

## Présentation des communications (par ordre alphabétique)

### Myriam Bouveret (univ. de Rouen, Lattice)

« La quasi-synonymie des verbes de séparation : étude contrastive français-anglais de *casser*, *briser*, *rompre* vs *break* »

En anglais les verbes *casser*, *briser* et *rompre* français se traduisent tous trois par le verbe *break*. Ce ne sont pourtant pas de simples équivalents du verbe *break*. Quelles sont leurs particularités de ces quasi-synonymes en français ? Nous étudierons les emplois relevés sur corpus en nous appuyant sur les ressources du CNRTL : Dicosyn, le TLF et Frantext (<https://www.cnrtl.fr>) et montrerons notamment que les trois verbes sont en relation de synonymie pour le sens I.1 de *casser* ‘mettre en morceaux’ ‘par pression ou ‘par choc’ en emploi transitif ou sous forme pronominale réfléchi (ex *La branche s’est cassée/s’est brisée/s’est rompue* ‘*the branch broke*’). Cependant, la spécificité qui apparaît le plus souvent dans les énoncés en corpus selon ce sens I.1. pour *rompre* est le sème ‘en deux’, tandis que pour *briser* c’est le sème ‘en multiples morceaux’, *casser* est plus générique dans sa décomposition sémantique. Les emplois concrets montrent par ailleurs qu’une synonymie peut s’établir notamment entre *rompre* et *briser* lorsqu’un lien est signifié (*rompre/briser des chaînes*) et par extension métaphorique un lien affectif : les relations affectives (domaine source) sont un lien (domaine cible) qui peut être brisé ou rompu, de même pour les relations amicales ou diplomatiques (*I. Ils n’osaient se parler, de peur de briser le charme* (Léon Gozlan 1836)/ *Le mouvement qu’ils avaient fait venait de rompre le charme de l’hésitation* (Pierre Loti 1896)). L’analyse syntaxique et sémantique mise en œuvre s’intéresse en particulier à la syntaxe des modificateurs, compléments prépositionnels ou adverbes (concordancier Word Sketch Engine, corpus web 126 850 281 mots) mais nous élargissons également aux emplois métaphoriques des trois verbes et constructions verbales selon le modèle Frame Semantics et la grammaire cognitive de construction (CCxG).

### Gilles Corminboeuf & Richard Huyghe (univ. de Fribourg, Suisse)

« Les alternances objet / oblique en français »

Cette présentation porte sur le cas où deux constructions, l’une à complément d’objet, l’autre à complément oblique, sont en concurrence pour un même verbe (ex. goûter X / goûter à X, chercher X / chercher après X, pousser X / pousser sur X). Nous définissons dans un premier temps les conditions exactes des alternances étudiées, puis nous examinons la contrepartie sémantique éventuelle de la double construction. Nous établissons une tripartition des alternances, fondée sur la littérature scientifique à disposition, et nous montrons que la classification proposée ne couvre pas l’ensemble des cas répertoriés en français. Nous proposons une approche multifactorielle du phénomène, qui ménage une place aux faits de variation inter-locuteurs.

**Catherine FUCHS (CNRS, ENS)**

**et Sylvie GARNIER (Université de Chicago à Paris)**

« Enseignement du lexique et synonymie verbale : étude contrastive de deux couples de verbes »

La présentation portera sur le traitement de la synonymie verbale dans un manuel de Français sur Objectifs Universitaires destiné aux étudiants étrangers avancés (Fuchs & Garnier, *Lexique raisonné du français académique*, Ophrys, 2020).

Après avoir présenté les choix théoriques et méthodologiques adoptés dans ce manuel, on montrera comment la comparaison systématique entre les verbes pris deux à deux permet de caractériser la distance sémantique plus ou moins grande entre deux verbes, en fonction des catégories et des types de noms qu'ils appellent de façon prototypique.

Cette approche sera illustrée sur les deux couples de verbes *éliminer / supprimer* et *mettre fin à / mettre un terme à*.

**Emmanuelle Guérin (univ. Sorbonne Nouvelle, DILTEC)**

« *Femme* » vs « *meuf* » : le prétexte de la synonymie pour ne pas voir la variation

Traditionnellement, la variation en français est pensée selon une représentation hiérarchisée de « variétés » cloisonnées qui tient, notamment, à des jeux d'équivalences instituant la binarité bon/mauvais usages. La considération de ce qu'il est commun d'appeler les « parlars jeunes » n'y échappe pas. Ainsi, les formes non-standard, qu'il s'agisse de mots construits par manipulation de la forme initiale (par exemple, le verlan), d'emprunts à une langue étrangère ou de métaphores récurrentes, sont communément perçues comme des concurrents (synonymes ?) illégitimés de formes standard, relativement à un procédé crypto-ludique. Ces supposées équivalences concourent à discréditer socialement les locuteurs, alors soupçonnés de ne pas adhérer aux référents communs. En m'appuyant sur un corpus constitué de données illustrant lesdits parlars en Île de France (*Multicultural Paris French*), je tenterai de montrer qu'il est davantage question de néologismes, enrichissant le lexique français d'items répondant à des besoins langagiers qui ne sont pas prévus par le lexique standard. Ainsi, « meuf » et « femme » ne sont pas concurrents (synonymes ?), le premier prenant un sens générique et le second un sens spécifique.

### **Pierre Jalenques (univ. de Rouen, Dylis)**

« La synonymie syntagmatique entre unités de catégories différentes : quelques exemples à partir des verbes préfixés en RE- et DÉ- »

La relation de synonymie (au sens large, incluant la quasi-synonymie) est habituellement définie à partir du critère de substitution. Cela amène à ne prendre en compte que les cas de synonymie entre des unités de même catégorie (puisqu'elles doivent pouvoir commuter dans un même co-texte). Cela met à l'écart tout un ensemble de phénomènes, à notre connaissance très peu étudiés en tant que tels, correspondant à ce que l'on peut appeler la *synonymie syntagmatique*. Il s'agit de relations de synonymie entre des éléments de catégories différentes se combinant sur l'axe syntagmatique et ne commutant jamais sur l'axe paradigmatique.

Nous nous intéresserons en particulier aux relations sémantiques syntagmatiques entre préfixes verbaux et bases verbales. Par exemple, en français, dans certains cas, l'ajout du préfixe RE- ou du préfixe DÉ- à une base verbale ne change pas notablement le sens global de l'énoncé, comme dans *je ferme / referme la parenthèse et je passe au point deux ; veuillez poser / déposer votre arme sur la table*. Dans ce type d'emploi, on considère généralement que le préfixe RE-/DÉ- est (quasi)-vide de sens, puisque son ajout à la base verbale ne modifie pas notablement le sens de l'énoncé. Pourtant, dans une approche holiste du sens des énoncés et de la synonymie, on peut montrer que ces préfixes ne sont pas vides de sens mais que la synonymie locale entre par exemple *poser* et *déposer* résulte elle-même d'une relation de synonymie syntagmatique entre le préfixe lui-même et la base verbale.

### **Dominique Legallois (univ. Sorbonne Nouvelle, Lattice)**

« L'identification automatique de la relation de synonymie: quelques apports des méthodes dites de plongement lexical (word embedding) ? Une observation à partir de trois corpus »

Les méthodes de plongement lexical ou le « word embedding » constituent une nouvelle technique de représentation des mots comme vecteurs dans un espace. Il s'agit-là, en quelque sorte, d'une métaphore permettant de mathématiser une conception linguistique développée indépendamment par Firth (*Papers in Linguistics 1934–1951* (1957)) et Harris (« Distributional structure », *Word*, vol. 10, no2-3, p. 146-162, 1954) dans les années 40-50 : la sémantique d'un mot  $W_0$  est nécessairement informée par les mots  $W_1, W_2, \dots, W_n$  partageant les contextes d'emplois de  $W_0$  (ses co-occurents). Dans le cadre de l'apprentissage profond (deep learning), un algorithme « apprend » les vecteurs à partir d'un corpus et les positionne dans un espace de façon à ce que les mots partageant des contextes semblables soient proches les uns des autres. Les vecteurs de mots les plus similaires sont identifiés à l'aide du calcul des cosinus.

Les études consacrées à ce type d'analyses qui sont assez peu souvent signées par les linguistes, font ressortir des mots « similaires », désignés parfois par l'étiquette « synonymes ».

Dans cette communication, nous proposons une présentation et une évaluation des méthodes informatiques de word embedding : à partir de trois corpus constitués par nos soins (dont on ne considère que les lemmes – les mots grammaticaux ayant été écartés) :

- un corpus de 79 romans Harlequin, de +1 800 000 lexèmes (ou lemmes)
- un corpus de débats à l'Assemblée nationale française, de +1 500 000 lexèmes
- un corpus de presse régionale (années 2000), de +3 600 000 lexèmes

nous comparons les résultats d'une recherche de synonymes verbaux (*saisir, s'emparer, prendre, tenir*, etc.). Les corpus utilisés sont de taille raisonnable ; ce sont même, d'une certaine manière, de petits corpus au regard des bases textuelles « tout venant » utilisées habituellement en TAL, qui ne respectent pas nécessairement le critère de l'unité générique. Mais nos corpus répondent aux pratiques et aux normes en vigueur en linguistique.

Cette recherche se fonde en fait sur deux méthodes :

- METH1 : une approche « word embedding » utilisant le programme Text2Vec une technique d'apprentissage par réseaux de neurones.
- METH2 : une approche plus « traditionnelle » décrite par J. Silge : <https://juliasilge.com/blog/tidy-word-vectors/> par construction d'une matrice de vecteurs de mots.

La synonymie est foncièrement une relation paradigmatique ; pourtant, force est de constater que les techniques employées pour identifier des relations en discours réussissent (dans une certaine limite qu'il convient de préciser) à circonscrire automatiquement des classes paradigmatiques de synonymes. Par exemple, l'application de METH2 sur le corpus Harlequin réussit à identifier, comme synonyme de *saisir* : *tendre, tirer, s'emparer, prendre, attraper*.

Nous évaluerons donc les résultats de cette première expérience en comparant les deux méthodes appliquées aux trois corpus. A terme, on peut penser que l'identification de synonymes dans des textes « contrôlés » par le linguiste, peut permettre aussi bien de constituer des ressources lexicales, que d'observer la sensibilité de la synonymie au genre textuel.

PS : les scripts R utilisés peuvent être diffusés aux collègues intéressés.

### **Alain Polguère (univ. de Lorraine, ATILF)**

« Paraphrases linguistique et conceptuelle : le cas de la (quasi-) synonymie des verbes collocationnels »

La distinction entre paraphrases linguistique et conceptuelle est bien connue : équivalence communicationnelle entre énoncés fondée (i) sur une (quasi-)identité de leurs significés linguistiques vs (ii) sur une même inférence pouvant être effectuée à partir des situations dénotées par chacun des énoncés. Dans notre intervention, nous nous intéressons à cette distinction sous l'angle de la (quasi-)synonymie des verbes collocationnels, c'est-à-dire des verbes phraséologiquement sélectionnés par un élément lexical avec lequel ils cooccurrent dans la phrase. Nous utilisons pour ce faire les outils théoriques et descriptifs de la Lexicologie Explicative et Combinatoire et, tout particulièrement, le système des fonctions lexicales.