

« Empirisme et sociolinguistique cladistique : analyse d'un réseau d'idiolectes salentins »

La cladistique (du grec *clados* = branche), ou *systématique phylogénétique*, est une approche classificatoire qui a émergé autour de 1950 avec les premières études de l'entomologue allemand W. Hennig sur l'analyse génétique-typologique des espèces vivantes. Elle est rarement appliquée à des données de "première main" et rarement utilisée en sociolinguistique. Les résultats obtenus dans le cadre de notre étude se prêtent d'une part à une analyse sociolinguistique réticulaire à la Labov ou à la Milroy, d'autre part ils offrent une perspective de taxinomie quantitative très solide, qui diffère de l'approche classique en termes de traitement de données. Les termes comme "distance cumulative, dissonance, isoglosse" laissent désormais la place à une interprétation plus fine de configurations à géométrie variable, visibles sur les cladogrammes générés par notre approche cladistique automatisée.

Cette méthode est utilisée avec un grand succès dans le domaine de la génétique. Ici, elle sert au linguiste pour observer les réseaux de communication parlée, représentés par les différents systèmes dialectaux qui sont dynamisés par l'analyse sociolinguistique que la cladistique permet d'intégrer.

Cette science dépasse donc la classification généalogique grâce à l'analyse taxinomique qui offre des moyens de classification automatique utilisables également en *microdialectologie* et en *sociolinguistique*.

Le corpus analysé ici se compose de 35 variables phonologiques fondamentales dans le champ philologique de référence, recueillies dans six points d'enquête faisant partie du Salento méridional. Le questionnaire dialectologique a été soumis à 64 locuteurs salentins différenciés par âge, genre, situation socioculturelle, niveau d'étude.

Avant l'analyse cladistique, les matériaux recueillis ont été soumis à une analyse diachronique qui a consisté à construire des arbres de dérivation à partir de l'étymon latin. Ce processus a permis de détailler les différentes étapes évolutives des variables phonologiques comme dans la figure 1 :

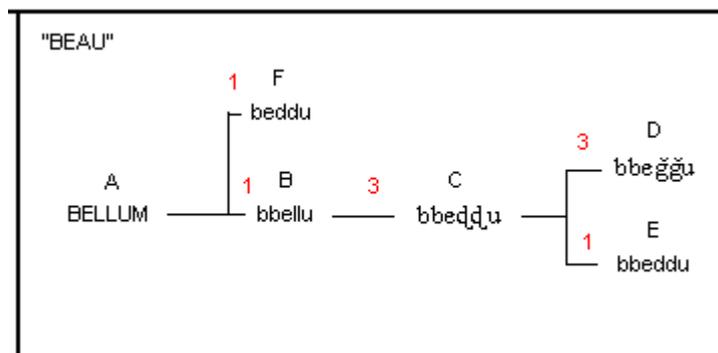


Figure 1 : arbre diachronique du mot BELLO < bellum (beau) ; les lettres indiquent les différents états évolutifs tandis que les chiffres en rouge indiquent l'indice de pondération appliqué en fonction du marquage phonologique (de 1 à 5).

Dans l'arbre évolutif du mot BELLUM chaque état (A, B, C etc.) peut correspondre à une génération précise, ou à une catégorie socio-culturelle, ou encore à une variante géographique. Les passages ont été pondérés en attribuant un poids de 1 à 5 en fonction des conditions de marquage phonologique ; ont suivi l'indexation et l'orientation des arbres (*A > B > C > D etc...), ainsi que l'explicitation des relations de réversibilité (A : B B : C A : D

A: ., où A peut évoluer en B, et B en C mais l'inverse n'est pas valable). Les matrices ainsi constituées ont généré des cladogrammes comme celui de la figure 2:

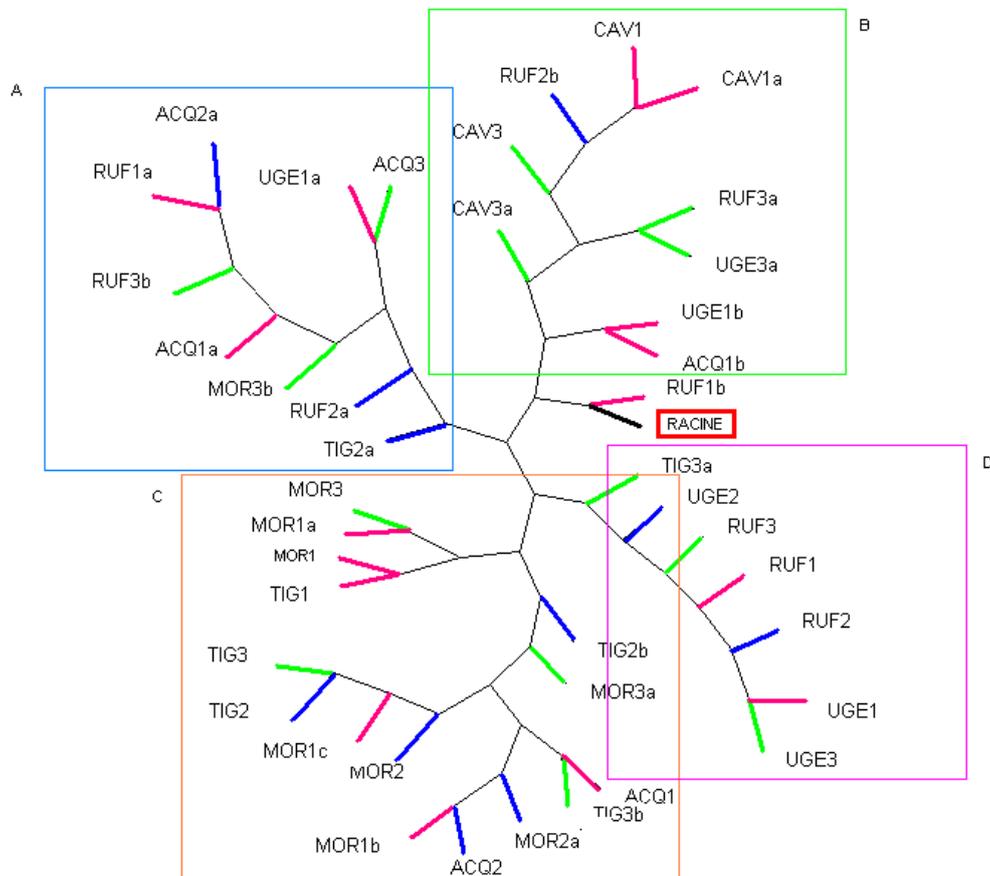


Figure 2 : cladogramme représentant la configuration sociolinguistique à géométrie variable de l'aire salentine méridionale. Les différentes couleurs indiquent les trois tranches d'âge : en bleu les « moyens », en fuchsia les « aînés » et en vert les « jeunes ». Les sigles indiquent les points d'enquêtes, tandis que les lettres en minuscule se réfèrent aux variantes relevées dans chaque variété étudiée.

L'analyse cladistique n'a pas pour but de mesurer des distances, ni d'établir des ressemblances ou correspondances régulières entre dialectes; elle vise plutôt à préciser le degré de parenté ou de convergence structurale entre les variétés dialectales. Après l'avoir appliquée à des diasystèmes dialectaux plus amples (Oïl e IRM¹), nous verrons que dans le cas du Salentin méridional, la cladistique permet d'analyser la granularité d'un réseau sociolectal restreint en croisant différents facteurs sociolinguistiques afin de rendre compte de la divergence typologique du paysage linguistique contemporain, interprétable à travers la ramification et la structure interne de chaque arborescence, tout en tenant compte du polymorphisme et des doublets, qui posent souvent des problèmes insurmontables aux approches taxinomiques classiques. Ceci n'aurait pas été possible avec une méthode d'analyse globale ou cumulative.

L'analyse cladistique appliquée devient ainsi un outil complémentaire des méthodes d'analyse de la sociolinguistique des réseaux de Milroy et Labov, car elle permet en effet de saisir la *réticularité* sociolinguistique à échelle *microdialectale*, avec des données de première main. Cette approche permettra donc de discuter et de revisiter la notion de réseaux sociolinguistiques, centrale dans le paradigme de la sociolinguistique moderne.

¹ Italo-roman méridional.

Mots clés : cladistique, salentin, dialecte, typologie, phonologie, lexique, dialectologie, réseau sociolinguistique, communauté linguistique, idiolecte.