Présentation

José DEULOFEU
Université de Provence

Le présent numéro de *Recherches sur le français parlé* s'inscrit dans les grandes lignes de recherche du groupe DELIC qui prolongent celles du G.A.R.S.

Un premier groupe d'articles prolonge l'entreprise de description du français à partir de corpus intégrant des données orales. Les deux grands domaines de la description (microsyntaxe et macrosyntaxe) sont abordés ainsi que leurs interrelations :

Dans le domaine de la microsyntaxe, Paul Cappeau propose, à partir de données orales, un classement des séquences qui peuvent figurer entre l'auxiliaire et le participe passé. La prise en compte de ce type de données lui permet à la fois de révéler de nouveaux faits (syntaxe de la négation) et de dégager les tendances structurales les plus remarquables qui se manifestent sur cette position.

André Valli à propos des constructions dites « datif partitif » souligne que le fait de retenir ou de ne pas retenir telles ou telles données a des conséquences dans la description syntaxique de la question : l'usage tel qu'il ressort de l'examen de corpus de langue parlée et de français écrit littéraire est en écart par rapport à l'usage puriste des grammairiens, et en revanche, conforme à l'usage qui était régulier dans les textes d'ancien et de moyen français.

Deux autres articles montrent que la description de certaines structures demandent que l'on articule finement les niveaux macro et micro syntaxiques.

Mylène Blasco-Dulbecco et Sandrine Caddéo montrent que des sous types de constructions présentant des dislocations doubles (pronoms forts + substantifs) peuvent être distinguées, à la fois par des propriétés micro (type de pronom) et par des propriétés macro (« éloignement » des éléments disloqués).

Frédéric Sabio propose de même une typologie de structures présentant des « contrastes » : *c'est pas l'appareil qui vaut cher. C'est les objectifs* ; *Ce ne sont plus les pays qui intéressent mais les gens. Pierre est venu, mais pas Marie* sur la base de la distinction entre propriétés d'ordre macro et propriétés d'ordre micro : certains contrastes imposent que la construction soit produite selon plusieurs unités macrosyntaxiques.

Deux autres articles abordent le problème de la description par une interrogation sur les données.

Bertille Pallaud s'interroge sur le statut des interruptions et des fragments de mots. Elle propose de les inclure dans la description dans la mesure où leurs occurrences répondent à des distributions contrainte et où « leur présence témoigne d'une réflexion ou d'une activité linguistique se traduisant par une interruption de morphèmes en cours d'énonciation ». C'est dire que l'articulation entre signifiant et signifié est présente dans ce domaine et qu'il est donc justiciable d'une étude proprement linguistique.

Marie-Noëlle Roubaud prolonge la réflexion du groupe sur les productions des locuteurs aphasiques, qui nous permettent de mieux cerner les limites du système. Elle aborde le problème de la présentation des transcriptions de telles données sous l'angle limité de l'emploi des signes API. « Le problème n'est pas simple : tout transcripteur, face à la recherche du mot par le locuteur aphasique, fait des hypothèses sur le morphème recherché ; que doit-il alors transcrire ? L'article expose ces difficultés et apporte des éléments de réponse pour l'usage des notations phonétiques ».

Un dernier groupe d'article propose de faire le lien entre l'activité de description linguistique et certaines applications pratiques qui lui sont liées.

La construction d'outils pour l'exploitation de corpus informatisés est une des nouvelles priorités du groupe. Au delà des concordanciers et des taggeurs, une question centrale pour aider la description syntaxique est de construire des « parseurs » adaptés aux données orales. Mireille Bilger et Estelle Campione proposent une étude préalable à la construction de tels outils. A l'inverse de l'approche dominante en TAL qui s'efforce de mettre au point des stratégies d'analyse partielle (ou shallow parsing) des énoncés, faisant apparaître des segments intermédiaires entre le mot et la « phrase », qui sont le plus souvent des suites de mots très courtes, elles proposent une stratégie d'analyse partielle non pas en segments minimaux, mais au contraire en « séquences » maximales, sans entrer dans le détail de leur composition.

Le groupe se pose la question de la présentation des résultats des recherches sous la forme synthétique d'une grammaire du français à partir de corpus. La question du métalangage et de la terminologie utilisée dans

un tel ouvrage est alors cruciale si l'on veut qu'il soit utilisable pour tous. Dans ce cadre, José Deulofeu présente une critique de la notion, en apparence indispensable, de « type de phrase » et montre la difficulté de la remplacer par une autre terminologie reflétant mieux l'état des descriptions linguistiques.