

# Modeling the Speech Community Through Multiple Variables: Trees, Networks and Clades

James A. Walker  
York University

Michael Dunn  
Uppsala University

Aymeric Daval-Markussen  
Aarhus University

Miriam Meyerhoff  
Victoria University of Wellington

**Introduction**

- Variationist sociolinguistics has been criticized for its focus on individual variables in isolation rather than taking into consideration the multiple variables used by members of the speech community to construct a 'lect'.
- Recent work has addressed the question of 'co-variation' by examining the interaction between multiple variables using statistical tests of correlation (e.g. Guy 2014; Tagliamonte & Waters 2011), though the results of such studies have proven inconsistent.
- Moreover, use of overall rates of occurrence in measuring correlation does not take into consideration effects of language-internal conditioning.
- Thus, correlation may not be able to tap into the linguistic system underlying the variation.
- We address the question of co-variation by taking a different statistical approach through the examination of an interrelated set of grammatical variables in a potentially polylectal speech community.

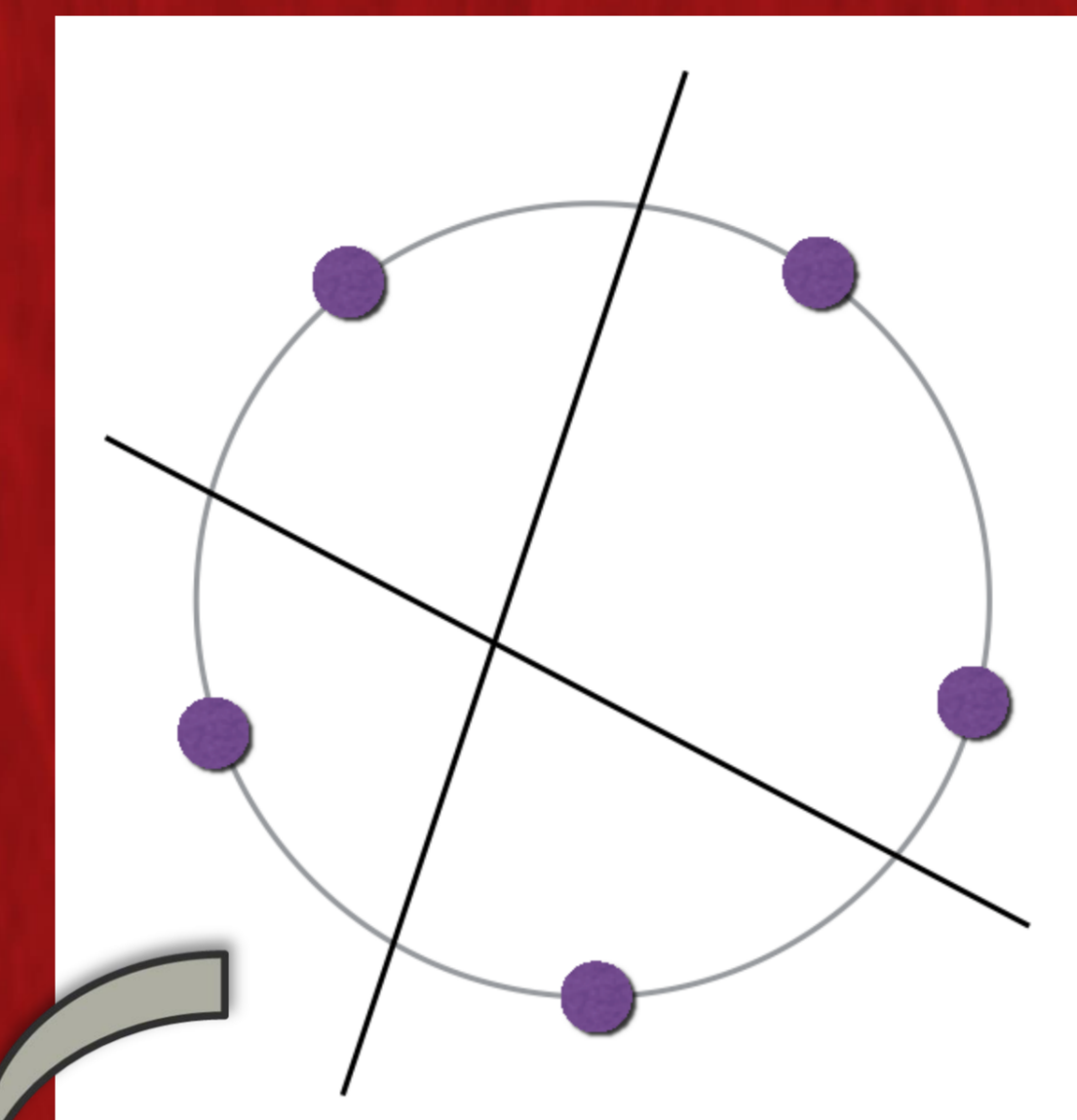
**Measuring distance between utterances**

1 = "utterance contains feature x"  
0 = "utterance does not contain feature x"

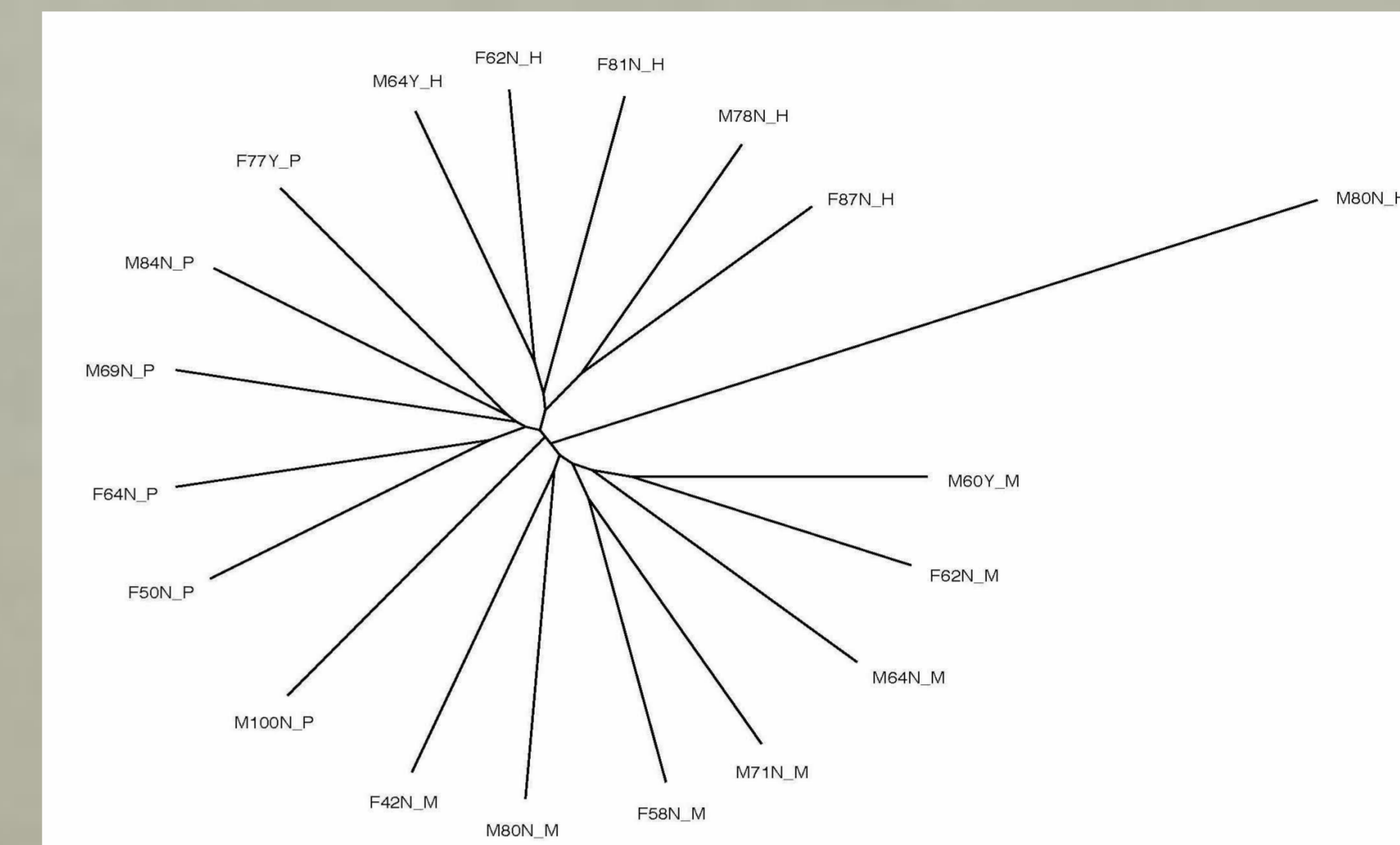
0011010000010100100100011111000  
0011100000001010000010101111000  
\* \* \* = 3/30  
distance = 0.10

001?01000101010100110101??101010  
1101010010010100001001?1111000  
\*\*? \*\* \*\*\*\* ?? \* \* = 10/27  
distance = 0.37

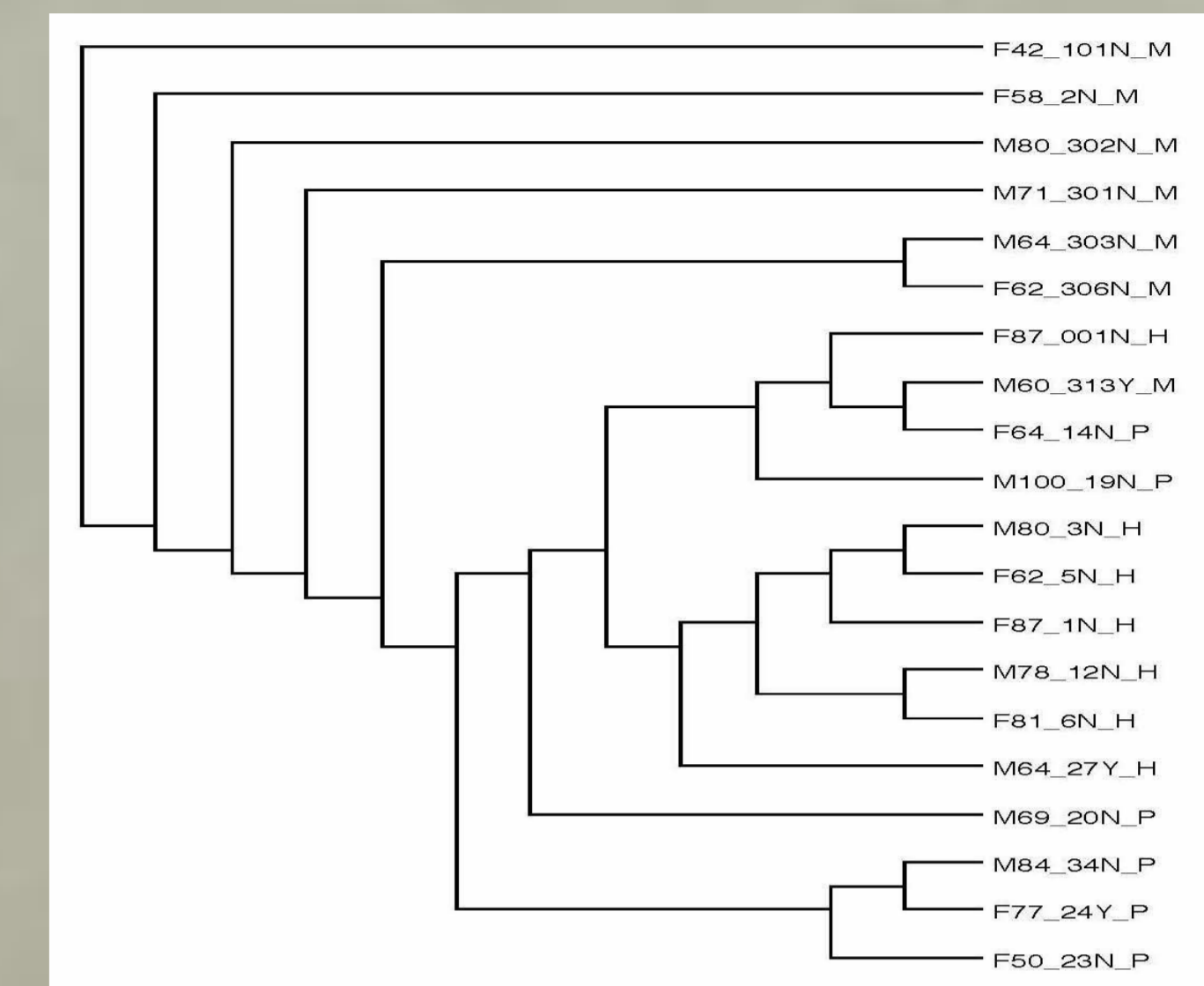
**NeighborNet graph:**



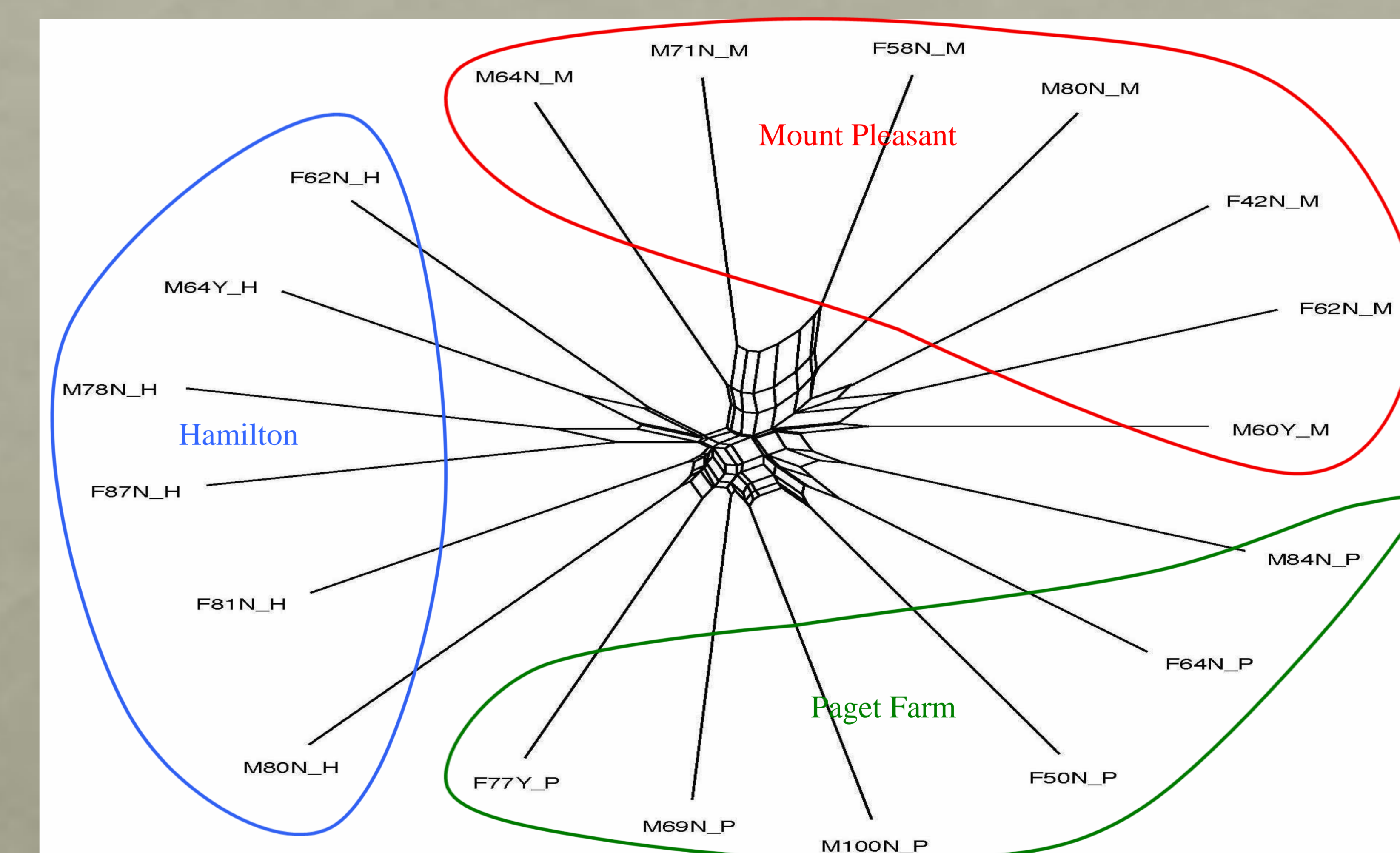
**Speaker Tree:**



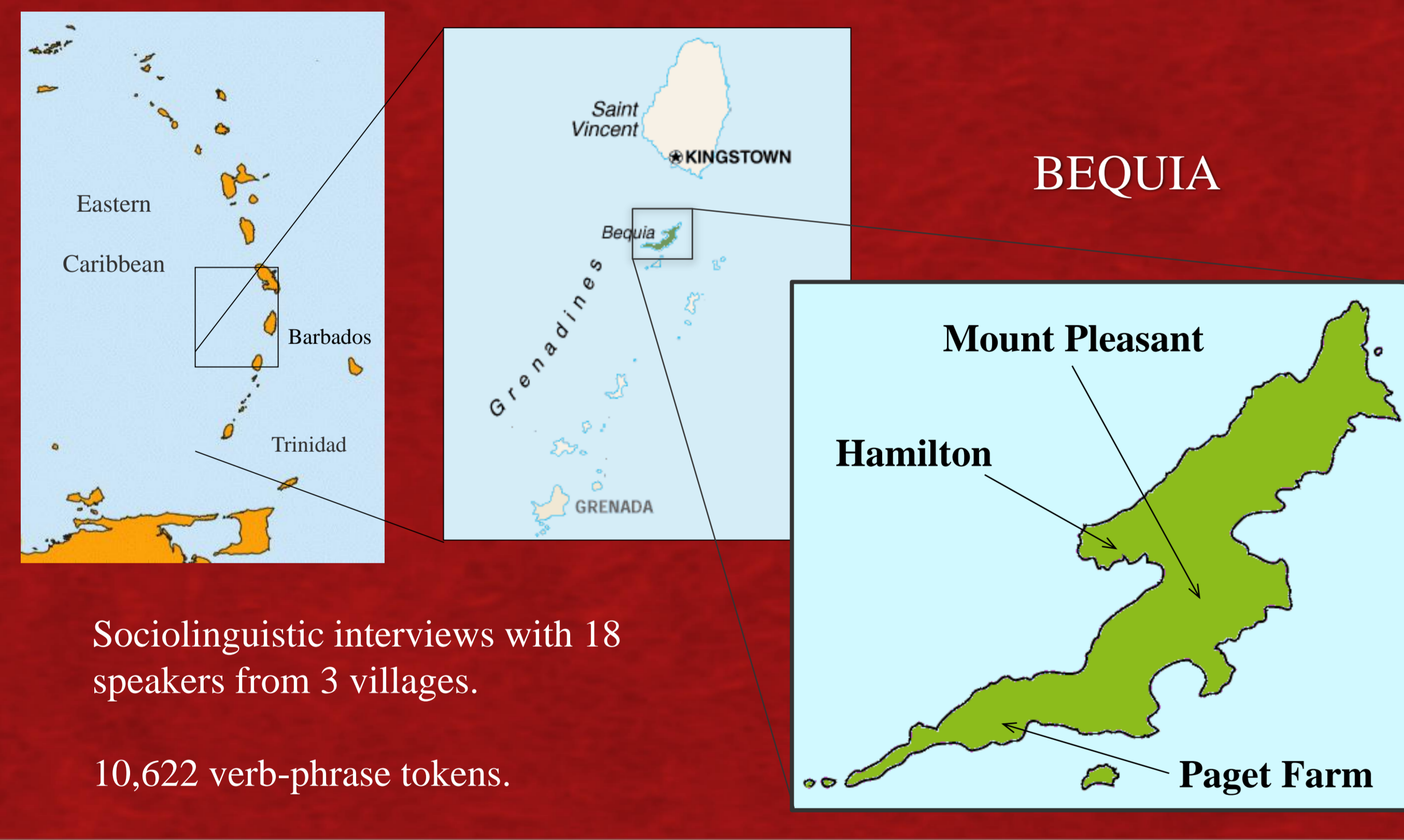
**Speaker Cladogram:**



**Speaker NeighborNet graph:**



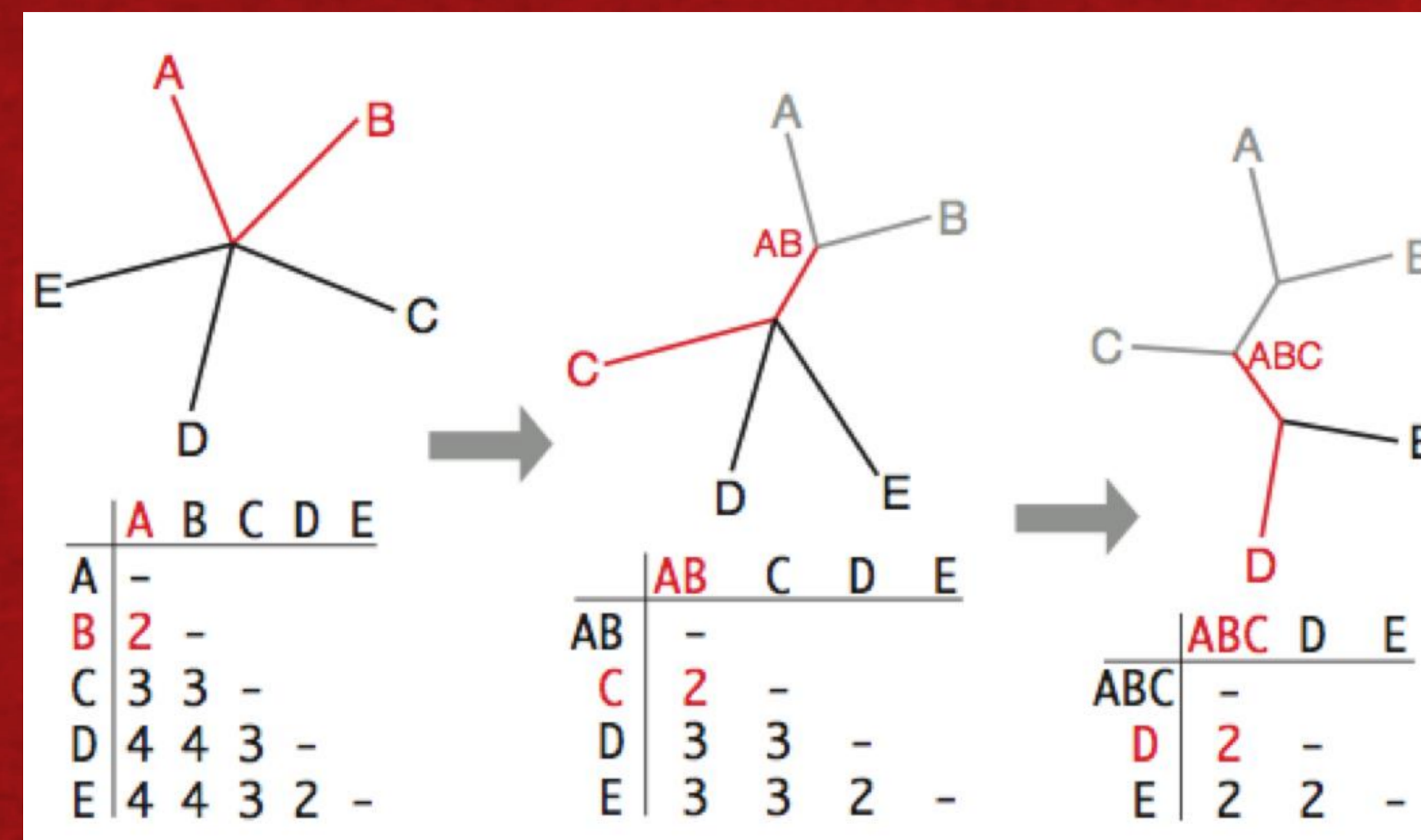
**St. Vincent and the Grenadines**



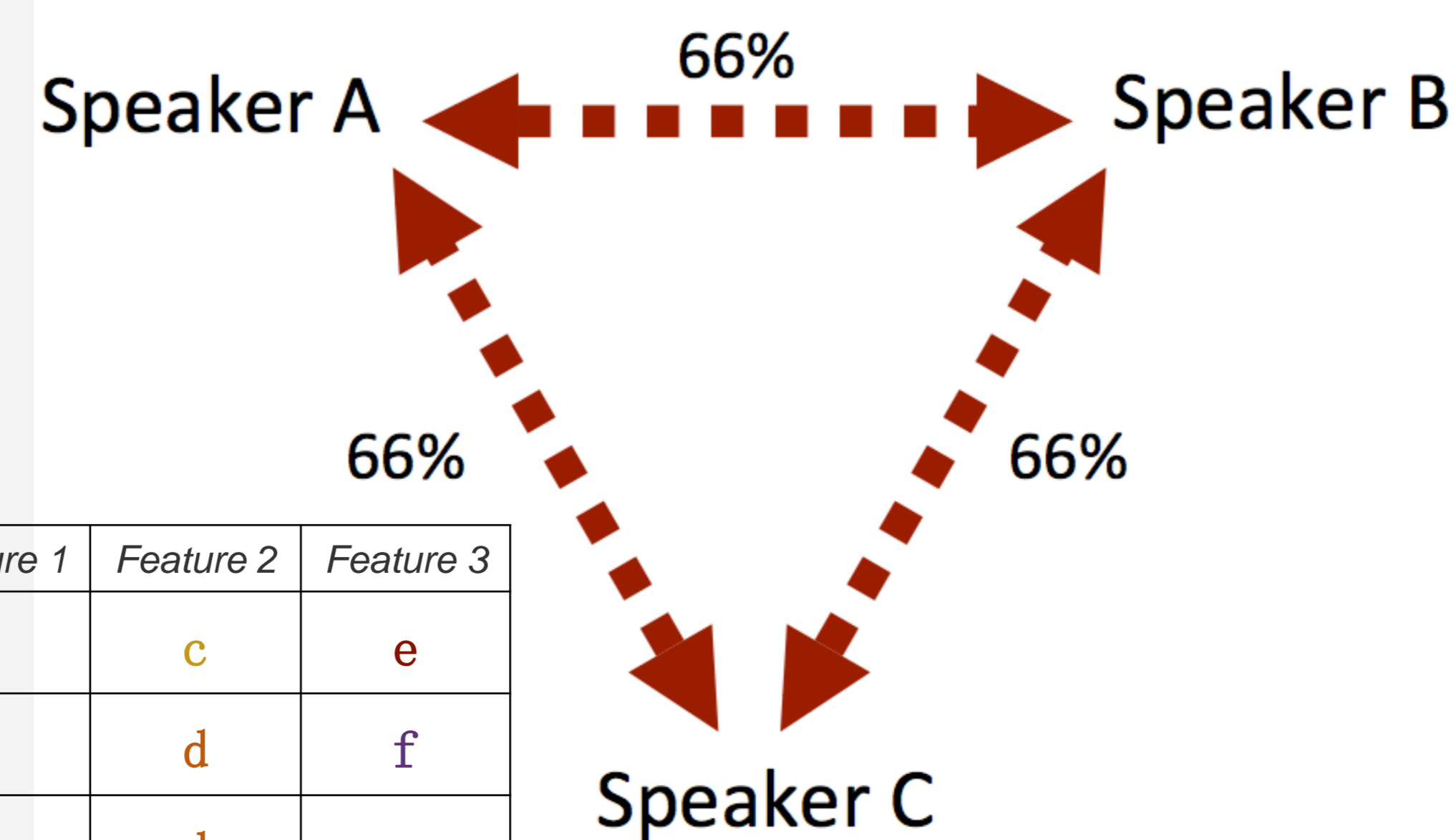
Sociolinguistic interviews with 18 speakers from 3 villages.

10,622 verb-phrase tokens.

**We could draw a tree using a neighbour joining algorithm ...**



**But not everything fits neatly into a tree ...**



**Conclusion**

- We have used neighbour-joining algorithms implemented in phylogenetic software to model the relationship between speakers in three ways: trees, cladograms and neighbour networks.
- The most consistent grouping of speakers according to the distribution of utterance types in their sociolinguistic interviews is according to the village they come from on Bequia.
- These results are consistent across models.
- These results are also consistent with those of previous studies of individual variables (see Meyerhoff & Walker 2013).
- We suggest that this approach may offer a more satisfactory answer to the question of co-variation.

**REFERENCES**

Bryant, D. & V. Moulton. 2004. NeighborNet, an agglomerative method for the construction of phylogenetic networks. *Molecular Biology and Evolution* 21: 255-265.

Meyerhoff, M. & J.A. Walker. 2013. Grammatical variation in Bequia (St Vincent and the Grenadines). *Journal of Pidgin and Creole Languages* 27: 209-34.

Guy, G.R. 2014. Bricks and bricolage: The question of sociolinguistic coherence. Keynote lecture, *New Ways of Analyzing Variation* 43, Chicago.

Tagliamonte, S. & C. Waters. 2011. Co-variation in the speech community: Methods for identifying innovators and their repertoires. Presented at Methods, University of Western Ontario.

Huson, D.H. & D. Bryant. 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* 23: 254-267.

Each token represents an utterance TYPE, a particular configuration of 'formal' and 'functional' characteristics:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
Tense/Aspect Marker	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Auxiliary/Modal	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Form of Copula	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Form of Lexical Verb	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Form of negation	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Following Grammatical Category	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Negative Concord	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Lexical Verb Code	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Temporal Reference	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Aspect	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Grammatical Person	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Subject Type	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Sentence Type	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Clause Type	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Line Number	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Context	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/

- Formal:**
- Tense/aspect marker
  - Auxiliary/modal
  - Form of copula
  - Form of lexical verb
  - Form of negation
  - Following grammatical category
  - Negative concord

- Functional:**
- Lexical verb
  - Temporal reference
  - Aspect
  - Grammatical person
  - Subject type
  - Sentence type
  - Clause type