# Salentinian sociolinguistics through cladistics: the geolinguistic and social complexity of a (micro)diasystem.

**Antonella Corvaglia (Sorbonne Paris 4-Inalco/Plidam)**

## INTRODUCTION

*Cladistics* (from Greek *kládos 'b*ranch'), also called phylogenetic systematics, is a classificatory method that emerged around 1950 when the German entomologist W. Hennig contrived a classificatory method for genetic analysis of living species based on typological clues ordered by derivational chains. Phylogenetic construction is based on the principle of "descent with modification": the characters observed in two or more species that indicate a close relationship are those inherited from their common ancestor. *Numerical Taxonomy* consists of estimating the linguistic distance between pairs of languages and calculating evolutionary trees or networks to produce linguistic classifications. This approach is generally used in dialectometry. But our approach here is closer to typological sociolinguistics than to genetic linguistics. We adopt a strategy enabling us to integrate linguistic hypotheses before making inferences on the evolution of linguistic traits and languages, and potentially to refute them.To test the heuristic value of this methodology, we apply cladistics to dialectal data from different sources, through an original coding of philological derivations.

## METHODOLOGY

The starting point of our study was a questionnaire constructed following the conventional criteria of Southern Italo-Romance dialectology and sociolinguistics and taking into account several phonological consonantal variables which are specific of the southern Salentinian dialects. The survey was undertaken by me (because I'm a native Southern Salentinian dialect speaker), with the collaboration of town councils. The Southern Salento social background is characterized by a network of small and densely populated towns of around 12,000 inhabitants each [19]. The towns surveyed are the following: Ugento (UGE), Ruffano (RUF), Acquarica del Capo (ACQ), Morciano di Leuca (MOR), Tiggiano (TIG), Cavallino (CAV). Cavallino was chosen as a northern point for the purposes of geolinguistic comparison. The dialect native speakers was 64: 6 men and 6 women in each town, categorized according to age (<30, 31-50 and >50), sex, profession, and education level. All informants were bilingual speakers of Italian and Salentinian. The stimuli were presented in Italian, but interactions with the interviewer were mostly in Salentinian.

### Salentinian consonantal corpus

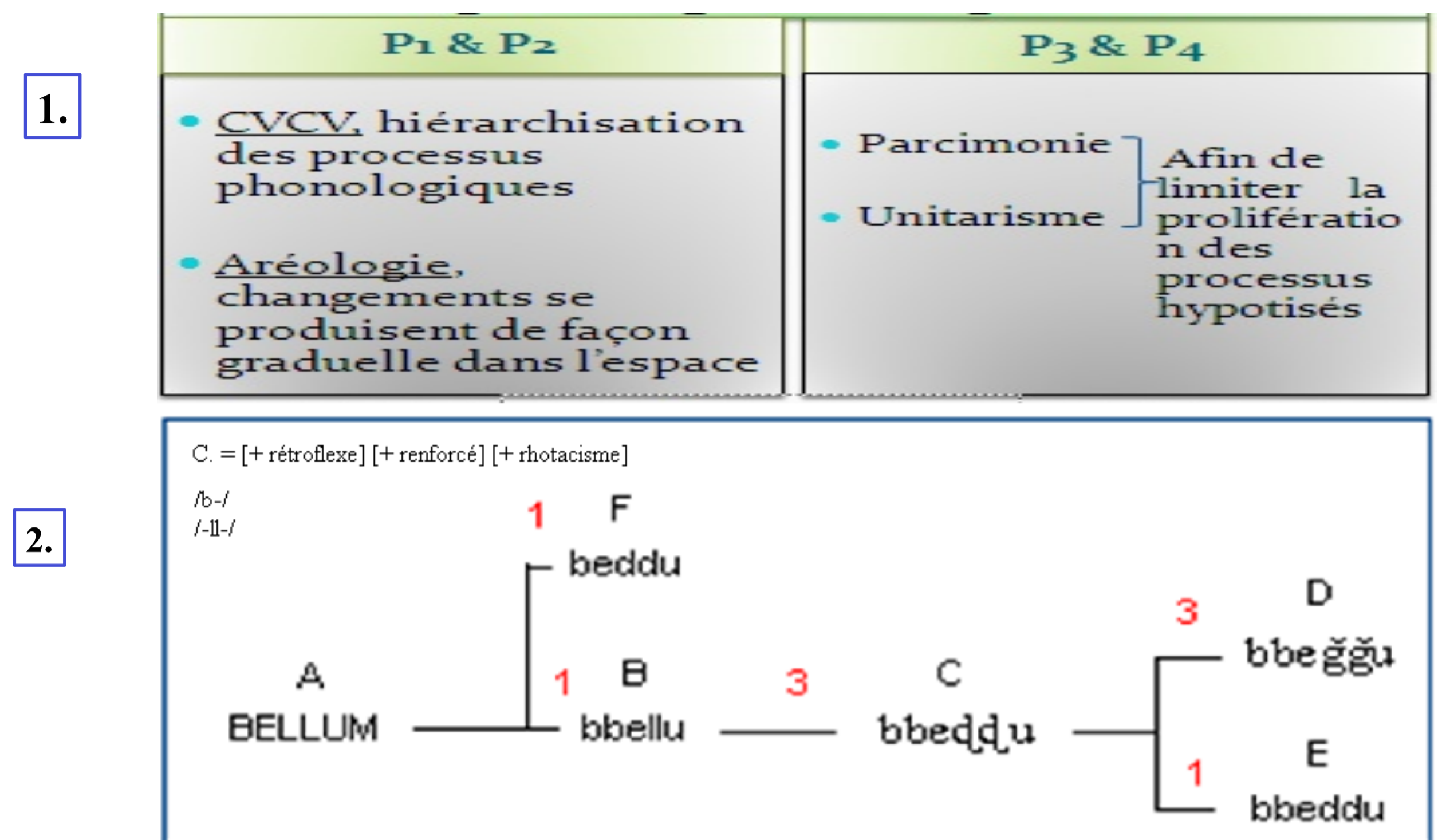314 phonological items including 95 consonantal variables from which we selected 35 for cladistic processing:

A. = Stops [α voiced]
{gatto, grande, fegato, litigare {dente, ditale, cado, credo}}

B. = Sibilants /s/ retraction + /str-/ reduction > Σ
{maestra, finestra, minestra {vostra, mostrare}}

C. = Laterals: retroflection, gemination and rhotacization
{gallo, quello, bello {capelli, cavalli, quelli}}

D. = Palatal laterals (palatalization of -lj-) [- lateral]
{figlio, famiglia, moglie, voglia}

E. = Palatal affricates
{giovedì, fuggire, gelo {ceci, cenere}}

F. = Voiced labial stops, with [O] [+continuant] [+tense] spirantization, gemination)
{bocca, braccio, basso, febbre, tavola {bere, battere}}

G. = [+ tense] [+ palatal] [- continuant, labial] [α voiced]
{vedo, vieni {vento, vomito}} (e.g., Lat. *vomitare* > Sal. [ˈommiku]/[ˈvomitu]/ [ˈvommuku]/[vummaˈkare]).

Each item on the question list was read in Italian once or twice to the informant, who was asked to translate and repeat the expected form in the Salentinian dialect, twice in isolation and once in a spontaneous sentence. The interviews were recorded with a SONY ECM-MS907 microphone and a SONY MZ-N710 minidisc recorder.
The phonetic transcription was performed using SoundForge 7 and Praat to check systematically the auditory impressions of the transcriber.

## A Cladistic-typology approach of socio-linguistic salentinian variation

### Diachronic analysis

Three principles *(1)* and Diachronic tree of *\*BELLUM (2)*



| P₁ & P₂ | P₃ & P₄ |
|---|---|
| 1. CVCV, hiérarchisation des processus phonologiques  Aréologie, changements se produisent de façon graduelle dans l'espace | Parcimonie  Unitarisme — Afin de limiter la prolifération des processus hypotisés |

2. C. = [+ rétroflexe] [+ renforcé] [+ rhotacisme]



The capital letters represent the different evolutionary states of the Latin variable, while the numbers in red indicate the weights applied according to conditions of phonological markedness in terms of frequency and articulatory difficulty (Index Weighting range is fixed from 1 to 5 points). In this way the data was indexed, weighted and directed to be processed with PAUP 4.0 (Phylogenetic Analysis Using Parsimony) in order to generate cladograms such as the one shown in the figure.

### Cladistic analysis and results

**Homogenization**

**Stabilisation**



**Not significant socio-linguistic Difference**
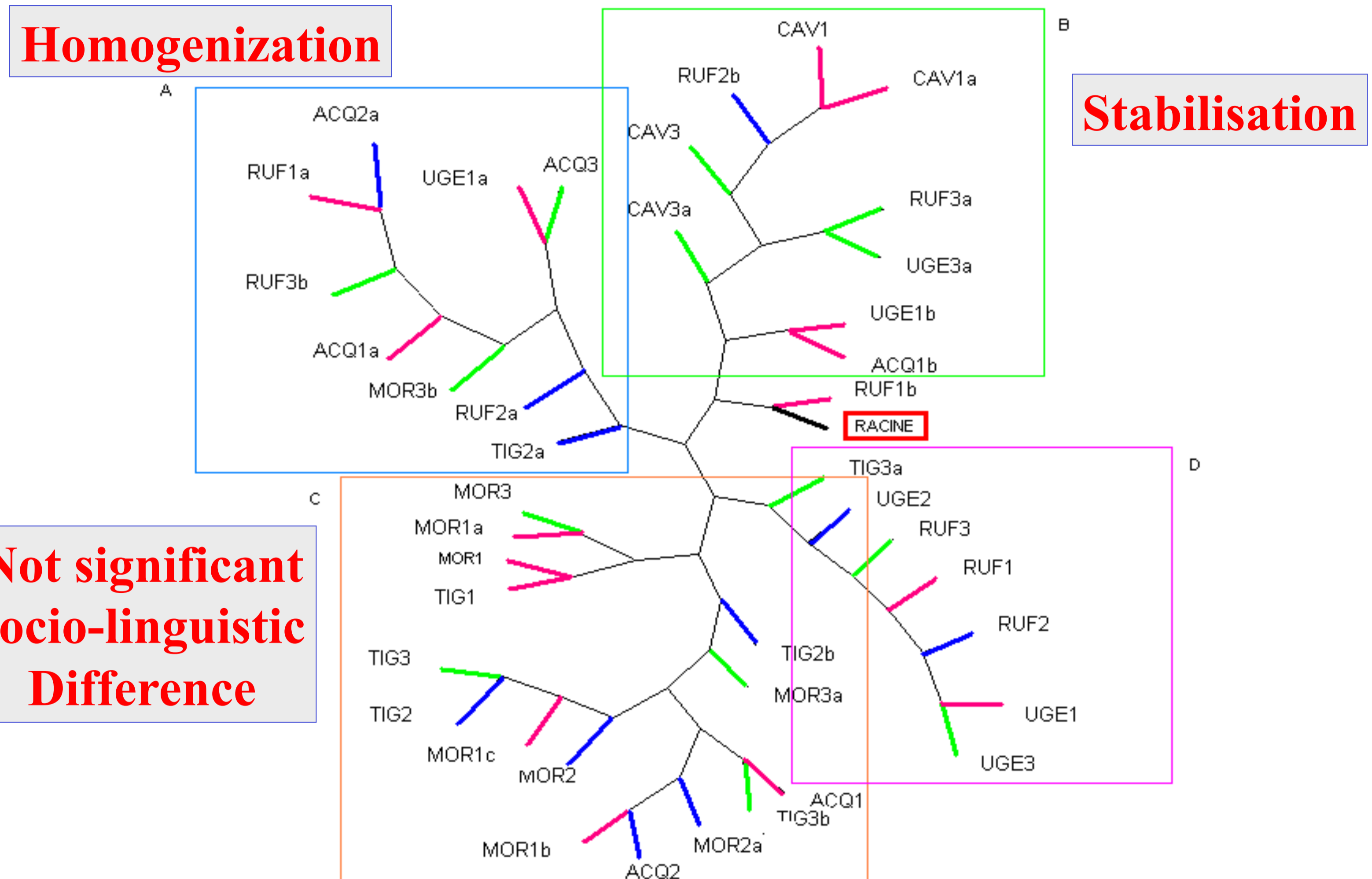
**EXPLANATION OF THE CLADOGRAM**
The cladogram shows the sociolinguistic shape of the region of Salento according to our cladistic analysis. Colors at the end of the branches indicate different generations (1, 2, 3); lowercase letters indicate idiolectal variants of each variety. For the word *BELLO* for example, in variety MOR and generation 1 (the oldest) we have 4 idiolectal variants: MOR1, MOR1a, MOR1b, MOR1c, corresponding to the heterogeneous answers given by the informants. The sociolinguistic interpretation of the cladogram must be made according to the presence/absence of sinapomorphy in each micro-diasystem. We have a clad (D) of the type Ugento-Ruffanese that groups all main variants (UGE 1-2-3 and RUF 1-2-3). This branch consists of the three generations of varieties geographically in the center of the studied area. It is adjacent to the branch of the central-northern type represented by B, the only type that includes the varieties of Cavallino accompanied by an idiolectal synapomorphy (RUF3a, UGE3a, UGE1b, ACQ1b, RUF2b). In the diasystem A sinapomorphy is perfectly balanced across generations (3 for each generation), with a dialectal characterization of type Acquarica-Ruffanese. Finally, the clad C is a little richer in sinapomorphy, with 15 variants, and represents a more important heterogeneity in comparison with the three other clads. Here we see a sort of *crescendo* of sinapomorphy: 4 for the young generation, 5 for the middle generation, and 6 for the older one. We also can see that sinapomorphy in green (3) belongs to the varieties MOR and TIG (that is to say Capo di Leuca, the most southern area); they are linked to the group in blue (2) of the same varieties, as well as ACQ and still to the same elements of group 1. As we see it, this clad belongs entirely to the extreme-southern type (MOR-TIG-ACQ) which is generationally and geolinguistically homogenous. The three generations of this cluster use the dialect in the same way, with no differentiation of age, sex, or socio-cultural level.
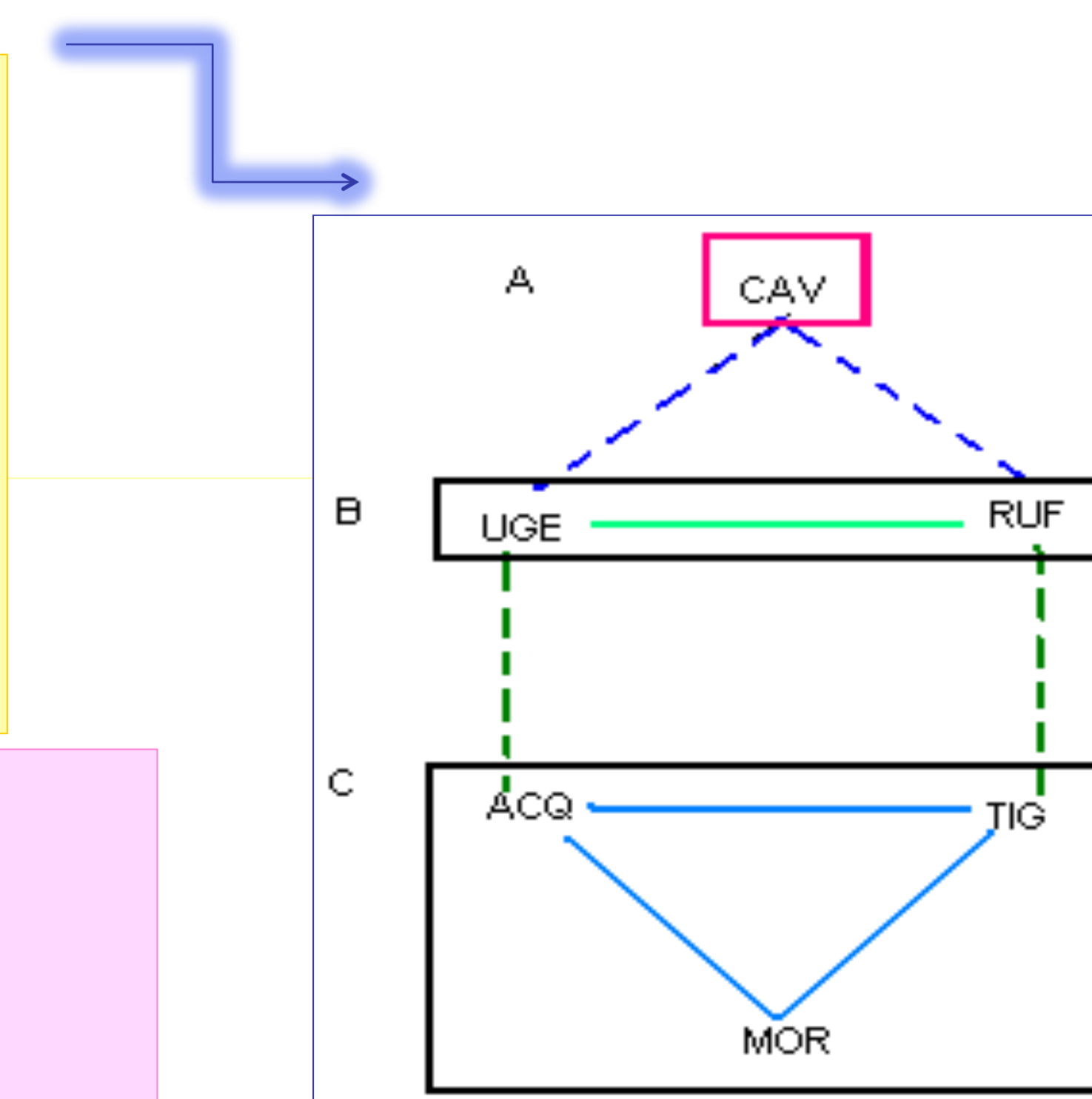
**CONCLUSION**
The Salento region's sociolinguistic variation does not seem to depend on age range as the three generational pools are often clustered. This can be explained by a high degree of cross-generational interactivity within the network as a trend to dialect norm synchronization, contrary to official findings which declare a diminution of the dialect's exclusive use and claim that the use of dialect is proportional to age (ISTAT data 2006 and 2007). This assumption is also disconfirmed by our surveys made in Salento where 77, 7% of speakers of every generation declared to use dialect every day. In other words, the generational distribution of sociolinguistic markers shows up as entirely homogenous and symmetrical in the Salento cladogram. The use of dialect, therefore, is not linked to age, a conclusion that agrees with the well known vitality of Southern Italian (and Salentinian) dialects.
these results imply that socio-cultural differences between speakers (differentiated by educational level, profession, age and sex, see § 2.1.1.) tend to be reduced by regular use of (and strong proficiency in) the local dialect. The Italianization of speakers' mother tongue has probably not yet played a very important role in this part of Salento, as the dialect remains the language of daily interaction and sociability.



**Same references :**
Calabrese, A. (1995). A constraint-based theory of phonological markedness and simplification procedures. *Linguistic Inquiry,* Vol. 26(3), pp. 373-463.
Gaillard-Corvaglia, A., Léonard, JL, Darlu, P. (2009). Approche cladistique des vocalismes d'Oïl à partir des données de l'ALF, dans BALI, *III Serie - Dispensa N. 32 2008*.
Gaillard-Corvaglia, A., Darlu, P., Léonard, JL. (2007). Testing cladistics on dialect networks and phyla (gallo-romance and southern italo-romance), *Ninth Meeting of the ACL Special Interest Group in Computational Morphology and Phonology*, 23–30 juin 2007, Prague.
Gaillard-Corvaglia, A. (2006). Lo spazio dialettale salentino: uno studio sociolinguistico. *Atti del convegno internazionale la Comunicazione Parlata*, Naples, février 2006.
Hennig, W. (1988). Systématique cladistique, *Société Française de Systématique*, 2ème Édition, Paris.