans la faculté des langues de l'université en organisant plusieurs ateliers et groupes de travail [↗](#). Depuis 2019, je suis aussi co-responsable de la formation annuelle *FieldLing* (Funded international school in linguistic fieldwork), qui est organisée par trois laboratoires du Labex EFL: le LACITO, le LLACAN et le SeDyL [↗](#). En 2020, j'ai reçu un financement ANR Jeunes Chercheurs pour le projet EVOGRAM: L'effet des facteurs linguistiques et non-linguistiques sur l'évolution des systèmes de classification nominale [↗](#).

J'ai été co-organisateur de quatre événements internationaux avec en moyenne 50 participants: En 2018, l'*Atelier de pratique pour les systèmes d'information géographique* à l'Université d'Uppsala en Suède. [↗](#). En 2017, le *Séminaire de lancement du réseau de recherche 'GIS for Language Study'* à l'Université d'Uppsala en Suède [↗](#). En 2016, l'*Atelier sur la classification nominale en Asie du Sud-Est* à l'Université Adam Mickiewicz en Pologne [↗](#). En 2016 les *29èmes journées de linguistique Asie orientale* à l'INALCO en France [↗](#).

Au cours de ces dernières années, j'ai été également sollicité pour participer à des comités de lecture. J'ai ainsi été évaluateur pour les revues et maisons d'édition suivantes (liste non-exhaustive): *Linguistic Typology*, *Studies in Language*, *Corpus Linguistics and Linguistic Theory*, *Oceanic Linguistics*, *Linguistic Variation*, *International Journal of American Linguistics*, De Gruyter Mouton, John Benjamins, Language Science Press. Par ailleurs, j'ai été évaluateur pour un certain nombre de conférences internationales (liste non-exhaustive): *Annual Meeting of the Societas Linguistica Europaea*, *Annual Paris Meeting on East Asian Linguistics*, *International Symposium on Chinese Theoretical and Applied Linguistics*, *Empirical Methods in Natural Language Processing*.

BIBLIOGRAPHIE

- Aikhenvald, A. Y. (2000). *Classifiers: A typology of noun categorization devices*. Oxford University Press, Oxford.
- Allasonnière-Tang, M. and Dunn, M. (2020). The evolutionary trends of grammatical gender in Indo-Aryan languages. *Language Dynamics and Change*, (in press).
- Allasonnière-Tang, M. and Kilarski, M. (2020). Functions of gender and numeral classifiers in Nepali. *Poznan Studies in Contemporary Linguistics*, 56(1):113–168.
- Basirat, A. and Tang, M. (2017). Neural network and human cognition: A case study of grammatical gender in Swedish. *Proceedings of the 13th Swedish Cognitive Science Society national conference*, pages 28–30.
- Basirat, A. and Tang, M. (2018). Lexical and morpho-syntactic features in word embeddings: A case study of nouns in Swedish. *Proceedings of the 10th International Conference on Agents and Artificial Intelligence*, 2:663–674.
- Basirat, A. and Tang, M. (2019). Linguistic Information in Word Embeddings. In van den Herik, J. and Rocha, A. P., editors, *Agents and Artificial Intelligence*, volume 11352, pages 492–513. Springer International Publishing, Cham.
- Corbett, G. G. (1991). *Gender*. Cambridge University Press, Cambridge.
- Corbett, G. G. (2013). Number of Genders. In Dryer, M. S. and Haspelmath, M., editors, *The World Atlas of Language Structures Online*. Max Planck Institute for Evolutionary Anthropology, Leipzig.
- Eliasson, P. and Tang, M. (2018). The lexical and discourse functions of grammatical gender in Marathi. *Journal of South Asian Languages and Linguistics*, 5(2):131–157.
- Gil, D. (2013). Numeral classifiers. In Dryer, M. S. and Haspelmath, M., editors, *The world atlas of language structures online*. Max Planck Institute for Evolutionary Anthropology, Leipzig.
- Grinevald, C. (2015). Linguistics of classifiers. In Wright, J. D., editor, *International Encyclopedia of the Social & Behavioral Sciences*, pages 811–818. Elsevier, Oxford.
- Her, O.-S. and Tang, M. (2020). A statistical explanation of the distribution of sortal classifiers in languages of the world via computational classifiers. *Journal of Quantitative Linguistics*, 27(2):93–113.
- Her, O.-S., Tang, M., and Li, B.-T. (2019). Word order of numeral classifiers and numeral bases. *STUF - Language Typology and Universals*, 72(3):421–452.
- Kemmerer, D. (2014). Word classes in the brain: Implications of linguistic typology for cognitive neuroscience. *Cortex*, 58:27–51.
- Kemmerer, D. (2017a). Categories of object concepts across languages and brains: the relevance of nominal classification systems to cognitive neuroscience. *Language, Cognition and Neuroscience*, 32(4):401–424.
- Kemmerer, D. (2017b). Some issues involving the relevance of nominal classification systems to cognitive neuroscience: response to commentators. *Language, Cognition and Neuroscience*, 32(4):447–456.
- Kilarski, M. and Tang, M. (2018). The coalescence of grammatical gender and numeral classifiers in the general classifier wota in Nepali. *Proceedings of the Linguistic Society of America*, 3(1):56.
- Lakoff, G. and Johnson, M. (2003). *Metaphors we live by*. The University of Chicago Press, London.
- Saikia, P. and Allasonnière-Tang, M. (2020). Nominal classification in Assamese: An analysis of function. In Kilarski, M. and Tang, M., editors, *Nominal classification in Asia: Functional and diachronic perspectives*. John Benjamins. (in revision), Amsterdam.
- Tang, M. (2017). Explaining the acquisition order of classifiers and measure words via their mathematical complexity. *Journal of Child Language Acquisition and Development*, 5(1):31–52.
- Tang, M. (2020). A simple introduction to programming and statistics with decision trees in R. *Teaching Statistics*, 42(2):36–40.
- Tang, M. and Her, O.-S. (2019). Insights on the Greenberg-Sanches-Slobin generalization: Quantitative typological data on classifiers and plural markers. *Folia Linguistica*, 53(2):297–331.