

GRAMM - R

ÉTUDES DE LINGUISTIQUE FRANÇAISE

Catherine Bolly

# Phraséologie et collocations

Approche sur corpus  
en français L1 et L2

P. I. E.  
PETER LANG



GRAMM - R

ÉTUDES DE LINGUISTIQUE FRANÇAISE

Catherine Bolly

# Phraséologie et collocations

Approche sur corpus  
en français L1 et L2

P. I. E.  
PETER LANG



## Complexité des unités phraséologiques

Bien qu'omniprésentes dans l'usage courant (par ex. : *avoir lieu, tirer la langue, fou rire, t'sais, à moins que*) et dans les langues de spécialité (par ex. : *rendre un arrêt* (juridique), *chambre stérile* (médical)), les unités phraséologiques ont longtemps été considérées comme des phénomènes linguistiques atypiques gravitant à la périphérie de la syntaxe. Les linguistes eux-mêmes n'ont pas encore fini de chercher à expliquer le fonctionnement de ces unités situées à l'interface du lexique et de la grammaire, qu'il s'agisse de les décrire à partir de critères internes à la langue (sémantiques, syntaxiques, lexicaux, etc.) ou de critères externes à celle-ci (cognitifs, sociaux, géographiques, diachroniques, stylistiques, etc.). Comment en effet expliquer que le locuteur natif préfère utiliser le verbe *tirer* – plutôt que le verbe *sortir, extraire* ou *tendre* –, pour exprimer verbalement l'action de « faire sortir » *la langue* de sa bouche ? Et comment expliquer cette différence de sens, que le locuteur natif perçoit de manière intuitive, entre l'enfant qui *tire la langue* pour faire une grimace à son voisin et l'enfant qui *tire la langue* pour que le médecin puisse l'ausculter ? Si aucune réponse simple ne semble à ce jour pouvoir être apportée à ces questions, on s'accorde néanmoins actuellement sur le fait que ces unités pour le moins complexes se laissent difficilement appréhender tant par le locuteur natif que par le locuteur non natif.

Partant de ce constat, nous tenterons de déterminer, dans cet ouvrage, en quoi les unités phraséologiques construites avec des verbes fréquents peuvent être comprises comme des phénomènes linguistiques « complexes », en français écrit langue première (L1) et langue étrangère ou seconde (L2). L'ouvrage est ainsi structuré en quatre parties qui visent chacune à répondre à l'une des questions suivantes : Qu'est-ce qu'une unité phraséologique (chapitre 1) ? Comment mesurer le degré de figement des unités phraséologiques (chapitre 2) ? Quelle est la place occupée par ces unités en acquisition d'une L2 (chapitre 3) ? Comment identifier et décrire les « erreurs » phraséologiques produites par des apprenants avancés du français L2 (chapitre 4) ?

Dans le premier chapitre, nous abordons le phénomène phraséologique sous l'angle de la description en français L1, afin d'en dégager les caractéristiques définitoires et de fournir à l'analyse des données de

corpus un ancrage théorique solide, héritier de la littérature tant franco-phonique qu'anglo-saxonne. Si nous devions les définir ici en quelques mots, nous dirions que les *unités phraséologiques*, appelées aussi *séquences (semi-)figées*, sont des unités linguistiques polylexicales (i.e. constituées de plusieurs unités lexicales) qui se sont conventionnalisées au fil du temps et dont les constituants co-apparaissent de manière contrainte dans l'usage.

Dans le second chapitre, nous proposons une méthode d'analyse permettant de mesurer le degré de figement des unités phraséologiques, à travers une étude de cas menée sur 200 séquences verbo-nominales construites avec le verbe *prendre* dans des corpus journalistiques et littéraires en français contemporain L1 (400 000 mots) (par ex., par ordre croissant de degré de figement : *prendre un fruit* > *prendre une décision* > *prendre contact* > *prendre le taureau par les cornes*).

Quelques notions issues du domaine des recherches en acquisition d'une L2 seront abordées dans le troisième chapitre. Une attention particulière y sera accordée au rôle que les unités phraséologiques jouent dans le processus acquisitionnel, et en particulier au stade avancé d'apprentissage d'une L2. Ces notions constituent un préalable théorique aux analyses sur corpus d'apprenants du français L2 du chapitre suivant.

L'objectif du quatrième chapitre est à la fois descriptif et méthodologique : il s'agit (i) de cerner les traits phraséologiques caractéristiques de l'interlangue avancée du français académique L2 en identifiant les « erreurs » spécifiques aux apprenants dans leur production d'unités phraséologiques (par ex. : *\*donner un coup d'œil* pour *jeter un coup d'œil*, *prendre un \*moment de réflexion* pour *prendre un temps de réflexion*) et (ii) de décrire étape par étape l'élaboration de l'outil d'analyse qui nous a permis d'identifier et de décrire ces « erreurs ». Cette analyse interlangagière consiste en l'analyse collocationnelle de 906 séquences verbo-nominales (dont 703 unités phraséologiques) construites avec les verbes à haute fréquence *prendre* et *donner*. Le corpus se compose de textes en français académique produits par des apprenants avancés anglophones (185 000 mots), néerlandophones (90 000 mots) et par des étudiants francophones natifs (185 000 mots). Nous évoquons en outre dans le dernier point de ce chapitre quelques pistes didactiques et pédagogiques.

Si nous avons restreint notre objet d'étude aux unités phraséologiques verbo-nominales construites avec les deux verbes *prendre* et *donner*, c'est que ces verbes de base se caractérisent par : (i) leur potentiel phraséologique important et leur haute fréquence dans l'usage qui, en compensant la rareté du phénomène phraséologique sous sa forme très figée (par ex. : *prendre des gants*) (par opposition au semi-figement

très répandu dans l'usage, par ex. *prendre la parole*), nous ont permis d'effectuer des analyses sur corpus intégraux de données phraséologiques authentiques en contexte ; (ii) l'existence de formes équivalentes pour ces verbes dits *nucléaires* dans de nombreuses langues (Viberg, 2002b), qui permet d'étendre l'étude de ces verbes à d'autres langues et/ou à d'autres verbes de ce type. D'un point de vue méthodologique, cette limitation de l'objet d'étude nous a par ailleurs permis : (i) de procéder à une analyse fine et exhaustive de données de corpus authentiques en L2 et en L1, en nous appuyant sur des bases de données lexicales et phraséologiques préexistantes disponibles pour le français L1 ; (ii) de développer des outils de travail reposant sur une approche multi-dimensionnelle du langage (syntaxique, sémantique et lexicale), incluant des critères phraséologiques qualitatifs (approche fonctionnelle de la phraséologie), mais aussi quantitatifs (approche statistique de la phraséologie) ; (iii) d'adopter une méthodologie s'inspirant, entre autres, de la méthode d'analyse paramétrique et statistique de données de corpus (décrite dans Degand et Bestgen, 2004), de la sémantique contextuelle (Rastier *et al.*, 1994), de la linguistique de corpus (McEnery et Wilson, 1996 ; Habert *et al.*, 1997 ; Kennedy, 1998), de l'analyse contrastive sur corpus informatisés d'apprenants (Granger, 2004) et de l'analyse d'erreurs (James, 1998).

Cette étude des unités phraséologiques s'inscrit dans une conception large du figement et porte une attention toute particulière aux séquences *semi-figées* (par ex. : *donner envie, prendre connaissance de qqch.*), puisque celles-ci représentent la quasi-totalité des données que nous avons extraites de nos corpus (en L2 et L1). L'étude de ce type d'unités phraséologiques met en avant un phénomène crucial en acquisition d'une L2 : les séquences semi-figées (*vs* figées), qui sont les plus fréquentes dans l'usage, sont aussi les moins évidentes à appréhender, tant du point de vue de leur description linguistique que des phénomènes acquisitionnels qui les sous-tendent.

Il est important de souligner que si la méthode adoptée tend bien à restreindre les biais méthodologiques liés aux analyses strictement quantitatives ou qualitatives, elle ne permet cependant pas d'éviter certains inconvénients liés à l'une ou l'autre approche. Outre le fait que nous avons constamment essayé de mettre le doigt, au fil des pages, sur ces inconvénients méthodologiques, nous mettrons en avant dans la conclusion de l'ouvrage les limites majeures de notre recherche, en même temps que nous y mettrons en lumière quelques perspectives susceptibles de prolonger ce travail.