

Le distributionnalisme

Introduction

Le distributionnalisme (ou linguistique distributionnelle) apparaît aux États-Unis dans les années 1930 grâce à ses deux fondateurs les linguistes américains Léonard Bloomfield et Zellig Harris. En effet, Bloomfield publie en 1935 à Londres son ouvrage intitulé « Language » (en français *langage*) et dans lequel il pose les bases de la linguistique distributionnelle. Quelques années plus tard, Harris publie en 1951 à Chicago un ouvrage intitulé « Methods in structural linguistics » (en français *Méthodes en linguistique structurale*). Dans son ouvrage, Harris présente la méthode scientifique de description des langues. C'est pourquoi on dit que la méthode du distributionnalisme est essentiellement descriptive.

A sa naissance, le distributionnalisme apparaît comme une critique de la « *grammaire fondée sur fondée sur une vue introspective de la langue et un psychologisme élémentaire, et comme un approfondissement et une systématisation d'une analyse combinatoire (des combinaisons possibles des éléments de la langue) connue depuis le XIXe siècle* » (Jean Dubois)

La linguistique distributionnelle s'est beaucoup inspirée de la linguistique européenne à travers ses courants représentatifs : le structuralisme saussurien, l'école de Prague, N. Troubetskoï et R. Jakobson.

Les principes théoriques du distributionnalisme :

Selon Jean Dubois, les principes théoriques et la méthodologie du distributionnalisme peuvent être définis comme suit :

1. L'objet de la linguistique est la description d'une langue, considérée comme un système formel, un ensemble de règles capables par leur fonctionnement de rendre compte des phrases produites. Le linguiste définit donc un corpus, c'est-à-dire un ensemble d'énoncés. Ce corpus est un échantillon de la langue, et c'est à partir de lui que seront définies les règles.
2. Si les règles dégagées permettent de rendre compte du fonctionnement d'une langue à un moment donné du temps, ceci signifie que la description du système qu'elles constituent sera **synchronique**. L'analyse **diachronique** se fera dans deux perspectives différentes : on pourra envisager le passage dans le temps d'un système à un autre système : autrement dit, on définira la structure des règles fonctionnant en 1660 (d'après un ensemble de textes) et celle des règles fonctionnant

en 1960 et on envisagera les règles de transformation qui permettent de passer de l'une à l'autre, selon une méthode proche de celle que la grammaire comparée connaissait mais avec cette différence essentielle que l'évolution est celle du système entier et non d'une de ses parties.

3. La linguistique distributionnelle se place donc devant un texte fini, représentatif, synchronique (le mot texte étant pris ici en un sens général). Ce dernier est devenu l'objet unique d'analyse. En empruntant la terminologie moderne, on dira que le linguiste distributionnaliste limite son objet aux performances, aux énoncés réalisés. Il va alors essayer d'en donner une description formalisée ; il en définira la structure par sans faire appel aux facteurs que constituent le sujet et la situation. Il s'agit donc d'une linguistique immanente.

La méthode du distributionnalisme :

Selon l'Encyclopédie Universalis, la méthode du distributionnalisme consiste à recueillir un « corpus » (c'est-à-dire un ensemble homogène d'énoncés considéré comme représentatif de la langue à étudier), puis à segmenter ce corpus.

La technique de segmentation repose sur la comparaison de morceaux d'énoncés grâce auxquels, de proche en proche, on identifie les unités minimales de signification, appelées morphèmes.

La notion clé est ici celle de « distribution » des unités sur la chaîne parlée ou écrite : la distribution d'un élément se définit comme la somme des environnements de cet élément (c'est-à-dire des autres éléments qui l'entourent) dans les énoncés du corpus.

Sur ces bases, la hiérarchie des constituants d'une phrase (depuis les morphèmes isolés jusqu'à la phrase entière) peut être représentée de façon rudimentaire ; ce type de représentation, où chaque constituant s'emboîte dans un constituant de niveau supérieur, est connu sous le nom d'« analyse en constituants immédiats ».