Virtual Distances between Languages

Methods from Dialectometry and Stylometry in Digital Humanities

Dafydd Gibbon

Universität Bielefeld
Germany

Bielefeld-Abidjan Cooperation
2015-11-16

Background

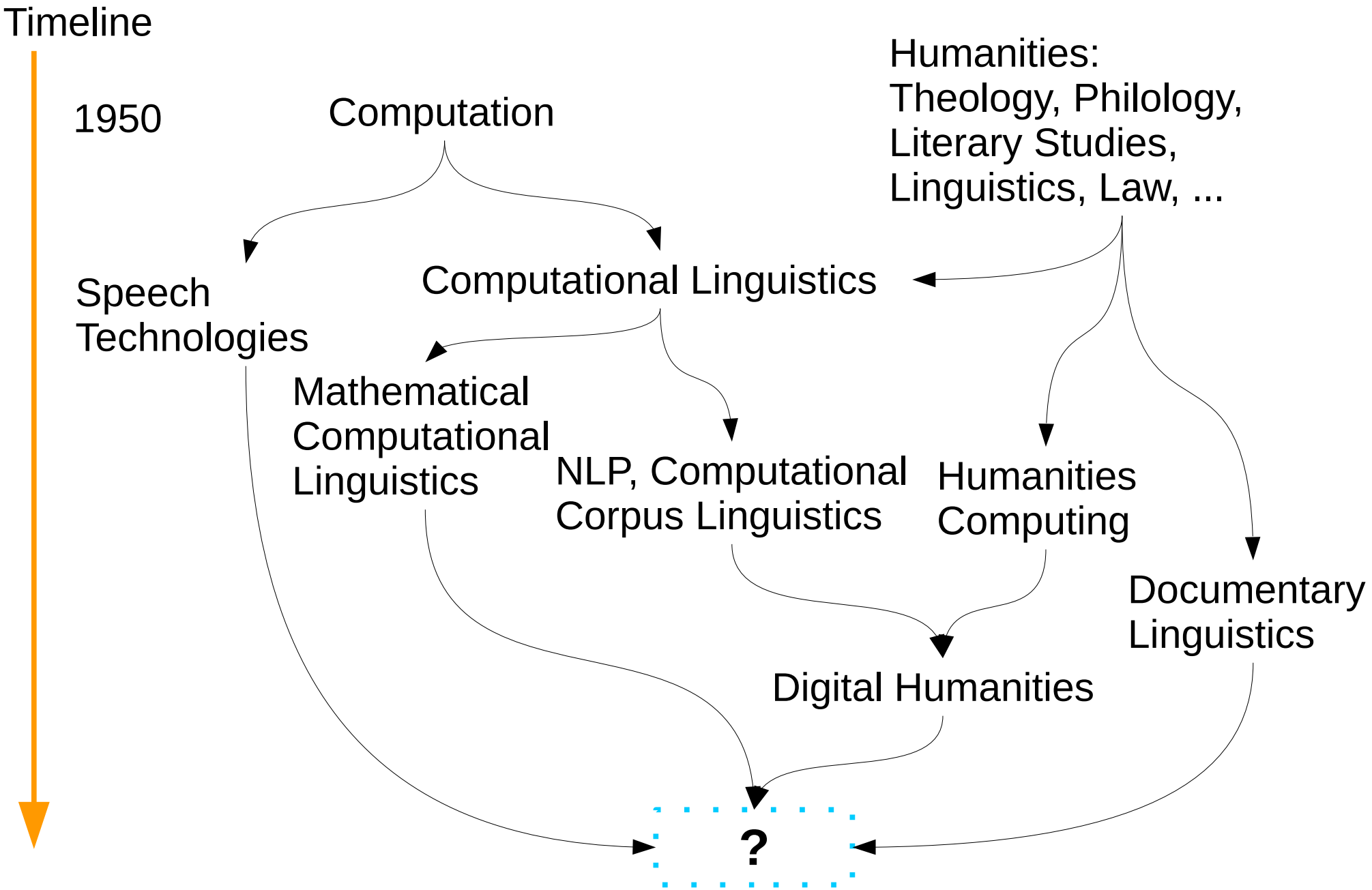
My work in teaching and in the field in West Africa has been mainly mainly in Côte d'Ivoire and Nigeria.

From fieldwork to Digital Humanities

Research goals:

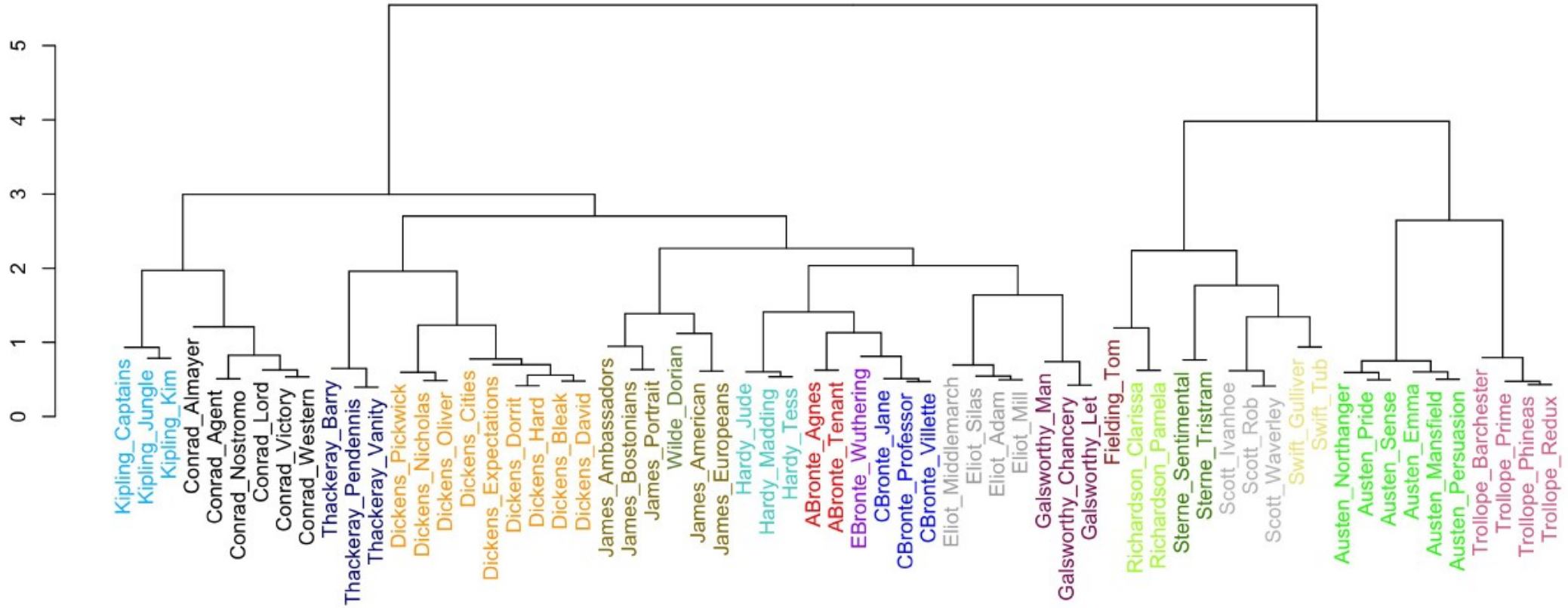
- Analysis of oral literature in a Digital Humanities framework.
- Phonetic analysis; sound-gesture synchronisation in speech and music.
- Visualisation of differences between languages as ‘virtual distances’.

A personal view of the ancestry of Digital Humanities

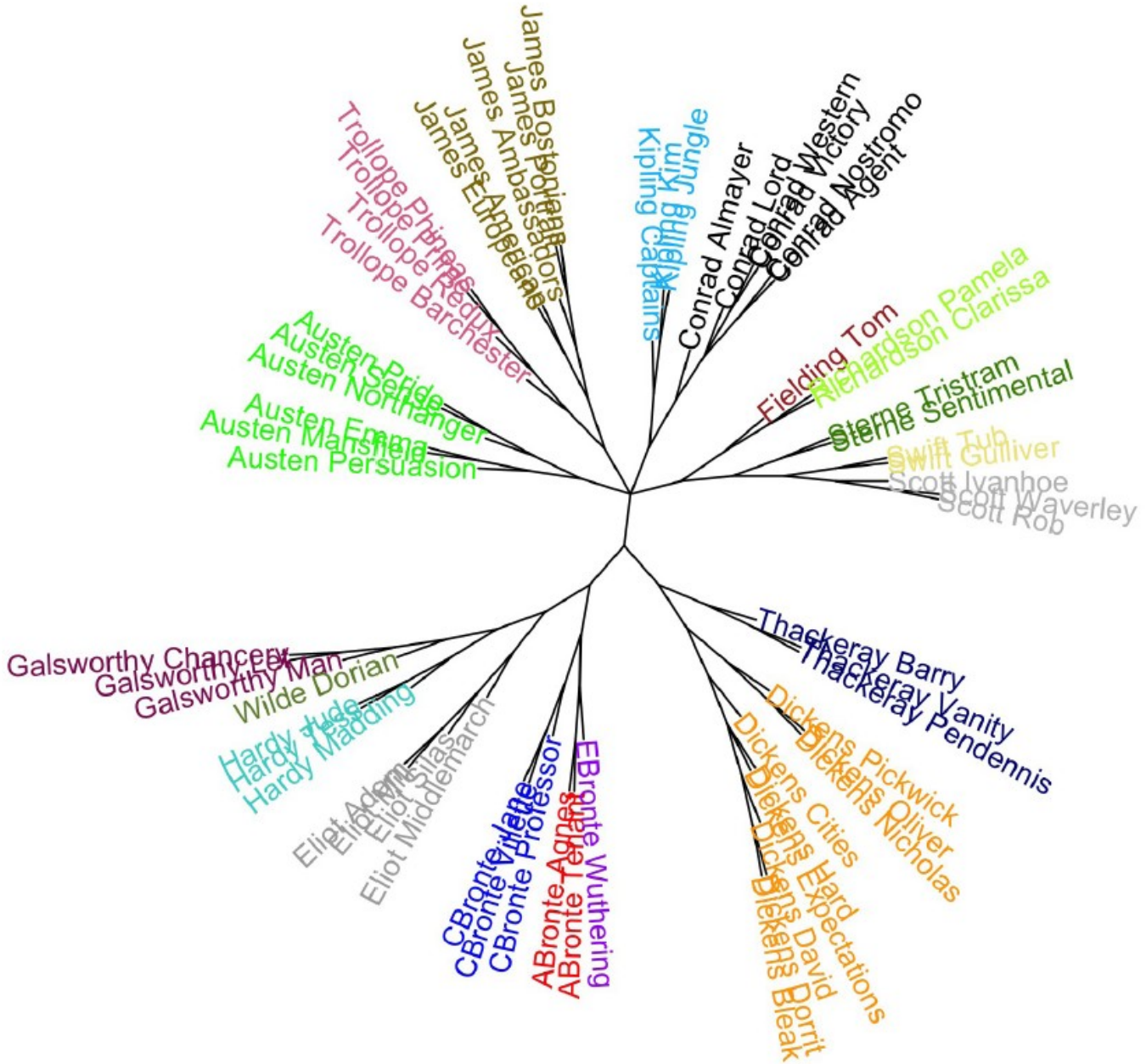


- Lexicostatistics
  - Morris Swadesh (Swadesh wordlist)
    - create wordlists for each language (interviews, corpus)
    - determine cognate words (related in form and meaning)
    - determine pairwise similarities → triangular distance table
    - create family tree using the table
    - Problems: accidental similarity, borrowing, taboo, ...
- Dialectometry
  - Two major centres in Europe:
    - Rijksuniversiteit Groningen
    - Universität Salzburg
  - Similar methodology in Digital Humanities:
    - Stylometry: authorship attribution, style typology
      - MFW: most frequent lexical words, most frequent grammatical words

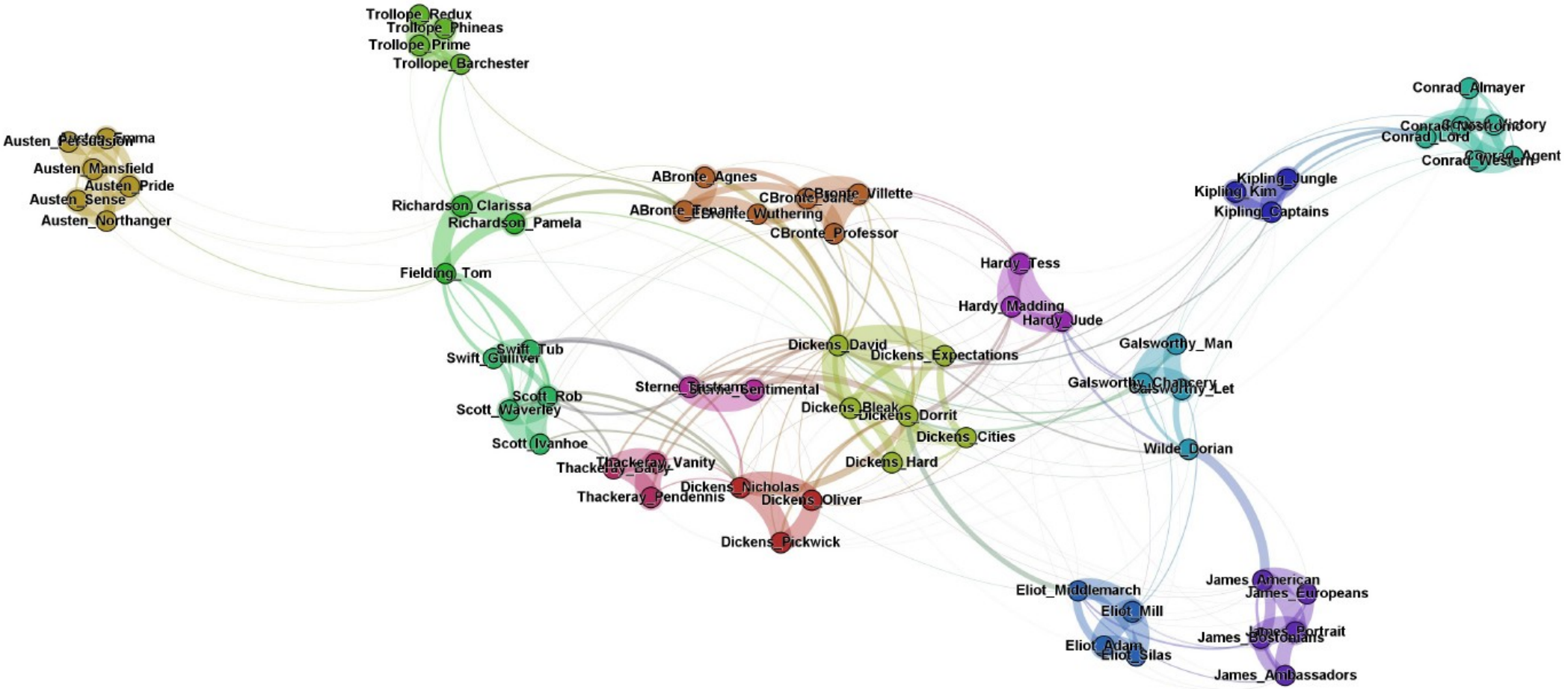
Stylometric visualisation – Maciej Eder, Kraków



Stylometric visualisation – Maciej Eder, Kraków



Stylometric visualisation – Maciej Eder, Kraków



Dialectometric visualisation – Dialectometry project, Salzburg

www.dialectometry.com

DIALEKTOMETRIE PROJEKT - SALZBURG | Dialektometrie | Ladinia | Frankreich | Toscana | Italien | Andere Projekte | Intern

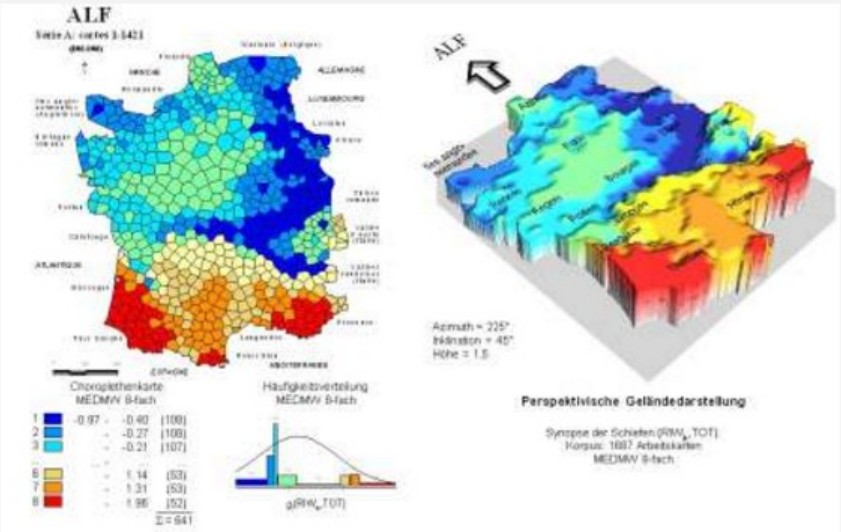
Dialektometrie (Einführungen in die Dialektometrie)

Die zentrale Aufgabe der Dialektometrie besteht in der Auffindung müglichst hochrangiger Ordnungsstrukturen in sprachgeographischen Netzen.

Das Forschungsinteresse richtet sich in erster Linie auf die regionale Distribution von Dialektühnlichkeiten, betrifft Dialektkerne und Übergangszonen, die durch eine mehr oder weniger geringe Dialektvarianz zwischen benachbarten Orten gekennzeichnet sind.

Als empirische Datenbasis dienen Sprachatlanten, die das dialektale Profil einer großen Anzahl von Erhebungsorten minuziös dokumentieren.

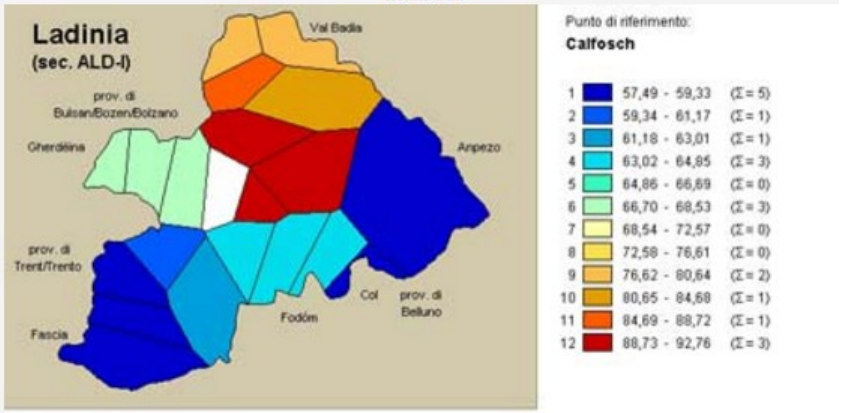
Verschiedene aus der numerischen Klassifikation gut bekannte Verfahren werden benutzt, um aus der immensen Datenfülle der Sprachatlanten zugrunde liegenden Basismuster zu abstrahieren und anschließend zu visualisieren.



Als Ergebnis darf nicht eine einzige Klasseneinteilung erwartet werden; vielmehr können durch verschiedene Verfahren jeweils andere Aspekte des gesuchten Basismusters aufgedeckt werden. Prinzipiell sind wir an der Vielfalt der taxometrischen Verfahren und den dazu gehörigen Resultaten bzw. an den daraus erfließenden linguistischen Interpretationen interessiert.

Dialektometrie Projekte

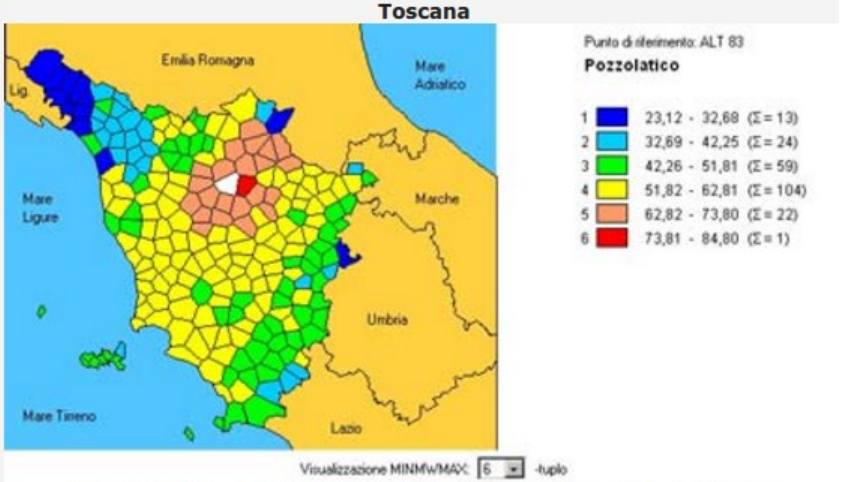
Ladinia



Visualizzazione MINMWMAX 12 -tuplo

Beschreibung Applikation Starten (Ein Projekt von Roland Bauer)

Toscana

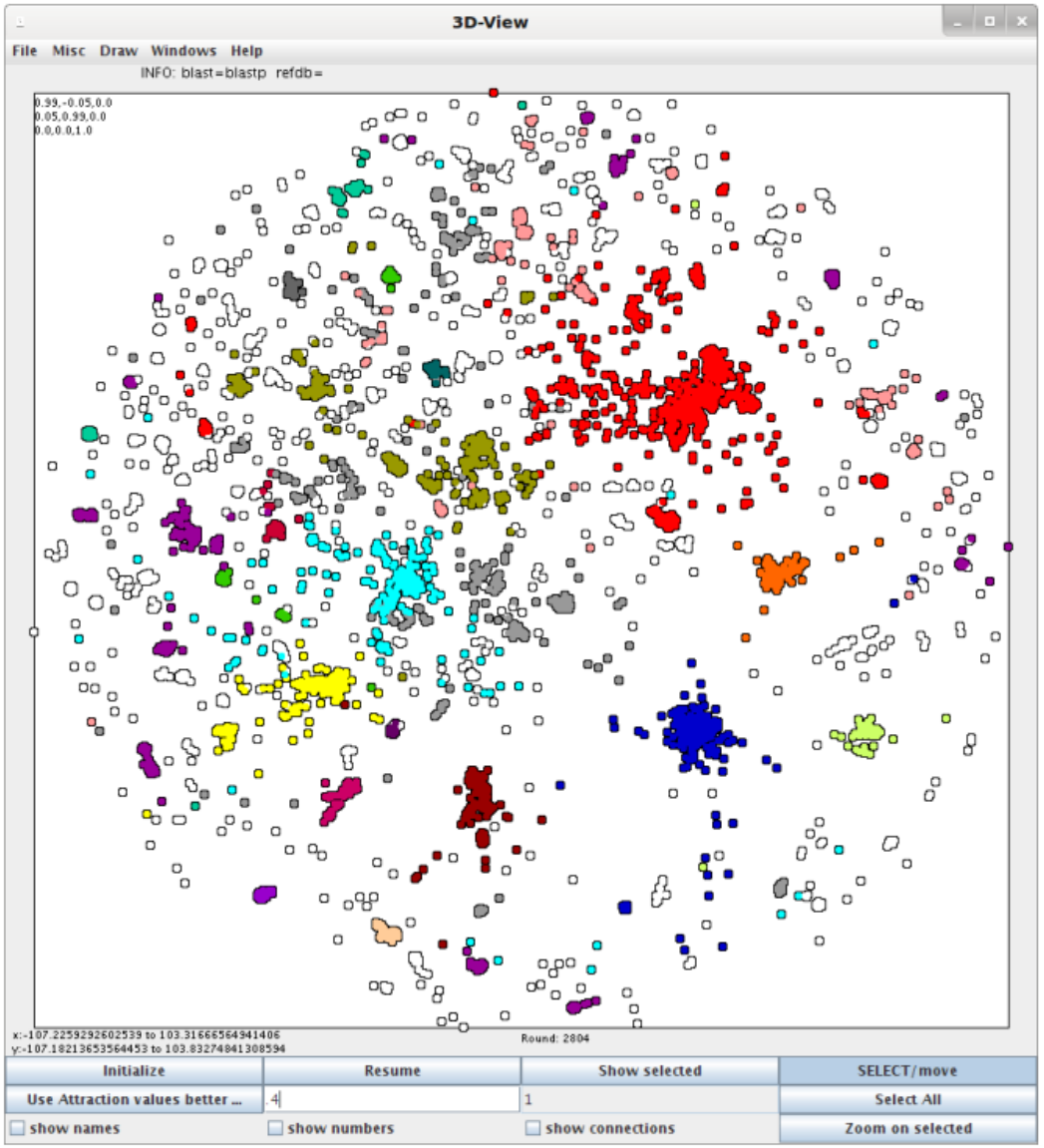


Visualizzazione MINMWMAX 6 -tuplo

Beschreibung Applikation Starten (Ein Projekt von Slawomir Sobota)

[Universität Salzburg, Fachbereich Romanistik, Erzabt-Klotz-Str. 1, A-5020 Salzburg, E-Mail: webmaster(at)dialectometry(dot)com

Language similarity visualisation – Gerhard Jäger, Tübingen



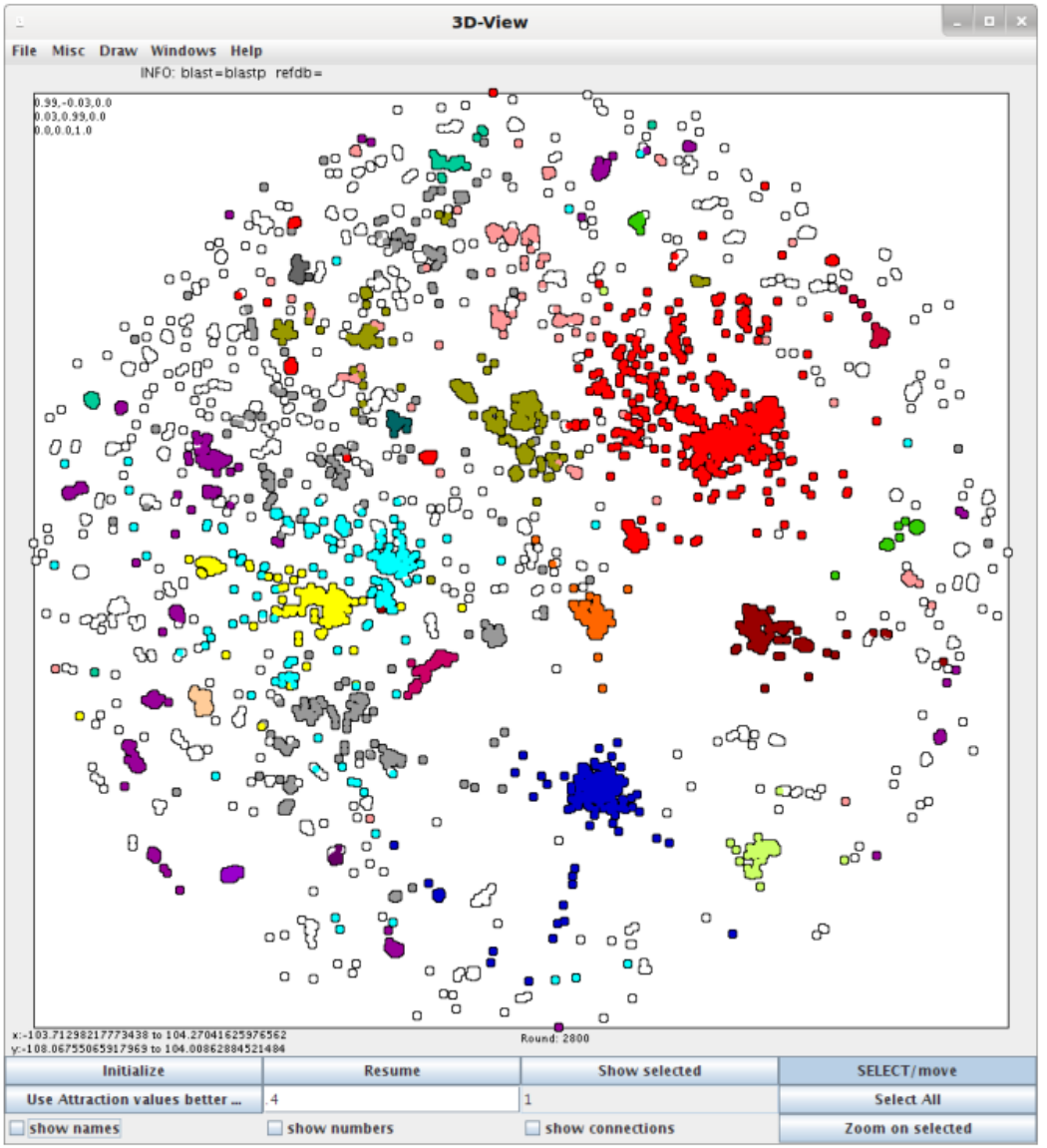


Figure 3: Languages of the world

Project: comparison of Kru languages

Hypothesis:

The geographical distances between the Kru languages are reflected in the differences of their consonant systems.

Method:

Data mining with legacy language atlases

Results:

Coming up

Here:

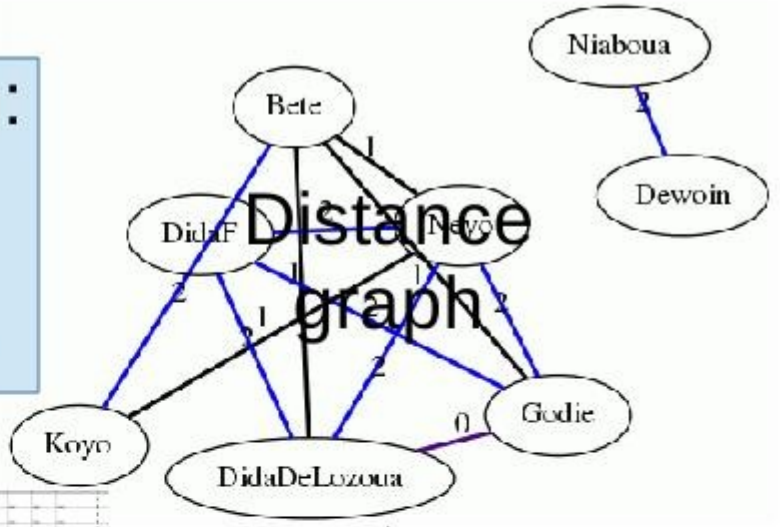
Visualisation of relations induced from the legacy data for 19 Kru languages

(of the 39 in the Ethnologue database of language metadata)

Côte d'Ivoire: Kru languages – consonants

Visualisation workflow:

Preprocessing:
1. scan/type
2. correct
3. create CSV



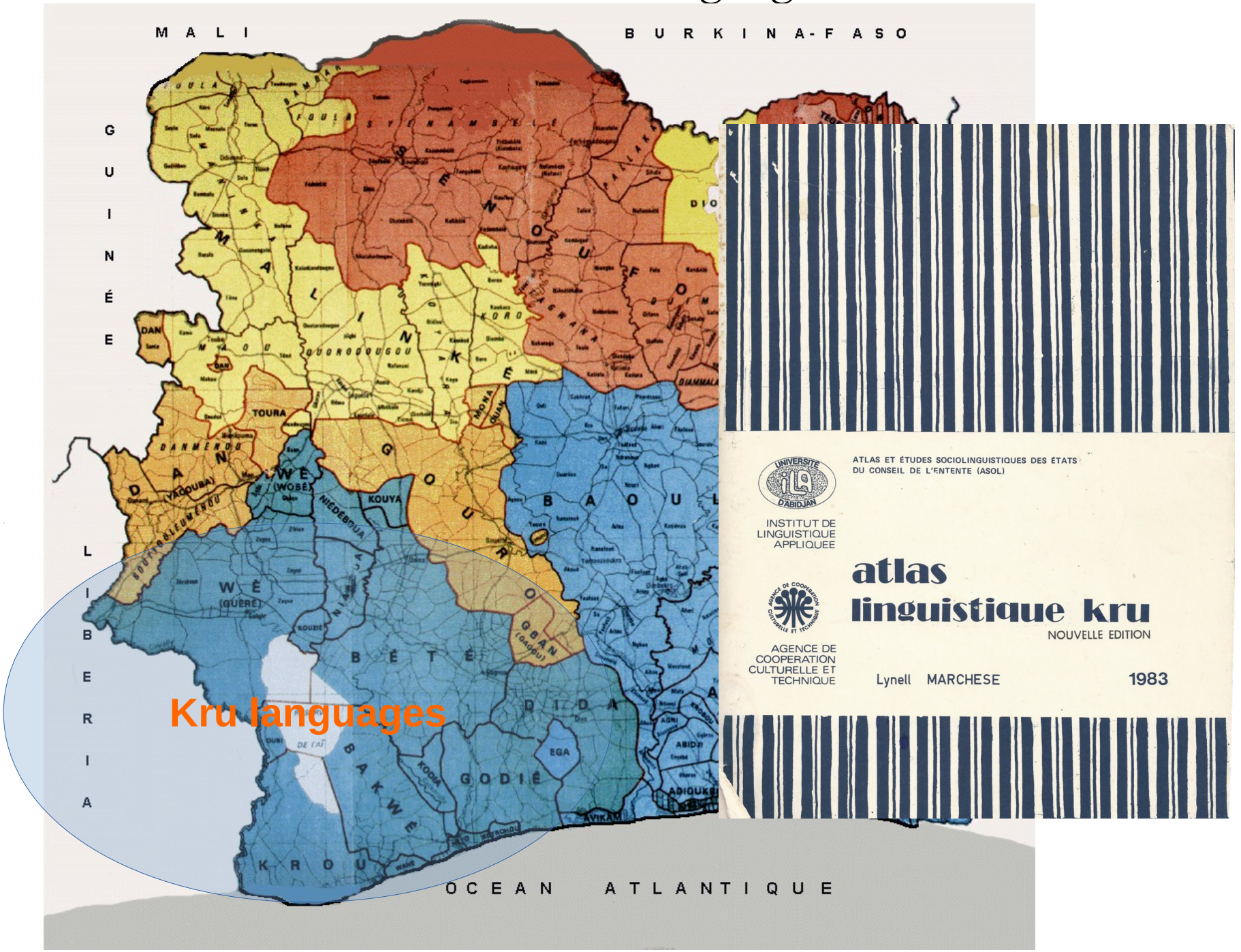
Data input: CSV

GraphViz

Levenshtein Edit Distance comparison

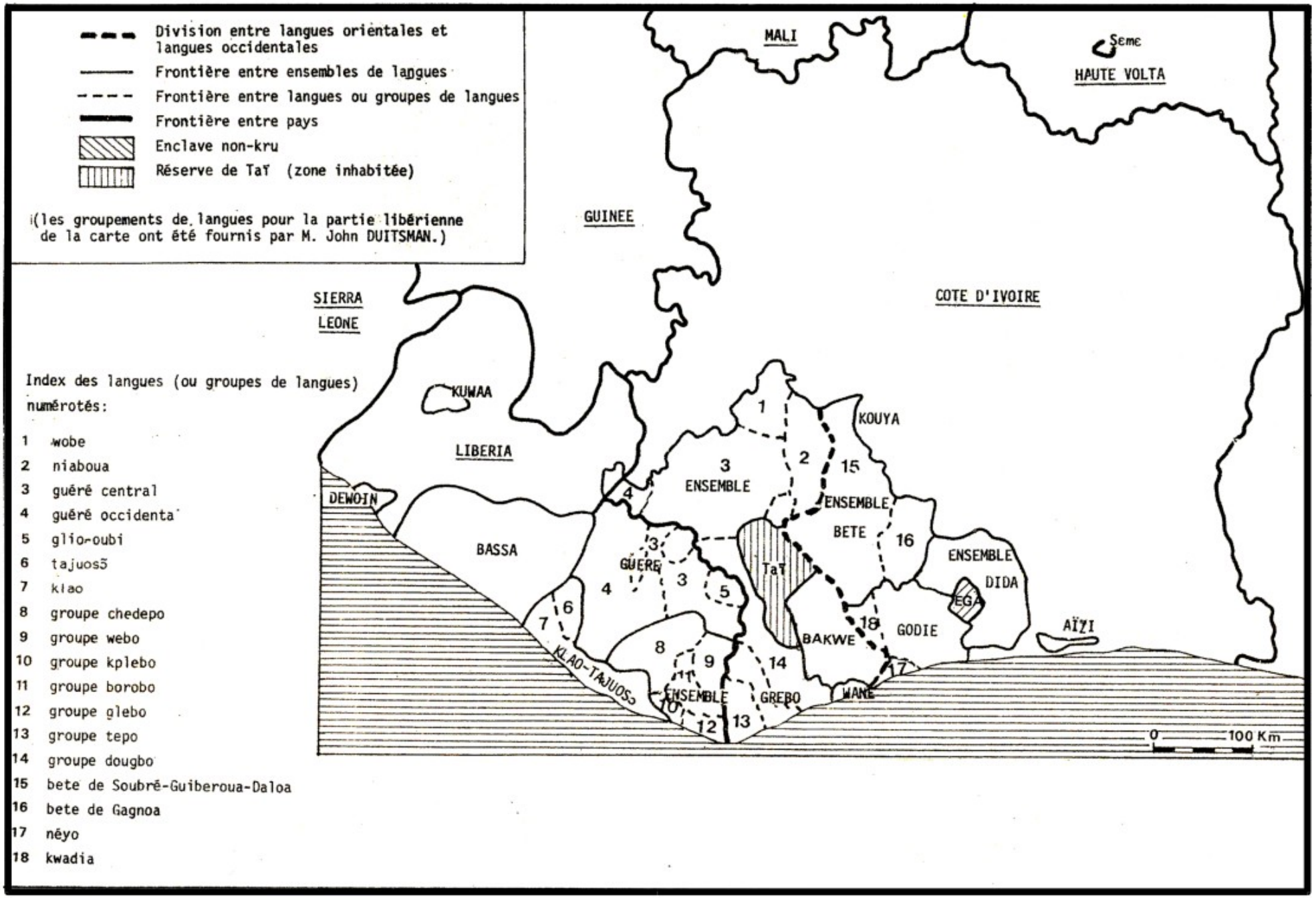
Distance matrix

Côte d'Ivoire: Kru languages



Côte d'Ivoire: Kru languages

Carte I : Les langues kru



Côte d'Ivoire: Kru languages – consonants

- 107 -

- SYSTEMES CONSONANTIQUES DE QUELQUES LANGUES ORIENTALES -

bété de Guibéroua (Werle, 1976)

p	t	c	k	kp	C^{w}(1)
b	d	ɟ	g	gb	
f	s				
v	z				
ɓ	l	j	ɣ	w	
m	n	ɲ	ŋ	ŋw	

Godié de Dakpadou et Lagako (Marchese, 1975)

p	t	c	k	kp	kw
b	d	ɟ	g	gb	gw
f	s				
v	z				
ɓ	l	j	ɣ		w
m	n	ɲ	ŋ		ŋw

Koyo (Kokora, 1976, p. 23)

p	t	c	k	kp	C^{w}	C^{j}
b	d	ɟ	g	gb		
f	s					
v	z					
ɓ	l	j	ɣ(2)		w	
m	n	ɲ	ŋ			

Néyo (Grah)

p	t	c	k	kp	C^{w}
b	d	ɟ	g	gb	
f	s				
v	z				
ɓ	l	j	ɣ		w
m	n	ɲ	ŋ		

Dida de Lozoua (Gratrix)

p	t	c	k	kp	kw
b	d	ɟ	g	gb	gw
f	s				
v	z				
ɓ	l	j	ɣ	w	
m	n	ɲ	ŋ	ŋw	

dida-f (vata) (Siméon, Dugas, Kaye, Koopman, 1981)

p	t	c	k	kp	kw	
b	d	ɟ	g	gb	gw	
f	s					
v	z					
m	n	ɲ	ŋ	ŋm(3)		
ɓ	l	j	ɣ	w		

(1) Voir section 2.1.4.4.

(2) En Koyo, [ɣ] n'apparaît que dans quelques lexèmes dont la plupart sont des emprunts.

(3) "Le ŋm résulte d'une assimilation nasale devant les consonnes labio-vélaires" (Siméon, et. al. 1980:107).

Côte d'Ivoire: Kru languages – consonants

- SYSTEMES CONSONANTIQUES DE QUELQUES LANGUES KRU OCCIDENTALES -

Wobé[1] (Link, 1975 p. 206)

p	t	c	k	kp	kw	
b	d	ɟ		gb		
f	s					
						w
m	n	ɲ		ŋm	km	ŋw

Guéré[2] (Fisher, 1976 p. 96)

p	t	c	k	kp	kw	
b	d	ɟ	g	gb	gw	
f	s					
v	z					
ɓ	l	j				w
	ɗ					
m	n	ɲ		ŋm	km	ŋw

Krahn (Duitsman)

p	t	c	k	kw	
b	d	ɟ		gb	
f	s				
	l			w	
m	n	ɲ			

cedepo (Laesch, c.p.)[4]

p	t	c	k	kp	kw	
b	d	ɟ		gb		
f	s		h			
	l					
m	n	ɲ		ŋm		

Klao (Duitsman, et.al, 1975, p. 92)

p	t	c	k	kp	kw	
b	d	ɟ		gb		
f	s					
	l	j		w		
m	n	ɲ	ŋm			

Niaboua (Bentinck, 1975 p.8)

p	t	c	k	kp	kw	
b	d	ɟ	g	gb	gw	
f	s					
v	z					
ɓ	l	j				w
m	n	ɲ				

Dewoin (Welmers)

p	t		k	kp	kw	
b	d	ɟ	g	gb	gw	
f	s					
v	z					
ɓ	l	j				w
m	n	n	ŋ			

Bassa[3] (Bertkau, et al.

p	t	c		k	kp	
b	d	ɟ	dj	g	gb	
f	s					
v	z					
ɓ	l				w	
m	n	ɲ			gw	
					hw	h
					hw	

Grebo (Innes, p. 14

p	t	c	k	kp			
b	d	ɟ	g	gb			
f	s						
	l	j			w		
m	n	ɲ	ŋ	ŋm		nw	
hm	hn					hw	h
	hl						

Tépo (Dawson, MS)

p	t	c	k	kw		
b	ɗ	ɟ	g	gb		
f	s					h
	l	j		w		
m	n	ɲ	ŋ	ŋm		

(1) La série de nasales en wobé, guéré, tépo et bassa n'est pas phonémique (voir section 2.2.1.2.)

(2) D'après Fisher, il y a une opposition entre l *et* ɗ *en guéré :* ɟɗí *singes / ɟlì' gallons*

(3) Il semble y avoir une opposition entre ɟ *et* dj *en bassa. Cette opposition n'a pas été relevée dans d'autres langues kru.*

(4) Nous ne savons pas pourquoi les semi-voyelles y *et* w *ne figurent pas sur les tableaux de cedepo et de bassa.*

Côte d'Ivoire: Kru languages – consonants

- 111 -

- SYSTEMES CONSONANTIQUES DES LANGUES KRU ISOLEES -

(LIBERIA)	(HAUTE-VOLTA)	(COTE-D'IVOIRE)
kuwaa (Thompson, p. 12)	Sɛmɛ (Prost, p. 346)	Aĩzi (Hérault p. 10)
p t k kp kw	p t c k kp	p t c k kp
b d ɟ	b d ɟ g gb	b d ɟ g gb
f s	f s (ă)	f s š
	v	v z ž
l j ɣ w	l l j w	l j w
m n ɲ ŋ	m n ɲ gm	m n ɲ ŋ
mb nd nɟ ŋg ŋmgb	(h semble être un allophone de f)	

Côte d'Ivoire: Kru languages – consonants

A practical systematisation procedure for a machine-readable database:
Step 1: A word processor or spreadsheet or DBMS <u>table</u>.
Step 2: Export as <u>CSV</u> (character/comma/tab separated value) table.
Step 3: Process <u>manually</u> or <u>automatically</u>: analyse and format as desired.

Bete	p t c k kp kw _ b d C _ g gb _ f s _ v z _ _ _ B _ l j x w m n J N Nw _ _ _ _ _ _ _ _ _ _ _
Godie	p t c k kp kw _ b d C _ g gb gw f s _ v z _ _ _ B _ l j x w m n J N Nw _ _ _ _ _ _ _ _ _ _ _
Koyo	p t c k kp kw kj b d C _ g gb _ f s _ v z _ _ _ B _ l j x w m n J N _ _ _ _ _ _ _ _ _ _ _ _
Neyo	p t c k kp kw _ b d C _ g gb _ f s _ v z _ _ _ B _ l j x w m n J N _ _ _ _ _ _ _ _ _ _ _ _
DidaDeLozoua	p t c k kp kw _ b d C _ g gb gw f s _ v z _ _ _ B _ l j x w m n J N Nw _ _ _ _ _ _ _ _ _ _ _
DidaF	p t c k kp kw _ b d C _ g gb gw f s _ v z _ _ _ B _ l j x w m n J N _ Nm _ _ _ _ _ _ _ _ _ _
Wobe	p t c k kp kw _ b d C _ _ gb _ f s _ _ _ _ _ _ _ _ _ _ _ w m n J _ Nw Nm km _ _ _ _ _ _ _ _ _
Guere	p t c k kp kw _ b d C _ g gb gw f s _ v z _ _ _ B D l j _ w m n J _ Nw Nm km _ _ _ _ _ _ _ _ _
Krahn	p t c k _ kw _ b d C _ _ gb _ f s _ _ _ _ _ _ _ _ _ l _ _ w m n J _ _ _ _ _ _ _ _ _ _ _ _ _
Cedepo	p t c k kp kw _ b d C _ _ gb _ f s _ _ _ _ h _ _ _ l _ _ _ m n J _ _ Nm _ _ _ _ _ _ _ _ _ _
Klao	p t c k kp kw _ b d C _ _ gb _ f s _ _ _ _ _ _ _ _ _ l j _ w m n J _ _ Nm _ _ _ _ _ _ _ _ _ _
Niaboua	p t c k kp kw _ b d C _ g gb gw f s _ v z _ _ _ B _ l j _ w m n J _ _ _ _ _ _ _ _ _ _ _ _ _
Dewoin	p t _ k kp kw _ b d C _ g gb gw f s _ v z _ _ _ B _ l j _ w m n J N _ _ _ _ _ _ _ _ _ _ _ _
Bassa	p t c k kp _ _ b d C dj g gb _ f s _ v z _ h hw B _ l _ _ w m n J _ Nw _ _ _ _ _ _ _ _ _ _ _
Grebo	p t c k kp _ _ b d C _ g gb _ f s _ _ _ _ h hw _ _ l j _ w m n J N Nw Nm _ _ hm hn hl _ _ _ _ _
Tepo	p t c k _ kw _ b d C _ g gb _ f s _ _ _ _ h _ _ _ l j _ w m n J N _ Nm _ _ _ _ _ _ _ _ _ _
KuwaaLiberia	p t _ k kp kw _ b d C _ _ _ _ f s _ _ _ _ _ _ _ _ _ l j x w m n J N _ _ _ _ _ _ _ mb nd nC Ng Nmgb
SemeHauteVolta	p t c k kp _ _ b d C _ g gb _ f s S v _ _ h _ _ _ l j _ w m n J _ _ _ _ gm _ _ _ _ _ _ _ _
AiziCdI	p t c k kp _ _ b d C _ g gb _ f s S v z Z _ _ _ _ l j _ w m n J N _ _ _ _ _ _ _ _ _ _ _ _

Côte d'Ivoire: Kru languages – consonants

A practical systematisation procedure for a machine-readable database:
Step 1: A word processor or spreadsheet or DBMS <u>table</u>.
Step 2: Export as <u>CSV</u> (character/comma/tab separated value) table.
Step 3: Process <u>manually</u> or <u>automatically</u>: analyse and format as desired.

Bete p t c k kp kw _ b d C _ g gb _ f s _ v z _ _ _ B _ l j x w m n J N Nw _ _ _ _ _ _ _ _ _ _ _
Godie p t c k kp kw _ b d C _ g gb gw f s _ v z _ _ _ B _ l j x w m n J N Nw _ _ _ _ _ _ _ _ _ _ _
Koyo p t c k kp kw kj b d C _ g gb _ f s _ v l j x w m n J N _ _ _ _ _ _ _ _ _ _ _ _
Neyo p t c k kp kw _ b d C _ g gb _ f s _ v _ _ _ _ _ _ _ _ _ _
DidaDeLozoua p t c k kp kw _ b d C _ g gb gw f s _ _ _ _ _
DidaF p t c k kp kw _ b d C _ g gb gw f s _
Wobe p t c k kp kw _ b d C _ _ gb _ f s _
Guere p t c k kp kw _ b d C _ g gb gw f s _
Krahn p t c k _ kw _ b d C _ _ gb _ f s
Cedepo p t c k kp kw _ b d C _ _ gb _ f s
Klao p t c k kp kw _ b d C _ _ gb _ f s
Niaboua p t c k kp kw _ b d C _ g gb gw f
Dewoin p t _ k kp kw _ b d C _ g gb gw f _
Bassa p t c k kp _ _ b d C dj g gb _ _
Grebo p t c k kp _ _ b d C _ g gb _ _
Tepo p t c k _ kw _ b d C _ g gb _ f s _ _ _ _ h _ _ _ _
KuwaaLiberia p t _ k kp kw _ b d C _ _ _ _ f s _ _ _ _ _ _ _ _ _ l j x w m n J N _ g Nmgb
SemeHauteVolta p t c k kp _ _ b d C _ g gb _ f s S v _ _ h _ _ _ l j _ w m n J _ _ _ _ gm _ _ _ _ _
AiziCdI p t c k kp _ _ b d C _ g gb _ f s S v z Z _ _ _ _ l j _ w m n J N _ _ _ _ _ _ _ _ _ _ _ _

Then: 171 language comparisons to do:
$(n^2 - n) / 2 = (19^2 - 19) / 2 = 171$
for 44 features each time: 7524.

That's a helluva lot.
So in comes the software …

How is the comparison done?

Côte d'Ivoire: Kru languages – consonants

A practical systematisation procedure for a machine-readable database:
Step 1: A word processor or spreadsheet or DBMS <u>table</u>.
Step 2: Export as <u>CSV</u> (character/comma/tab separated value) table.
Step 3: Process <u>manually or automatically; analyse and format as desired.</u>

Bete	p t c k kp kw _	_ _ _ _ _ _ _
Godie	p t c k kp kw _	_ _ _ _ _ _ _
Koyo	p t c k kp kw k	_ _ _ _ _ _ _
Neyo	p t c k kp kw _	_ _ _ _ _ _ _
DidaDeLozoua	p t c k kp kw _	_ _ _ _ _ _ _
DidaF	p t c k kp kw _	_ _ _ _ _ _ _
Wobe	p t c k kp kw _	_ _ _ _ _ _ _
Guere	p t c k kp kw _	_ _ _ _ _ _ _
Krahn	p t c k _ kw _	_ _ _ _ _ _ _
Cedepo	p t c k kp kw _	_ _ _ _ _ _ _
Klao	p t c k kp kw _	_ _ _ _ _ _ _
Niaboua	p t c k kp kw _	_ _ _ _ _ _ _
Dewoin	p t _ k kp kw _	_ _ _ _ _ _ _
Bassa	p t c k kp _ _	_ _ _ _ _ _ _
Grebo	p t c k kp _ _	hn hl _ _ _ _ _
Tepo	p t c k _ kw _	_ _ _ _ _ _ _
KuwaaLiberia	p t _ k kp kw _	_ _ mb nd nC Ng Nmgb
SemeHauteVolta	p t c k kp _ _	_ _ _ _ _ _ _
AiziCdI	p t c k kp _ _	_ _ _ _ _ _ _

Probably the most well-known algorithm for pairwise comparison of sequences:

Levenshtein Edit Distance

Definition:

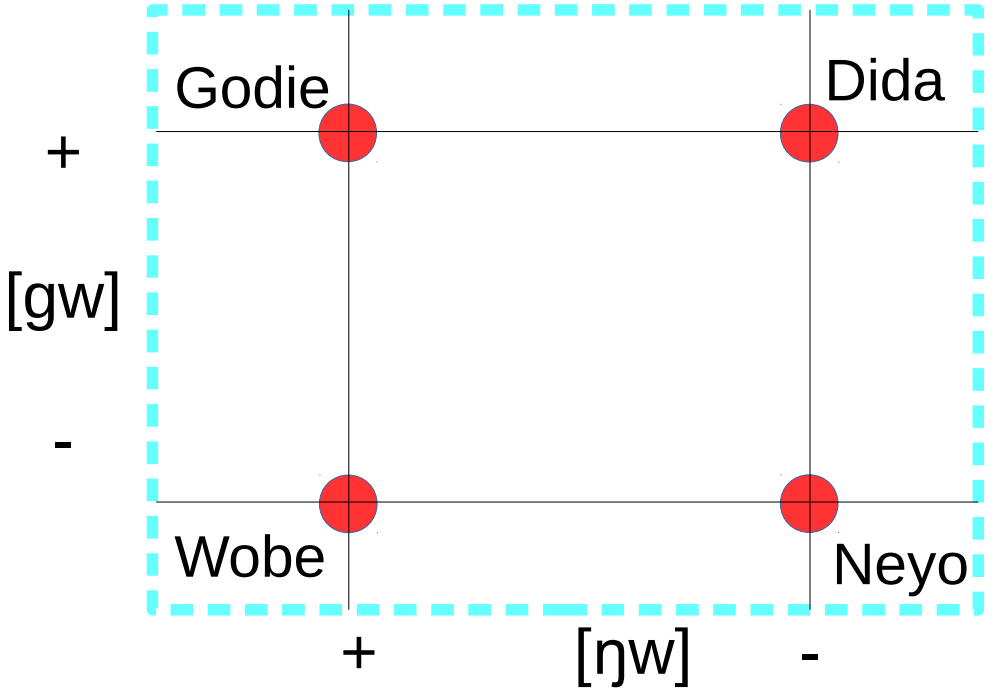
The minimum number of <u>deletions</u>, <u>insertions</u> and <u>substitutions</u> needed to change one sequence into another.

Other distance measures can be used.

In stylometry, numerical distances are used (Comparision of MFW, most frequent words, in texts).

Côte d'Ivoire: Kru languages – consonants

- But: the phoneme data matrix is deceptively 2-dimensional: 19 languages x 44 consonants:
- The 19 objects are actually located in a 44 dimensional quality space. Here are 2 of these dimensions, applied to the 4 languages Godie, Dida, Wobe and Neyo:



- Even distinctive features involve around 12 dimensions.
- How to visualise all 44 dimensions in 2 dimensions?

Côte d'Ivoire: Kru languages – consonants

Bete	0	1	2	1	1	3	10	6	9	11	8	4	4	7	11	8	12	9	6
Godie	1	0	3	2	0	2	11	5	10	12	9	3	3	8	12	9	13	10	7
Koyo	2	3	0	1	3	3	12	8	9	11	8	4	4	9	13	8	12	9	6
Neyo	1	2	1	0	2	2	11	7	8	10	7	3	3	8	12	7	11	8	5
DidaDeLozoua	1	0	3	2	0	2	11	5	10	12	9	3	3	8	12	9	13	10	7
DidaF	3	2	3	2	2	0	11	5	10	10	7	3	3	10	12	7	13	10	7
Wobe	10	11	12	11	11	11	0	8	6	6	4	10	12	12	11	8	14	11	12
Guere	6	5	8	7	5	5	8	0	11	11	8	4	6	9	13	10	18	11	10
Krahn	9	10	9	8	10	10	6	11	0	4	3	7	9	10	12	5	11	8	9
Cedepo	11	12	11	10	12	10	6	11	4	0	3	9	11	10	10	5	13	8	11
Klao	8	9	8	7	9	7	4	8	3	3	0	6	8	11	9	4	10	7	8
Niaboua	4	3	4	3	3	3	10	4	7	9	6	0	2	7	13	8	14	7	6
Dewoin	4	3	4	3	3	3	12	6	9	11	8	2	0	9	13	8	12	9	6
Bassa	7	8	9	8	8	10	12	9	10	10	11	7	9	0	10	11	19	8	9
Grebo	11	12	13	12	12	12	11	13	12	10	9	13	13	10	0	7	17	10	11
Tepo	8	9	8	7	9	7	8	10	5	5	4	8	8	11	7	0	12	7	8
KuwaaLiberia	12	13	12	11	13	13	14	18	11	13	10	14	12	19	17	12	0	15	14
SemeHauteVolta	9	10	9	8	10	10	11	11	8	8	7	7	9	8	10	7	15	0	5
AiziCdI	6	7	6	5	7	7	12	10	9	11	8	6	6	9	11	8	14	5	0

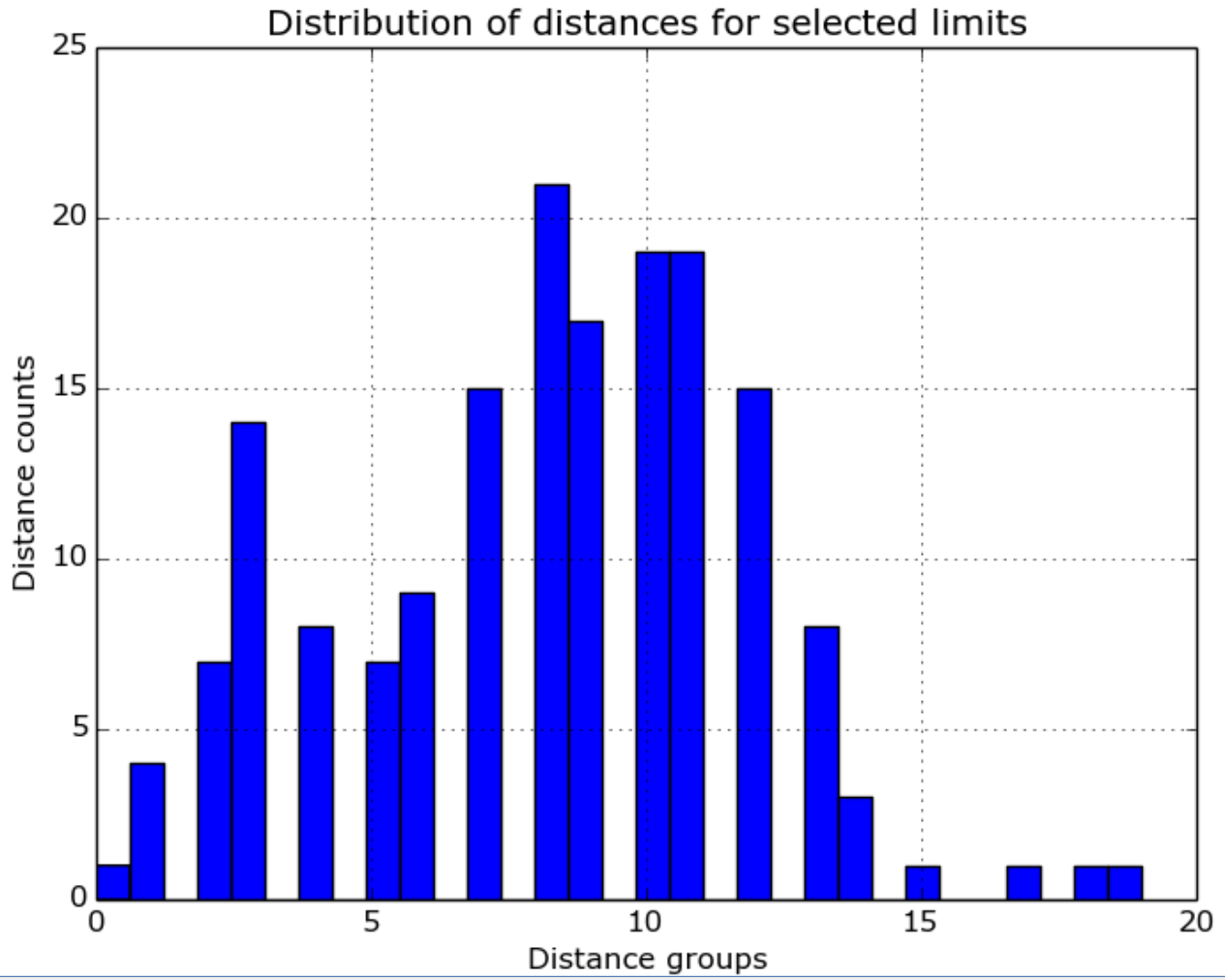
Côte d'Ivoire: Manding languages – phonological rules

TABLEAU N° 2

CALCUL DE DISTANCE ENTRE LES 23 PARLERS MANDING de COTE D'IVOIRE

N°S des parlers	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	•	10	7	8	6	3	4	6	6	4	6	3	4	6	7	8	5	7	8	7	8	6	4
2	10	•	8	11	9	4	5	7	7	4	6	4	5	5	3	6	3	4	4	3	3	5	4
3	7	8	•	11	13	6	5	4	4	6	3	6	6	9	3	6	5	6	2	5	6	5	3
4	8	11	11	•	12	3	4	5	5	5	4	5	4	6	2	5	4	5	3	2	3	4	3
5	6	9	13	12	•	5	4	3	3	4	2	5	5	8	2	5	4	5	1	4	4	4	2
6	3	4	6	3	5	•	13	11	11	9	8	12	9	7	4	6	3	4	4	5	5	8	12
7	4	5	5	4	4	13	•	12	12	9	9	11	9	6	3	5	5	5	5	3	4	8	13
8	6	7	4	5	3	11	12	•	14	9	10	11	9	4	3	5	3	3	5	2	2	7	11
9	6	7	4	5	3	11	12	14	•	9	10	11	9	4	3	5	3	3	5	2	2	7	11
10	4	4	6	5	4	9	9	9	9	•	10	12	9	7	5	6	6	5	4	4	4	7	10
11	6	6	3	4	2	8	9	10	10	10	•	9	7	3	3	4	4	3	6	2	2	6	11
12	3	4	6	5	5	12	11	11	11	12	9	•	11	7	4	6	4	4	4	3	3	7	10
13	4	5	6	4	5	9	9	9	9	9	7	11	•	7	2	6	5	5	5	3	4	8	8
14	6	5	9	6	8	7	6	4	4	7	3	7	7	•	7	9	8	9	6	9	10	5	5
15	7	3	3	2	2	4	3	3	3	5	3	4	2	7	•	11	7	8	10	12	11	7	3
16	8	6	6	5	5	6	5	5	5	6	4	6	6	9	11	•	7	7	9	11	10	7	4
17	5	3	5	4	4	3	5	3	3	6	4	4	5	8	7	7	•	9	10	10	10	7	6
18	7	4	6	5	5	4	5	3	3	5	3	4	5	9	8	7	9	•	8	10	9	5	5
19	8	4	2	3	1	4	5	5	5	4	6	4	5	6	10	9	10	8	•	10	10	7	5
20	7	3	5	2	4	5	3	2	2	4	2	3	3	9	12	11	10	10	10	•	13	7	4
21	8	3	6	3	4	5	4	2	2	4	2	3	4	10	11	10	10	9	10	13	•	7	4
22	6	5	5	4	4	8	8	7	7	7	6	7	8	5	7	7	7	5	7	7	7	•	7
23	4	4	3	3	2	12	13	11	11	10	11	10	8	5	3	4	6	5	5	4	4	7	•

Côte d'Ivoire: Kru languages – consonants



Spread of differences between 19 Kru consonant inventories for 44 features, which we want to visualise.

Useful strategy: interpret and map differences as distances in quality space.

Strategy #1:

Squash to 2 dimensions!

- Differences are interpreted as distances
- Distances are represented spatially as a distance map
- The dimensions are squashed – like a system of springs – into 2 dimensions
- Further dimensions may be represented by colours, etc.

• Strategy #2:

Select elite features!

- Check the features for their importance in distinguishing objects
- Randomly start with an important feature and build a hierarchy of features distinguishing between sets of objects until all are distinguished
- Different choices lead to different results, different insights

Strategy #1:

Squash to 2 dimensions!

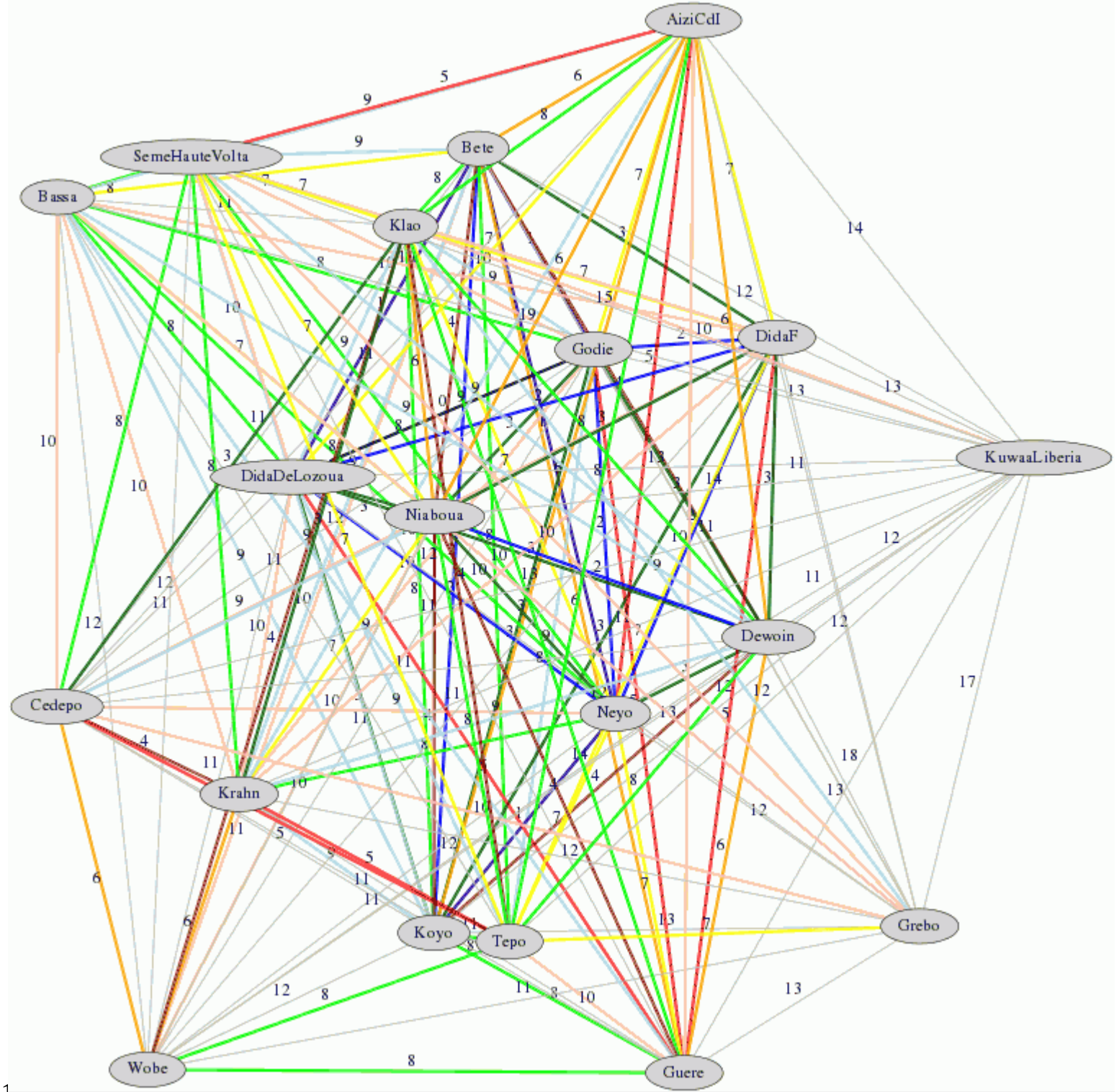
- Differences are interpreted as distances
- Distances are represented spatially as a distance map
- The dimensions are squashed – like a system of springs – into 2 dimensions
- Further dimensions may be represented by colours, etc.

And visualise the results!

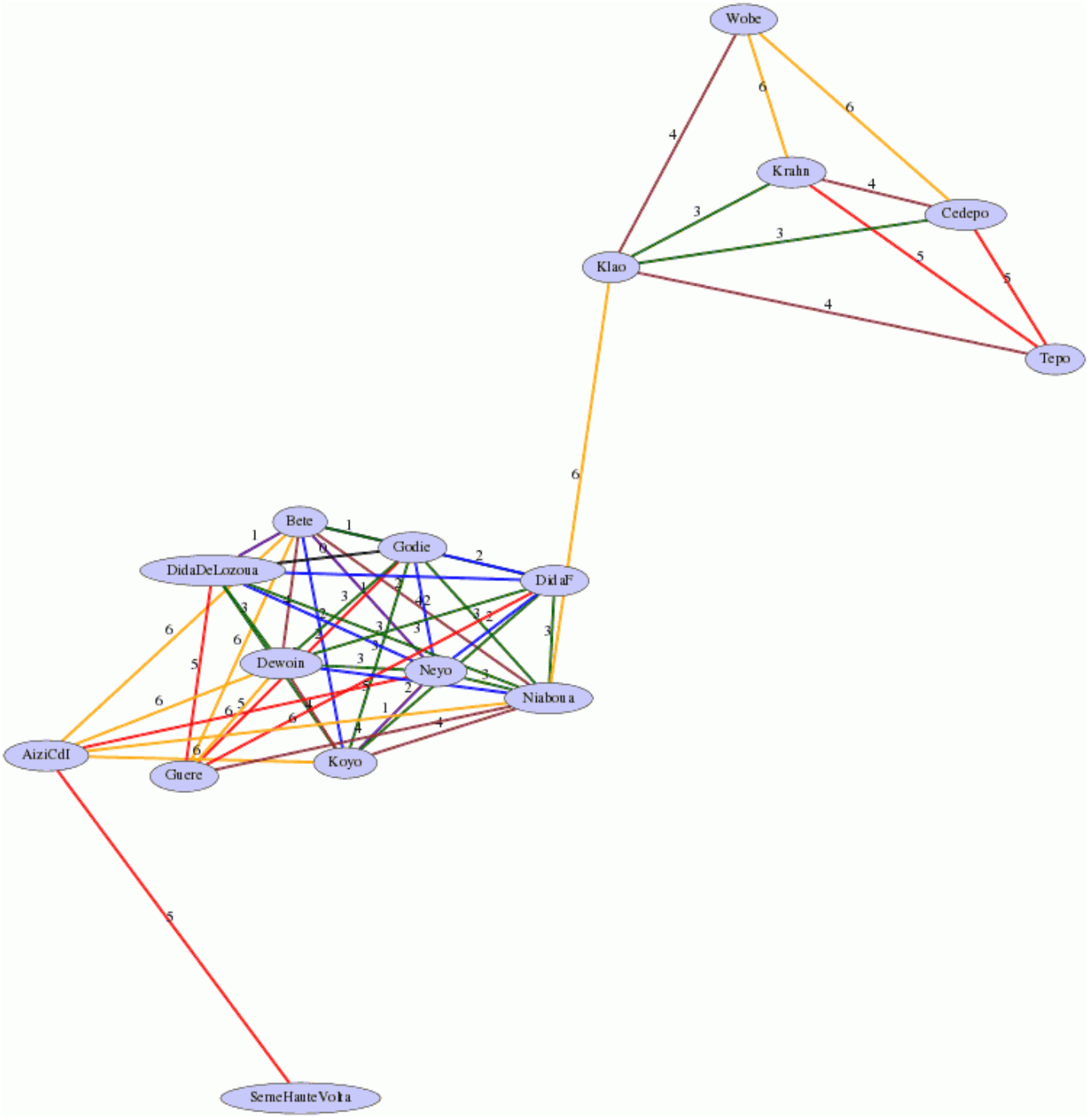
- Strategy #2:

Select elite features!

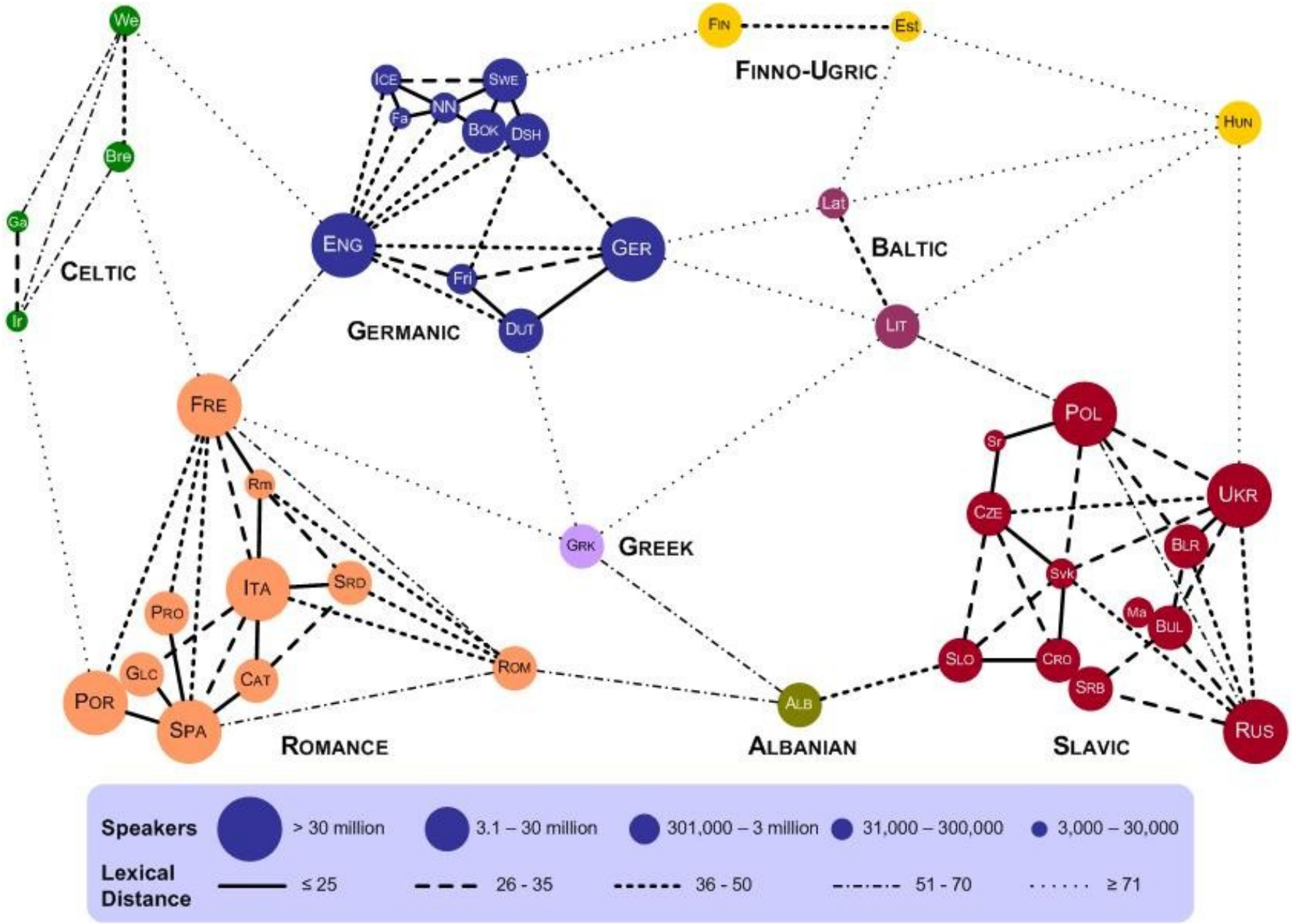
- Check the features for their importance in distinguishing objects Randomly start with an important feature and build a hierarchy of features distinguishing between sets of objects until all are distinguished
- Different choices lead to different results, different insights


AiziCdI
SemeHauteVolta
Bete
Bassa
Klao
DidaF
Godie
KuwaaLiberia
DidaDeLozoua
Niaboua
Dewoin
Cedepo
Neyo
Krahn
Grebo
Koyo
Tepo
Wobe
Guere

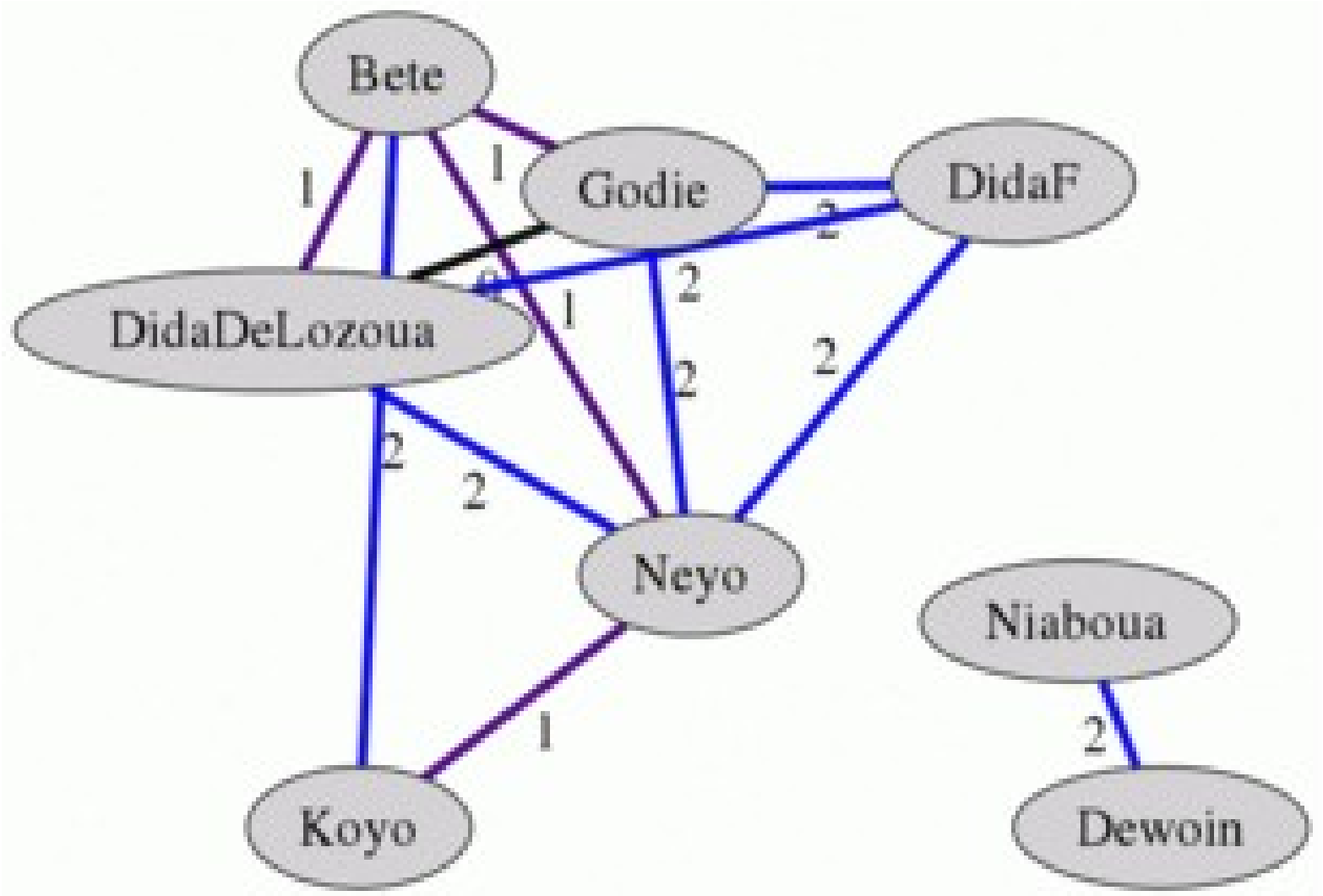
Strategy 1: Squash those dimensions!



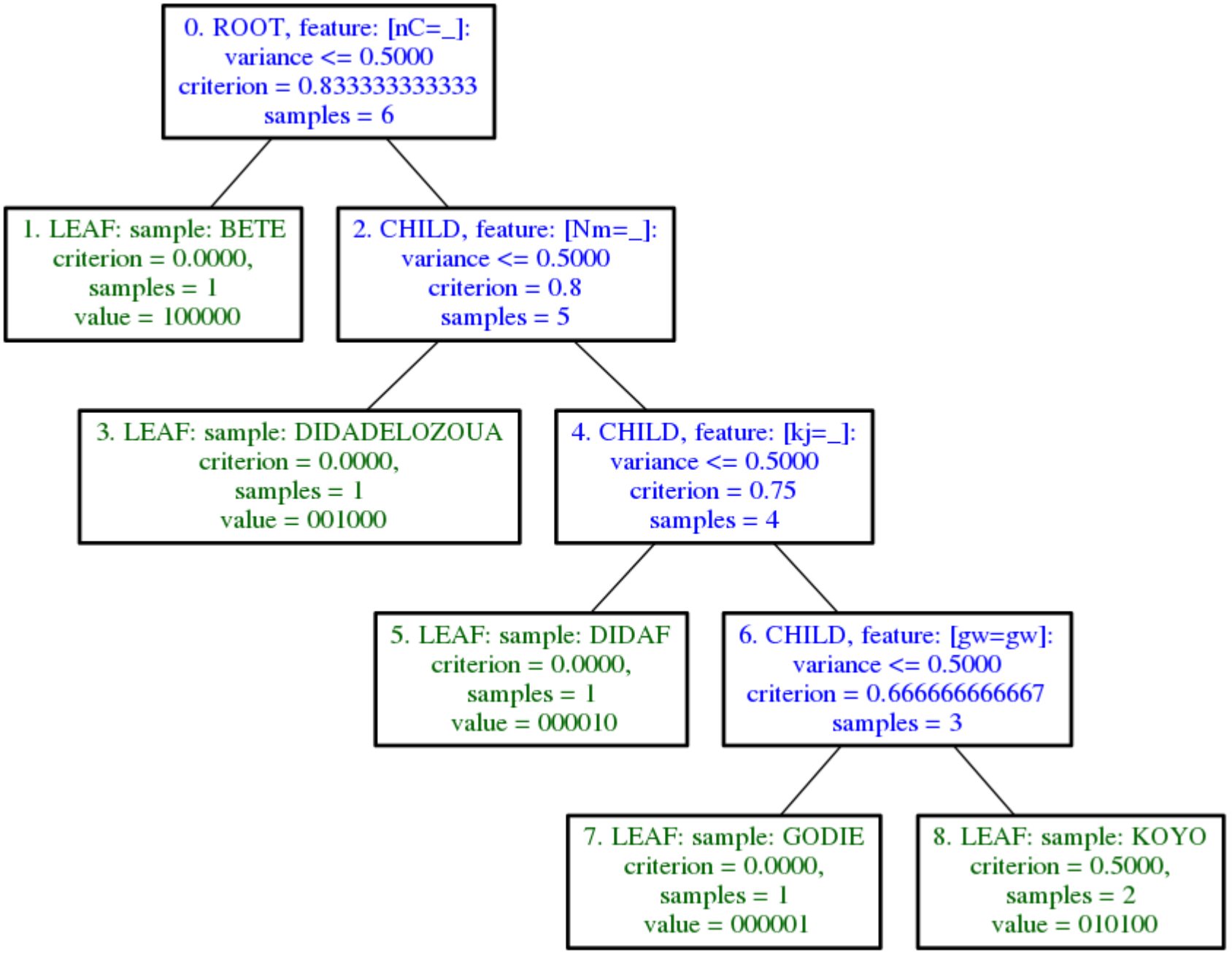
Strategy 1: Squash those dimensions!



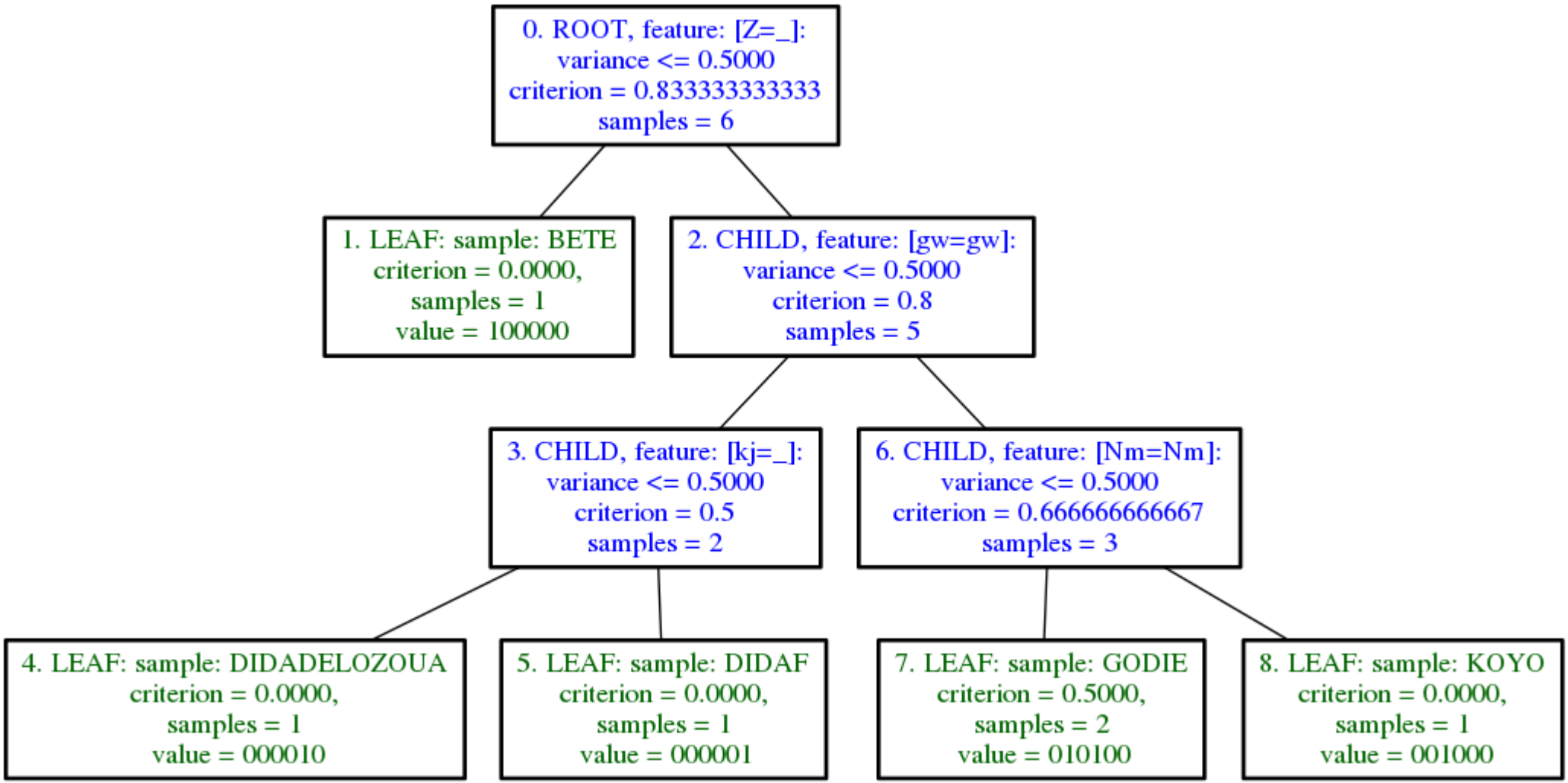
Strategy 1: Squash those dimensions!



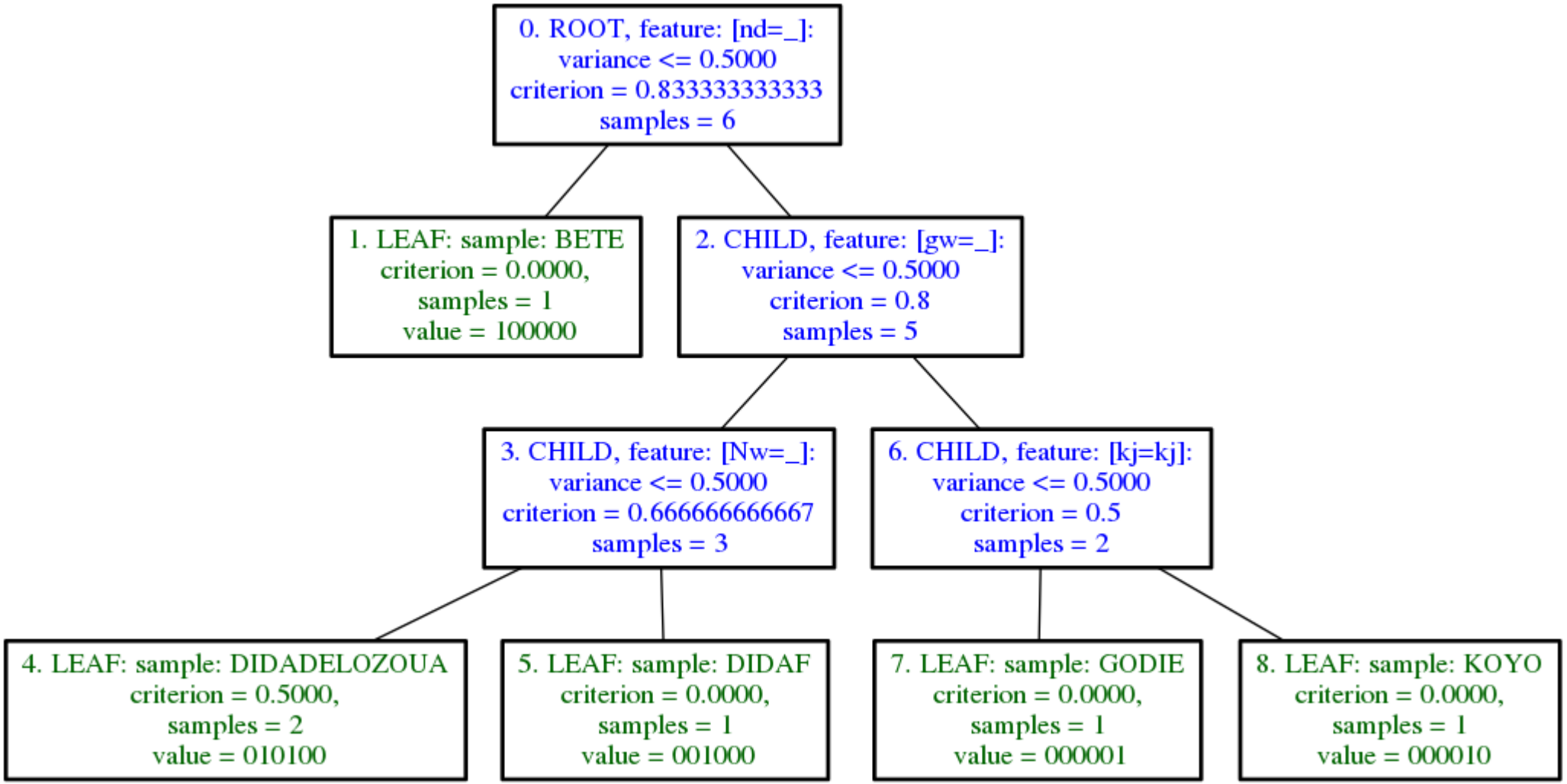
Strategy 2: Pick out the best features!



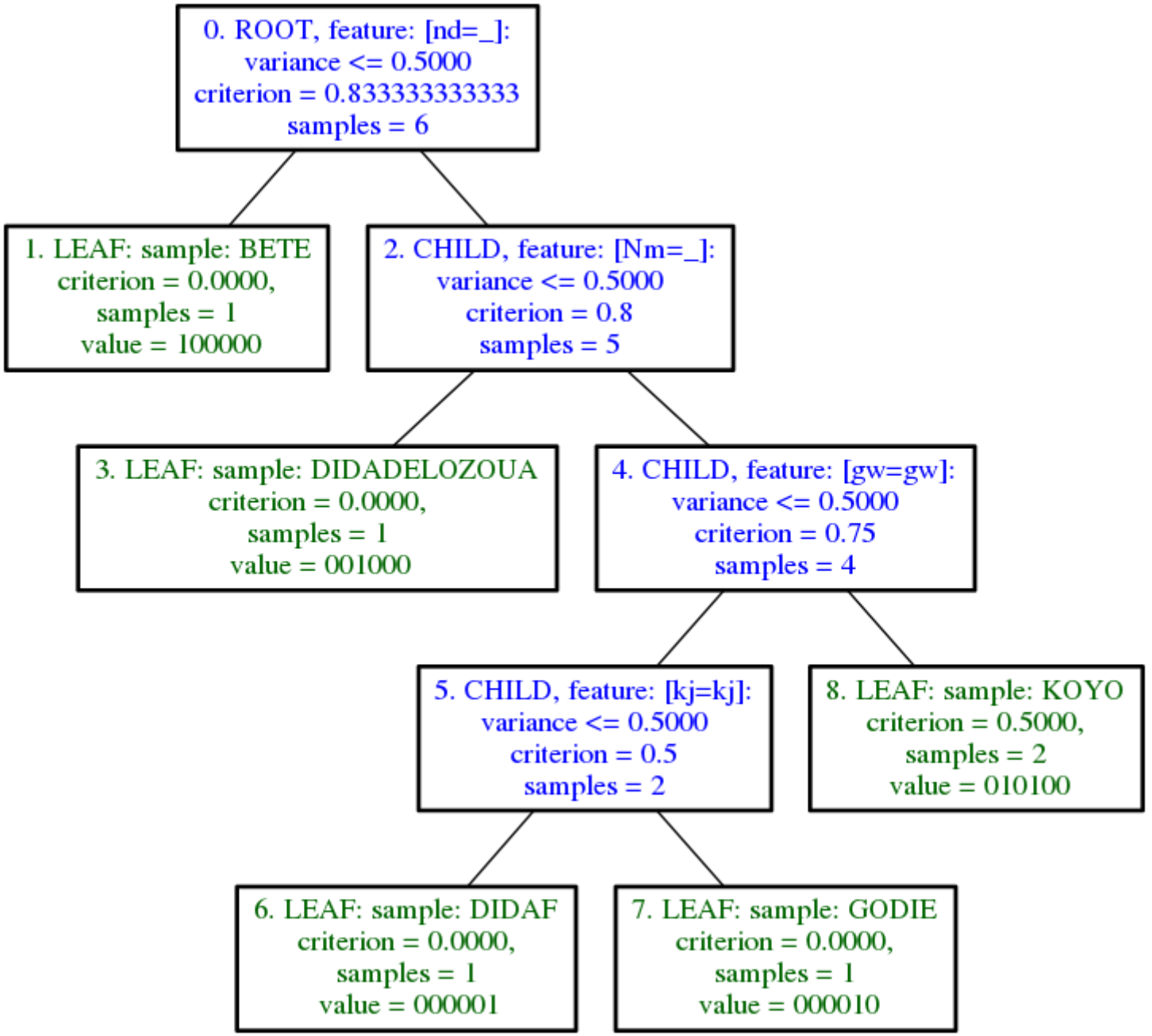
Strategy 2: Pick out the best features!



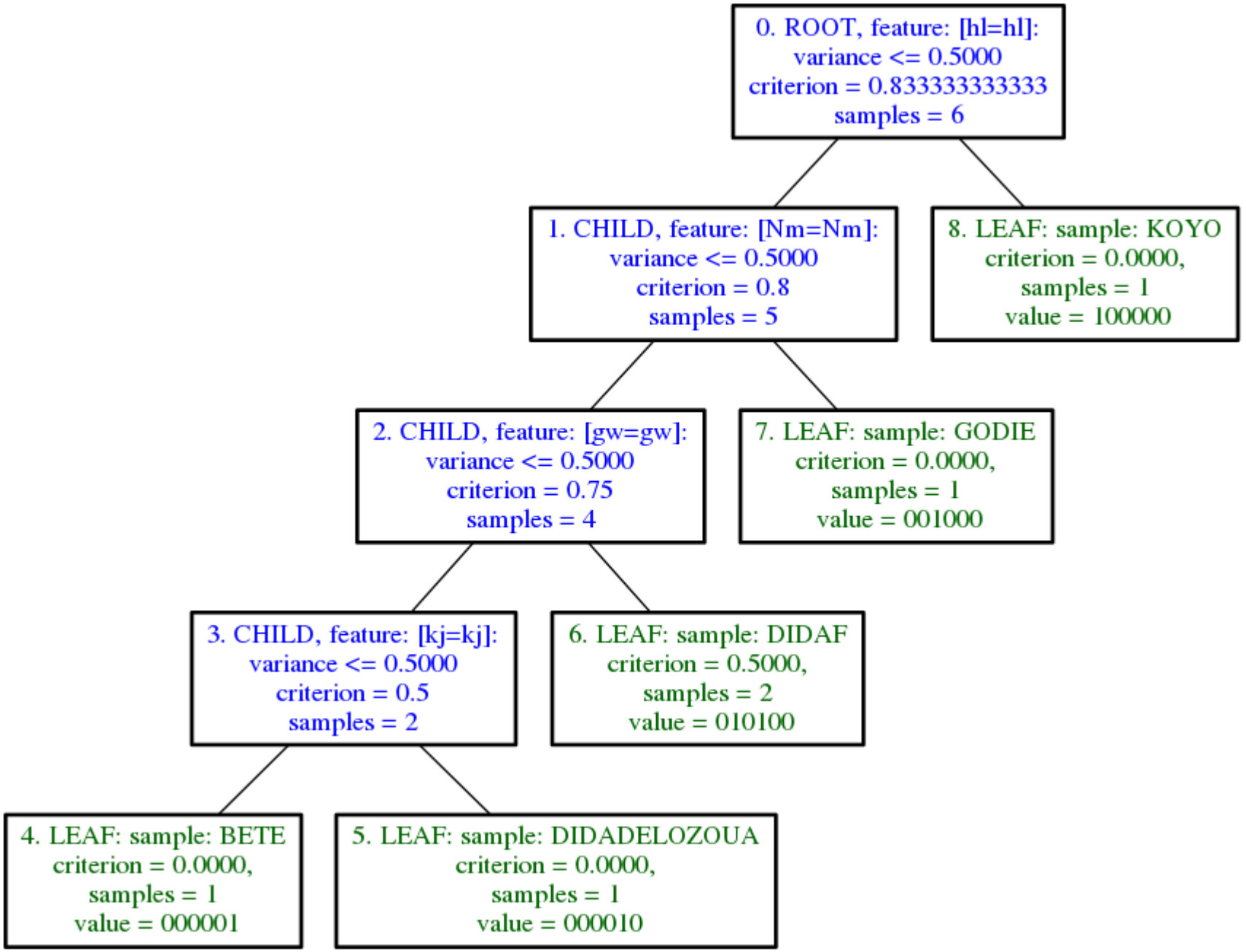
Strategy 2: Pick out the best features!



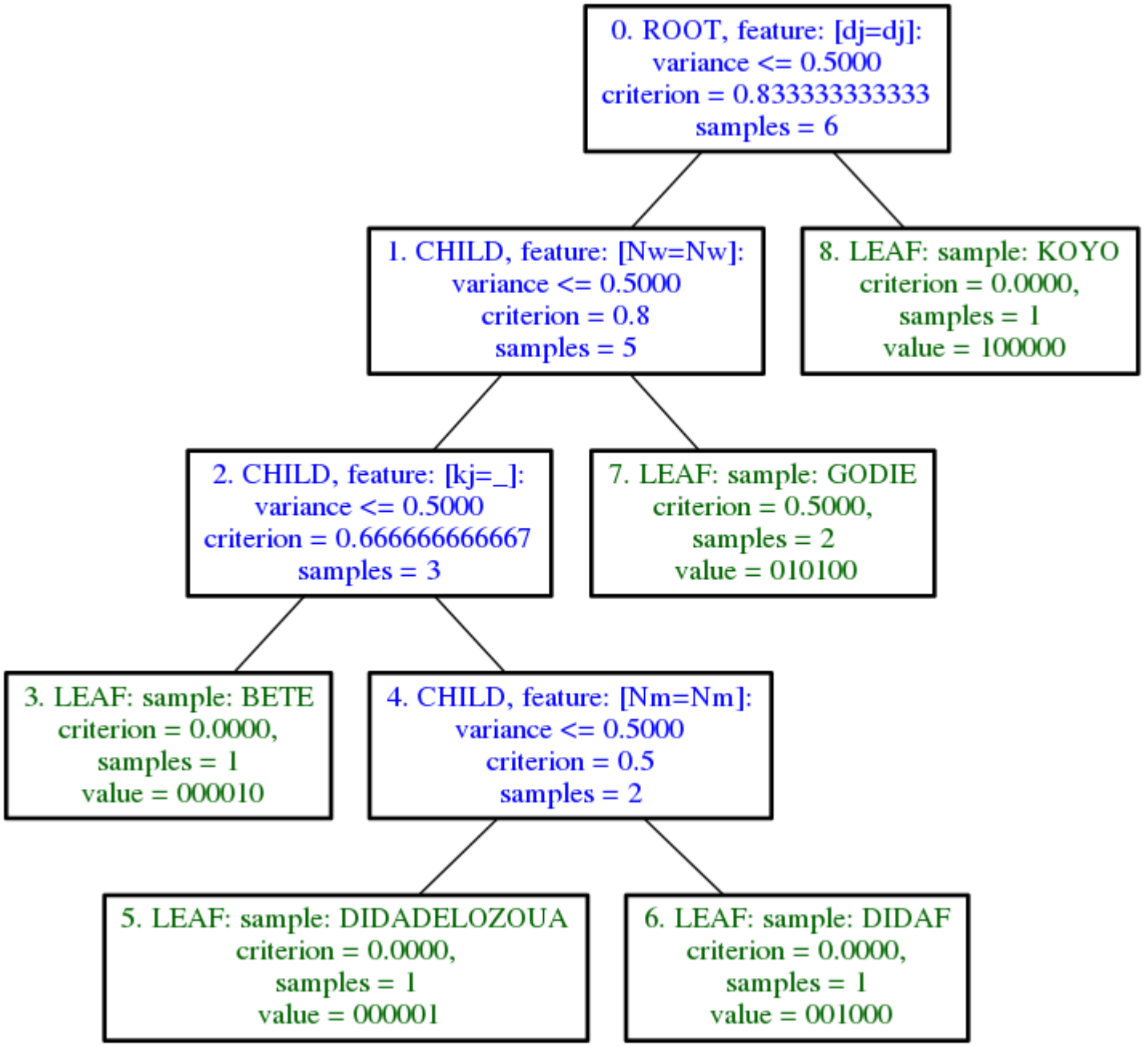
Strategy 2: Pick out the best features!



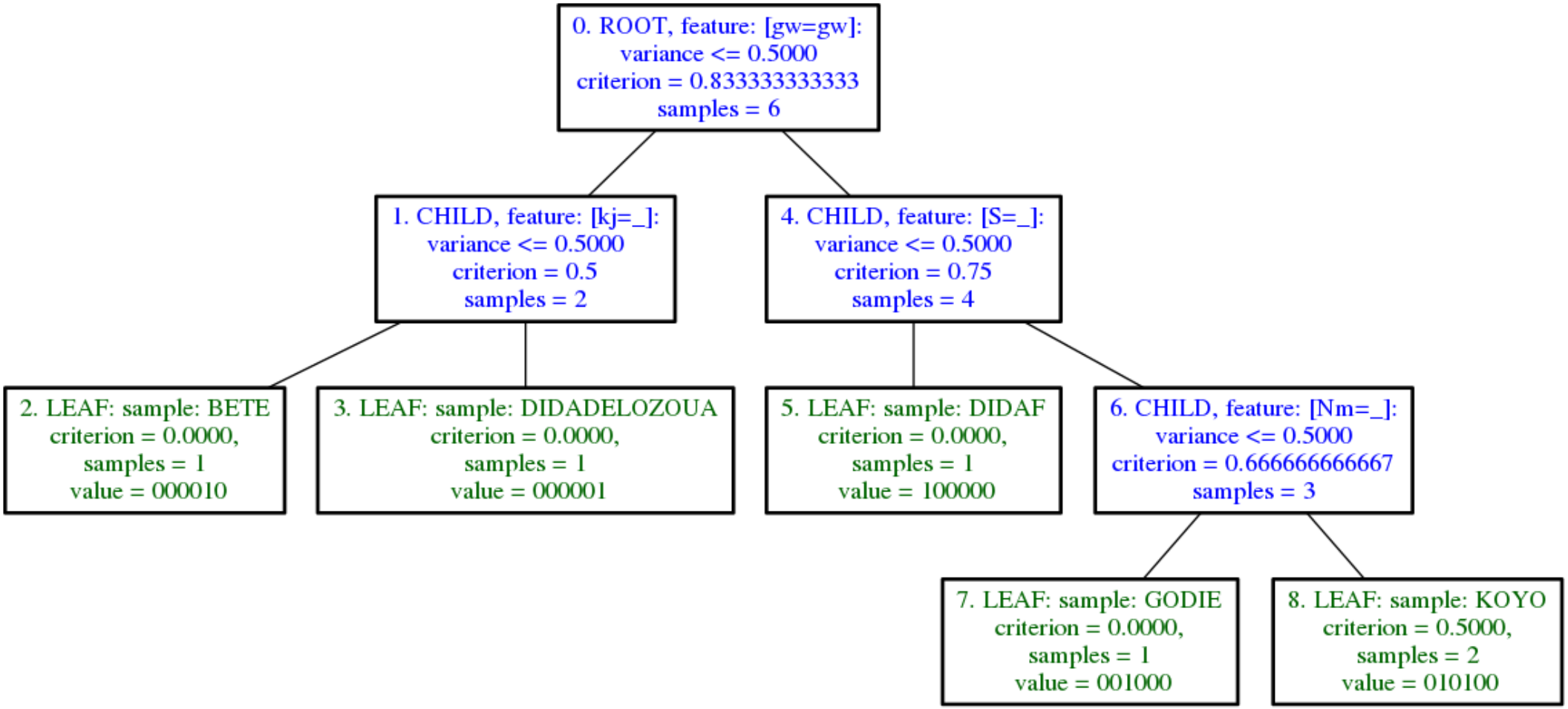
Strategy 2: Pick out the best features!



Strategy 2: Pick out the best features!



Strategy 2: Pick out the best features!



The moral of this story is that

Legacy data in linguistic atlases can be given a new lease of life and a solid quantitative foundation in addition to any further research on dialect relations and history which may be pursued.

Standard arrangements of quantitative information (e.g. tables) may be useful.

Graphical visualisations are helpful in either suggesting or underlining lines of investigation.

Demo: DistGraph

Many thanks – and looking forward

to further discussions

and future cooperation!