

## ***L'Encyclopédie Grammaticale du Français (EGF): présentation et illustration***

Le texte sur l'interrogation directe présenté par Aidan Coveney s'inscrit dans le cadre d'une vaste entreprise, *L'Encyclopédie Grammaticale du Français* (EGF). Ce projet, initié par Claire Blanche-Benveniste, ensemble avec quatre de ses collègues, tous spécialistes en linguistique française<sup>1</sup>, vise à rassembler, de manière commode et pédagogique et sous forme électronique, le patrimoine des acquis de la recherche en linguistique française, à partir de synthèses régulièrement mises à jour et de documents de première main.

Le projet résulte d'une série de réflexions en commun sur la linguistique française, au moment où cette discipline s'ouvre largement à la recherche interdisciplinaire (informatique, neurolinguistique, sciences cognitives). Il cherche à remédier à certaines faiblesses et lacunes:

- (a) Une quantité énorme de travaux descriptifs ont paru sur le français au cours du XXe siècle (thèses, monographies, articles de revues). Mais, d'une part, en raison de leur nombre même, et de leurs degrés de visibilité très inégaux, beaucoup d'informations pertinentes restent noyées dans la masse. D'autre part, les recherches n'ont pas eu suffisamment d'effet cumulatif : on n'a pas systématiquement relevé les analyses sur lesquelles tout le monde semble être d'accord (même si les terminologies différentes masquent parfois ces consensus). Une des conséquences désastreuses en est que chaque jeune chercheur se croit obligé de recommencer à zéro les mêmes démonstrations.
- (b) On déplore toujours l'absence de données suffisantes en quantité et en qualité dans certains domaines de recherche (la langue parlée p.ex.). Il n'y a pas eu non plus de réflexions systématiques sur le tri des données linguistiques qui servent à l'établissement des descriptions grammaticales : phénomènes attestés ou non attestés, liés à un « genre » ou non, fréquents ou non, etc. Une des conséquences regrettables est que l'on voit mal la stratification des phénomènes, souvent présentés sur le même niveau.
- (c) Nous ne disposons pas de bibliographies commodes et minimalement exhaustives du domaine, qui tiendraient compte des recherches mal diffusées (thèses, articles de revues peu accessibles) et qui permettraient d'en avoir une vue critique, bien hiérarchisée. Un bilan permettrait du même coup de découvrir les zones d'ombre dans la recherche et des sujets de thèses possibles.

En réponse à cet état de choses, l'objectif principal de l'EGF est de constituer une synthèse des savoirs grammaticaux sur le français actuel. C'est-à-dire de faire la **somme des acquis descriptifs** qui, au-delà des effets de mode, peuvent être tenus pour des connaissances solides, sinon incontestables, et candidates au statut de « savoir commun », tant à l'intérieur qu'à l'extérieur de la

---

<sup>1</sup> Marie-José Béguelin, Alain Berrendonner, José Deulofeu, et Dominique Willems

science linguistique. Cette somme, au demeurant, est vouée à évoluer au fil du temps et recevra des mises à jour périodiques, que l'outil informatique rend aisées et dont il permet de conserver la trace.

Présenter une synthèse des acquis suppose évidemment qu'une **sélection critique** soit opérée parmi les travaux existants. C'est pourquoi une large place est réservée dans l'EGF à la réflexion sur les pratiques d'analyse en vigueur, au débat sur les présupposés théoriques et les notions primitives généralement assumées, ainsi que sur les usages terminologiques et les divers métalangages de description. Il ne s'agit évidemment pas de faire table rase de toutes les descriptions établies, mais de les ramener aux thèses et opérations dont elles procèdent, afin que chacun puisse en évaluer la pertinence.

L'effort **d'explicitation méthodologique** entrepris ici devrait avoir pour effet d'intensifier les échanges entre écoles ou pratiques de recherche, et donc de favoriser leur décroisement. Par là, l'EGF vise en quelque sorte une unification non seulement des savoirs, mais encore des dynamiques productrices de savoirs. Là où existent des controverses (désaccords profonds entre modèles, divergences théoriques plus ou moins radicales), les rédacteurs de l'EGF se donnent pour premier objectif de dresser un tableau des positions en présence, qui rende compte de l'hétérogénéité des connaissances accumulées.

Une option fondamentale de l'EGF est **la place réservée aux données**. Toute description linguistique doit être évaluée d'abord sur sa façon de sélectionner *a priori* ses observables : les données dont elle fait état sont-elles attestées ou controuvées, triviales ou inédites, prototypiques ou marginales, fréquentes ou non ? etc. Le seul moyen de s'en éclaircir est de confronter les descriptions à des corpus, oraux et écrits. À l'heure actuelle, on assiste d'ailleurs sur ce point à un renversement significatif des mentalités grammaticales : les investigations sur corpus tendent à se développer même dans le cadre de théories qui les réputaient jusqu'alors inutiles, et s'en remettaient entièrement à l'intuition du grammairien.

L'observation précise des données de corpus permet non seulement d'attester l'existence de structures syntaxiques, mais aussi de caractériser la façon dont s'y distribue le lexique. Elle permet aussi de mesurer la fréquence ou le rendement de chaque construction en fonction de divers paramètres environnementaux (types d'interaction, genres discursifs, buts pragmatiques...) ou intrinsèques (complexité opérative...) Ces mesures sont précieuses pour le grammairien, qui peut en tirer une image hiérarchisée des phénomènes. Elles sont aussi utiles pour tous ceux qui ont à développer des applications pratiques, que ce soit en traitement automatique du langage ou dans l'apprentissage du français langue étrangère. Enfin, elles permettent d'identifier les domaines de recherche dans lesquels est à déplorer un manque de données suffisantes, en quantité et en qualité.

Les sources bibliographiques actuellement disponibles ne rendent pas des services optimaux, pour des raisons classiques : excès de silence (p. ex. thèses non publiées) et excès de « bruit » (articles redondants). Il a donc paru utile de proposer une **sélection de références bibliographiques** ciblées, qui ne vise pas à l'exhaustivité documentaire, mais soit plutôt fondée sur des critères qualitatifs de

pertinence, de fécondité conceptuelle et de solidité empirique (travaux fondés sur des dépouillements). Il s'agit entre autres de mettre en valeur des apports méconnus, et de contribuer au décloisonnement territorial de la recherche, en faisant place aux études publiées hors des pays francophones, et en intégrant une perspective contrastive.

En conclusion, L'EGF n'est pas une grammaire, et n'a pas été conçue sur le plan habituel d'une grammaire française. C'est un instrument qui répond principalement à quatre objectifs : (1) présenter une synthèse du savoir grammatical sur le français; (2) dresser un bilan des méthodes appliquées en grammaire française; (3) mettre à disposition un corpus de données écrites et orales; (4) aider à la recherche bibliographique.

Les cinq promoteurs du projet ont élaborés les documents préparatoires essentiels à son démarrage : une table des matières, une introduction méthodologique et surtout un modèle type de notice pour la présentation de l'information. C'est ce format qui a été rigoureusement adopté par Aidan Coveney dans le texte qui suit. Il constitue la première notice établie selon les principes de l'EGF<sup>2</sup>.

---

<sup>2</sup> D'autres suivront, qui feront, dans un premier mouvement, également l'objet d'une publication papier. Dès qu'un nombre suffisant de notices aura été élaboré, le site électronique d'EGF sera ouvert, offrant au public des chercheurs les commodités de consultation, de recherche et de navigation liées à l'informatisation et à l'hypertexte.