African Language classification

Autumn 2009

Subject matter: classification

• Why language classification?
  – Historical interest: how are languages genetically related?
  – Classifications taking language contact into account: areal classifications
  – Typological interest: building meaningful categories of languages on the basis of shared properties
  – Referential classifications: identifying languages
Referential classifications

- Malcolm Guthrie (1948, 1967-71)
  - Bantu languages
  - Zones A-S, groups 10, 20, 30…
  - Individual languages:
    - G42 Swahili
    - R31 Herero
    - S32 Northern Sotho
- Methods...
- Cf. updated source:
  http://goto.glocalnet.net/mahopapers/nug/online.pdf

Typological classifications 1/2

- Interest in the distribution of certain grammatical/phonetic-phonological/syntactic features:
  - Grammatical gender
  - Passive strategies
  - Word order of main constituents (Subject, Verb, Object) in declarative main clauses
  - Distribution of specific sounds/sound classes
  - …
Typological classifications 2/2

• “Shared features of languages of one type (=from one typological class) may have arisen completely independently.”
  Source: http://en.wikipedia.org/wiki/Natural_language#Typological_classification

• What is possible in languages? Which features/properties tend to co-occur?
  – SOV word order and postpositions
  – VSO word order and prepositions
Co-occurrence point at universal laws governing the structure of language.

Areal classification 1/2

• Languages of a given geographical area sometimes share a considerable amount of linguistic features, despite the fact that they belong to different (sub)families.
  – Balkan Sprachbund; Kaukasus; East Asia; etc.
• Reason: Speaker contact, multilingualism
• Caution: Geographical names for language groupings…
Areal classification 2/2

• “[R]esearch on contact-induced linguistic relationship is still in its infancy. What makes areal language classification particularly difficult are problems such as the following: (a) there are no reasonable findings to guide the student of areal linguistics as to how many features would be required to define an areal group, or how to determine its boundaries; from the little we know, boundaries of areal groups are notoriously fuzzy; (b) there are also no ready-made methods and models to classify languages according to contact-induced relationship.”

Source: (Heine??) in Dixon Aikhenvald, p 409

Genetic classifications

• Examples: Indo-European, Semitic, Finno-Ugric
• Genetic classification = family trees
• Key motivations
  – Language history
  – Cross-linguistic evidence for difficult phenomena
• Method
  – Find similarities in different languages!
  – Mass comparison of lexical items
  – Historical-comparative reconstruction
Comparison

• How do these classifications relate to each other?
  – Genetic classification primordial
  – Search for cross-linguistic similarities; these are evidence for genetic relationship
  – Areal features and typological universals are important to correct wrong assumptions

Aims and objectives

• Knowledge of the main existing proposals concerning the historical (=genetic) classification of African language families
• Clear understanding of different types of classification
• Familiarisation with major recurrent typological features in African languages
• Techniques: reconstruction, dialectology
Tasks, assignments, test

- Max. absence: 2 sessions
- Familiarise yourself with one of the (putative) major language families in Africa
- Prepare a presentation pointing out major typological features and examples for reconstructed items; assess the validity.
- Test at the end of the course. Monday Oct 19.

Earliest attempts at classifying African languages

- Unity of Bantu languages long recognized
- Hamitic race and Negroid populations
  - Biblical origins: Noah’s sons
  - C.G. Seligman (1930 [sic!])
  - Johann Ludwig Krapf (1810-1881):
    - Nilo-hamitic (Bantu) vs.
    - Nigro-hamitisch (W African)
Earliest history of classification

Since 10th cent.: Mention of African languages in arabic documents. Relationship between Hebrew, Arabic and Aramaic known to Jewish/Arabic scholars
1538 G. Postel first European to recognize this relationship.
17th cent. First scholarly work on African languages in Europe: Coptic (1636), Nubian (1638), (Ki-)Kongo (1652), Nama (1643), Ge’ez (1661) and Amharisch (1698)
1700 H. Ludolf recognized Ge’ez and Amharic as related to Hebrew, Arabic, Aramaic
1776 L.B. Proyart recognized genetic relations among Bantu
1778 W. Marsden more comprehensive recognition of the Bantu family and the degree of their relationship (published in 1816)

1781 von Schlözer introduces the term ‘Semitic’
1808 H. Lichtenstein distinguishes between Bantu and Nama languages in southern Africa
1820s Champollion: similarities Old Egyptian and Semitic
1826 A. Balbi early overview in Atlas ethnographique du globe ou classification des peuples anciens et modernes d’après leurs langues
1850 J.L. Krapf invents the terms ‘Hamitische Sprachen’ (Hamitic languages) for the non-semitic African languages
From Krapf to Lepsius and Meinhof

- Friedrich Müller named the traditional Hamito-Semitic family in 1876 in his Grundriss der Sprachwissenschaft, and defined it as consisting of a Semitic group plus a "Hamitic" group containing Egyptian, Berber, and Cushitic; he excluded the Chadic group.
- Egyptologist Karl Richard Lepsius (1810–1884)
  - Hamitic = non-Semitic languages in Africa with gender system:
    - Ancient Egyptian, Berber, Cushitic (Beja), Chadic (Hausa), Nama
    
    Source: Lepsius, Richard: Nubische Grammatik. Mit einer Einleitung über die Völker und Sprachen Afrika's, Berlin, 1880

Carl Meinhof (1857-1944)

- 1912 Die Sprachen der Hamiten (The Languages of the Hamites)
  - Grammatical gender
  - Cattle-herding speakers
  - Brighter skin complexion

- Meinhof expanded Lepsius' model, adding Fula (Peul), Maasai, Bari, Nandi, Sandawe, and Hadza.
- Nilo-Hamitic debate
Criticism

Diedrich Westermann 1875-1956:
• Relationship between Bantu languages and languages in W African Sudan belt.
• Does not challenge the relationship with E African Sudan languages

Klingenheben, 1930s:
Ful (Fulfulde, Peul) shown to be related to W-Sudansprachen, thus not Hamitic

Historical linguistics

- Initial observation: systematic similarities
  - Requirement: similar in sound and meaning
  - Explanation: inherited from common ancestor

<table>
<thead>
<tr>
<th></th>
<th>Swahili</th>
<th>Herero</th>
<th>Zulu</th>
<th>Lucazi</th>
</tr>
</thead>
<tbody>
<tr>
<td>'person'</td>
<td>mtu</td>
<td>omundu</td>
<td>omuntu</td>
<td>muntu</td>
</tr>
<tr>
<td>'people'</td>
<td>watu</td>
<td>abandu</td>
<td>abantu</td>
<td>bantu</td>
</tr>
<tr>
<td>'guinea-fowl'</td>
<td>kanga</td>
<td>onganga</td>
<td>impangele</td>
<td>(li)kanga</td>
</tr>
<tr>
<td>'healer'</td>
<td>mganga</td>
<td>onganga</td>
<td>igqirha</td>
<td>cimbanda</td>
</tr>
</tbody>
</table>

...Bantu

- Swahili G 42
- Lucazi K 13
- Herero R 31
- Zulu S 42

http://goto.glocalnet.net/mahopapers/nuglonline.pdf
Historical linguistics

• Initial observation: systematic similarities
  – Requirement: similar in sound and meaning
  – Explanation: inherited from common ancestor

<table>
<thead>
<tr>
<th></th>
<th>Swahili</th>
<th>Herero</th>
<th>Zulu</th>
<th>Lucazi</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘person’</td>
<td>mtu</td>
<td>omundu</td>
<td>omuntu</td>
<td>muntu</td>
</tr>
<tr>
<td>‘people’</td>
<td>watu</td>
<td>abandu</td>
<td>abantu</td>
<td>bantu</td>
</tr>
<tr>
<td>‘guinea-fowl’</td>
<td>kanga</td>
<td>onganga</td>
<td>impangele</td>
<td>(li)kanga</td>
</tr>
<tr>
<td>‘healer’</td>
<td>mganga</td>
<td>onganga</td>
<td>igqirha</td>
<td>cimbanda</td>
</tr>
</tbody>
</table>

Ancestor languages

• Correspondences in form and meaning

<table>
<thead>
<tr>
<th>Sprache</th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kongo, Ganda, Rwanda, Zulu</td>
<td>m-ntu</td>
<td>ba-ntu</td>
</tr>
<tr>
<td>Herero</td>
<td>m-ndu</td>
<td>va-ndu</td>
</tr>
<tr>
<td>Lingala</td>
<td>m-o-to</td>
<td>ba-to</td>
</tr>
<tr>
<td>Sotho</td>
<td>m-d-ho</td>
<td>b-d-ho</td>
</tr>
<tr>
<td>Swahili</td>
<td>m-tu</td>
<td>wa-tu</td>
</tr>
</tbody>
</table>

• Methods
  – Mass comparison of lexical items
  – Historical-comparative reconstruction
Exercise

• Go through the following list and establish the family tree

Ancestor languages

• Correspondences in form and meaning

<table>
<thead>
<tr>
<th>Sprache</th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kongo, Ganda, Rwanda, Zulu</td>
<td>m-ntu</td>
<td>ba-ntu</td>
</tr>
<tr>
<td>Herero</td>
<td>m-ndu</td>
<td>va-ndu</td>
</tr>
<tr>
<td>Lingala</td>
<td>mo-to</td>
<td>ba-to</td>
</tr>
<tr>
<td>Sotho</td>
<td>m-bi-thó</td>
<td>bô-thó</td>
</tr>
<tr>
<td>Swahili</td>
<td>m-tu</td>
<td>wa-tu</td>
</tr>
</tbody>
</table>

• Methods
  – Mass comparison of lexical items
  – Historical-comparative reconstruction

• Shortcomings and challenges
  – Where to start?
  – Bifurcation
  – Plausibility of change
Likely direction of change

- Bantu 7V & 5V systems
- Prenasals

What is the *proto-root?

- *-thu 'person'
- *-tu 'person'
- *-tô 'person'
- *-ntô 'person'
- *-ntu 'person'
- *-nthu 'person'
- *-ndu 'person'
- *-du 'person'

<table>
<thead>
<tr>
<th>Sprache</th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kongo, Ganda, Rwanda, Zulu</td>
<td>mu-ntu</td>
<td>ba-ntu</td>
</tr>
<tr>
<td>Herero</td>
<td>mu-ndu</td>
<td>va-ndu</td>
</tr>
<tr>
<td>Lingala</td>
<td>mo-to</td>
<td>ba-to</td>
</tr>
<tr>
<td>Sotho</td>
<td>mō-thō</td>
<td>bā-thō</td>
</tr>
<tr>
<td>Swahili</td>
<td>m-tu</td>
<td>wa-tu</td>
</tr>
</tbody>
</table>

Next exercise: homework

- Try to establish as many regular correspondences, and suggest a likely proto-phoneme.
- Describe problems that you encounter and reflect possible reasons.

Handout 1: Hand-written list of random Bantu languages
Handout 2: Copies from
Kavango languages

- Oribi  
  *Ourebia ourebi*
- Impala  
  *Aepyceros melampus*
- Pangolin  
  *Manis temminckii, Manis gigantea*

<table>
<thead>
<tr>
<th></th>
<th>Kwangali</th>
<th>Geiriku</th>
<th>Mbulushu</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘bull’</td>
<td>ntwedu (9/10)</td>
<td>ntwedu (9/10)</td>
<td>ntwedho (9/10)</td>
</tr>
<tr>
<td>‘calf’</td>
<td>sitana (7/8)</td>
<td>kantana (12/13)</td>
<td>ndana (9/10)</td>
</tr>
<tr>
<td>‘impala’</td>
<td>mpara (9/10)</td>
<td>mpara (9/10)</td>
<td>mara (9/10)</td>
</tr>
<tr>
<td>‘new-born calf’</td>
<td>sikombe (7/8)</td>
<td>nkembe (9/10)</td>
<td>dingombe (5/6)</td>
</tr>
<tr>
<td>‘orbi’</td>
<td>ntimg (9/10)</td>
<td>ntimg (9/10)</td>
<td>nting (9/10)</td>
</tr>
<tr>
<td>‘pangolin’</td>
<td>nkaka (9/10)</td>
<td>nkaka (9/10)</td>
<td>naka (9/10)</td>
</tr>
<tr>
<td>‘udder’</td>
<td>sawo (9/10)</td>
<td>mashwe (6)</td>
<td>thitango (7/8)</td>
</tr>
</tbody>
</table>
### Kavango languages

<table>
<thead>
<tr>
<th>Semantic field</th>
<th>Kvangah</th>
<th>Gciriku</th>
<th>Mbbokshu</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘Impala’</td>
<td>mpara (9/10)</td>
<td>mpara (9/10)</td>
<td>mpara (9/10)</td>
</tr>
<tr>
<td>‘Pangolin’</td>
<td>nkaka (9/10)</td>
<td>nkaka (9/10)</td>
<td>nkaka (9/10)</td>
</tr>
<tr>
<td>‘Oribi’</td>
<td>ntungu (9/10)</td>
<td>ntungu (9/10)</td>
<td>ntungu (9/10)</td>
</tr>
<tr>
<td>‘Calf’</td>
<td>sitana (7/8)</td>
<td>kantana (12/13)</td>
<td>ndana (9/10)</td>
</tr>
<tr>
<td>‘New-born calf’</td>
<td>sikembe (7/8)</td>
<td>nkembe (9/10)</td>
<td>dingembe (5/6)</td>
</tr>
<tr>
<td>‘Bull’</td>
<td>ntwedu (9/10)</td>
<td>nwedu (9/10)</td>
<td>nwedlu (9/10)</td>
</tr>
<tr>
<td>‘Udder’</td>
<td>sawo (9/10)</td>
<td>mazhwe (6)</td>
<td>bizzanga (7/8)</td>
</tr>
</tbody>
</table>

#### Regularly corresponding sounds
- /nl/, /al/, /el/, ...
- What about /nt/, /mp/, /nk/?

**But:**
- ntwedu~ntwedu~ntwedho
- sitana~kantana~ndana
Things to bear in mind

• How many correspondences mean "regular"?
• How to explain double or multiple series of correspondences?

Reasons for similarities 1/2

• Assumption: Genetically inherited
   But possibly other reasons:
• Borrowing
   – Semantic criteria sometimes helpful
   – Sound historical evidence can help
   – Non-linguistic knowledge of historical situation
Another example

<table>
<thead>
<tr>
<th>Language A</th>
<th>Language B</th>
<th>GLOSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>dus-</td>
<td>dyster-</td>
<td>dark</td>
</tr>
<tr>
<td>hak-</td>
<td>hakk-</td>
<td>cut, split</td>
</tr>
<tr>
<td>d'aan-</td>
<td>deen-</td>
<td>stretch</td>
</tr>
<tr>
<td>kif-</td>
<td>kip-</td>
<td>turn over; pour</td>
</tr>
<tr>
<td>warb-</td>
<td>werf-</td>
<td>throw</td>
</tr>
<tr>
<td>waas-</td>
<td>wets-</td>
<td>sharpen</td>
</tr>
<tr>
<td>zimk-</td>
<td>saug-</td>
<td>suck</td>
</tr>
<tr>
<td>naas-</td>
<td>nass</td>
<td>(be) wet</td>
</tr>
<tr>
<td>rik'-</td>
<td>raq-</td>
<td>seize, grasp</td>
</tr>
<tr>
<td>tarf-</td>
<td>tropf-</td>
<td>drip</td>
</tr>
<tr>
<td>sul(hu)b'-'</td>
<td>shlypf-</td>
<td>ship (v.)</td>
</tr>
<tr>
<td>wander-</td>
<td>wander-</td>
<td>hike</td>
</tr>
<tr>
<td>k'arfi</td>
<td>kraft</td>
<td>force</td>
</tr>
<tr>
<td>gurbii</td>
<td>gruube</td>
<td>ditch</td>
</tr>
</tbody>
</table>

Reasons for similarities 2/2

• Nursery words, onomatopoeia
  – mama
  – croak

• Chance: accidental similarity
  – Hausa and German (previous slide) are almost certainly **not** related to each other
Layerings…

• German Kopf 'head'
• Bantu *-piú 'knife'

Internal reconstruction 1/2

• Words for 'eat' in Bantu:
  -lia ~ -rya ~ -la ~ -dya [~ etc.]
• Note: final /-a/ segmentable!
  -li- ~ -r/- ~ -l- ~ -d/- [~ etc.]
• What is the root?
Internal reconstruction 2/2

- Swahili -refu 'long, tall'

Examples
- wa-tu wa-refu 'tall people'
- m-tu m-refu 'tall person'
- ji-cho Ø-refu 'long eye'
- ndege ndefu 'long aeroplane'

9. bird 9. long (N 'cl. 9'; -dege 'bird')

Relative chronologies

- Remember: 7 vowel systems and 5 vowel systems
- Contrast in both front and back high vowel /i/, /ɪ/, /u/, /ʊ/
- In some Bantu languages, stops spirantised before tense ĭ.
  - *ki > ki, and *kį > cį
- In some Bantu languages, stops spirantised before tense ĭ and cardinal i.
  - *ki, *kį > cį
- In many Bantu languages, the contrast between i and ĭ was lost (reduced from 7 > 5 vowels). As a result the system becomes opaque.

Current views

- Greenberg 1963, 1966

Point of departure

- Against Hamitic-Sudanic-Bantu classification
- Anti-“Hamitentheorie” (& anti-Meinhof)
- Greenberg is a “splitter”, not a “lumper”
- Methodological “innovation”: Greenberg recommends
  - to discard non-linguistic criteria;
  - greatest possible care with typological criteria!
Greenberg’s methods

• Three principles
  – Linguistic evidence only
  – Resemblance in sound and meaning
  – Mass comparison (not pairs of languages)

Significance of mass comparison

• Wide-scale genetic classification feasible?
  – Recurrent resemblances through common origin
  – Probability of chance resemblance minimal
  – Less likelihood to miss meaningful connections

• Are such classifications not arbitrary?
  – Greenberg challenges critics to provide series of etymologies and similar morphemes, but…
Greenberg in a nutshell

• Afroasiatic
• Niger-Kordofanian
  – Niger-Congo
• Nilo-Saharan
  – Chari-Nile
• Khoisan
Niger-Congo

- Westermann (1911, 1927)
  - Bantu, W Sudanic
- Greenberg
  - Eastern Sudanic excluded
  - Bantu’s classificatory relevance minor
  - Sub-divisions with W Sudanic

Greenberg

**Niger-Congo (1500+)**

- West Atlantic (Atlantic)
- Mande
- Voltaic (Gur)
- Kwa
- Benue-Congo
  - ...
  - Bantu (several 100)
- Adamawa-Eastern (Adamawa-Ubangi)
Revised version(s)…

• Ethnologue
Afroasiatic

- Clear relation
  - Berber
  - Old Egyptian
  - Semitic
  - Cushitic
- Greenberg discusses
  - Fulani
  - “Nilo-Hamitic”
  - Hottentot (Nama)
  - Hausa (and similar)
- Undo Hamito-Semitic

Greenberg’s Afroasiatic (374)

- Berber (25)
- Old Egyptian
- Semitic (79)
  - Cushitic
- Cushitic (74)
  - East (33)
  - North (1; Beja)
  - South (7)
  - Central (4)
  - West (=Omotic; 29)
- Chadic (almost 200)
  - Hausa
Afroasiatic (rev. Greenberg)

Khoisan

- Departs from...
  - Bushman languages
  - Hottentot as Hamitic
- D. Bleek
  - Nama, Naro, and others
- Greenberg:
  - South African Khoisan (25)
    - Northern (6)
    - Central (13)
    - Southern (6)
  - Sandawe
  - Hatsa (Hadza)
Chari-Nile (170)

**Chari-Nile**
- Central Sudanic (60)
- Kunama
- Berta
- Eastern Sudanic (100)
  - Nubian
  - …
  - Nilotic (60)

- Rejects Hamitic
  - Cognate morphology of Nilotic
  - Identifying so-called hamitic features as lexical borrowings from Cushitic

- Bottom-up approach
  - Nilotic (W, E, S)
  - With 9 others, forming Eastern Sudanic

Nilo-Saharan (approx. 200)

- Songhai
- Saharan
  - Kanuri, Kanembu, Teda, †Berti, Zaghawa
- Maban
- Fur
- Chari-Nile (as seen)
- Koman
Methods, notions, developments

• Mass comparison, that is clear
• Geographic origin
  Principle of least moves
  – Debate on Bantu
  – Debate on Afroasiatic
  – Debate on Ful

Lexicostatistics

• Rationale:
  The more shared vocabulary, the closer the genetic link between languages
• Sub-groupings characterised by a percentage of shared vocabulary
• Lexicostatistics is an approach to comparative linguistics that involves quantitative comparison of lexical cognates. Lexicostatistics is related to the comparative method but does not reconstruct a proto-language.
Method of lexicostatistics

• Create word list
  – A list of universal culture-free meanings. Words are then collected for these meaning slots for each language being considered. (200 originally, reduced further (but including some meanings that were not in his original list), giving Swadesh’s 100-item list.

• Determine cognacies
  – Cognacy decisions to be made by a trained and experienced linguist; however it should be noted that the decisions may need to be refined as the state of knowledge increases. However, lexicostatistics does not rely on all the decisions being correct. For each pair of lists the cognacy of a form could be positive, negative or indeterminate. Sometimes a language has two words for one meaning, e.g. small and little for not big.

• Calculate lexicostatistic percentages

Method of lexicostatistics

• Percentages relate to the proportion of meanings for a particular language pair that are cognate, i.e. relative to the total without indeterminacy. This value is entered into a N x N table of distances, where N is the number of languages being compared. When complete this table is half-filled in triangular form. The higher the proportion of cognacy the closer the languages are related.

• Create family tree
  – all lists are placed in a pool
  – the two closest members are removed and form a nucleus which is placed in the pool
  – this step is repeated
  – under certain conditions a nucleus becomes a group
  – this is repeated until the pool only contains one group.

• Calculations need to be made of nucleus and group lexical percentages.
Glottochronology

• …attempts to use lexicostatistical methods to estimate the length of time since two or more languages diverged from a common earlier proto-language.

• An application of lexicostatistics that assumes a relatively constant rate of change for basic lexical items.

• Highly controversial (or rather, uncontroversially rejected…)

Raimund Kastenholz 1991/92

• Mande classification

• Significance for us:
  • Deals with the shortcomings/challenges in lexicostatistics: how to interpret the values
  • Introduces (arguably) the remotest branch in the Niger-Congo phylum
Typological characteristics

• Mande languages generally no noun-class system or verbal extensions of the Atlantic-Congo languages (Bɔɔ has causative and intransitive forms of the verb).

• Some have initial consonant mutation on nouns (perhaps a remnant from former noun class prefixes).

• Plurality is often marked with tone, as for example in Sembla.

Typological characteristics

• Pronouns commonly have alienable–inalienable and inclusive–exclusive distinctions.
  – N´ fà túlo ‘My father’s ear’
    1s father ear
  – Cε’à fàli kù ‘The man’s donkey’s tail’
    Man POSS donkey tail

• Word order in transitive clauses is S–(AUX–)O–V(–adverb/-OBL).
  – N´ be’ bamanankan mɛ’n dɔɔ’nin
    I understand a little bit of Bambara (lit: I AUX Bambara hear small-small)
  – Perfective aspect: no AUX! Proper example of SOV…
Typological characteristics

• Both prepositions and postpositions are used (Manding rather PostP):
  – *bàma kùn* ‘the crocodile’s head’
  – *sɔ kùn* ‘on top of the house’
  – *n’ kɔnɔ* ‘my belly’
  – *Bàmakɔ bɛ’ Mài kɔnɔ* ‘Bamako is in Mali’

• Within noun phrases, possessives generally come before the noun, adjectives and plural markers after, while demonstratives are found with both orders. (Williamson & Blench 2000)

Lexicostatistics

[Diagram showing the relationships between different linguistic groups, including Mande, Sya, Soninke, Susu, Vai, etc.]
Percentages

• Which percentage to go by?
  – Nearest neighbour
  – Furthest neighbour
  – Branch average

• Examples
  – Soninke-Bozo 30%
  – Bozo-Bambara & Bozo-Susu 35%
  – Bozo-Ligbi & Bozo-Vai 20-24%

Möhlig 1984/85

• Classification of Swahili and Sabaki (Mijikenda, Pokomo, Ilwana)
• Significance for us:
  – Deals with dialect situations
  – Introduces an improved lexicostatistical method for a first assessment:
    Dialectometry
Swahili and relatives

- Dialect continuum, although with identifiable units
- Pokomo (Tana river valley)
- Ilwana (upper Tana)
- Mijikenda (coastal hinterland in S Kenya)

Möhlig’s principles

- Skepticism against family trees
- Dialectological principle: no blank spaces on the map
- Integration of various historical models
  - “divergence” (split as captured by comparative method)
  - “convergence” (borrowings, contact-induced change, “wave-model” type developments)
  - language shift as a possible historical development
- Relative chronologies of relevant sound-historical developments
Dialectometry 1/2

- Identify dialects in an area
- Collect 600 word lists, and basic grammatical information for all of them
- Calculations on the basis of a shorter list
  - Compare each pair of dialects
  - Lexical values differentiated from phonological values
- “Sophisticated version of lexicostatistics”

Dialectometry 2/2

- Phonological indexes
  - Establish sufficient features to distinguish all realisations of a diaphoneme in the area under study
  - Count the number features for all pairs of dialects that they share, and divide by maximum amount
- Lexical indexes
  - Compare lexical entries in the wordlist and assess similarity
    - 0 points for non-cognate
    - 1 point for accumulated deviations
    - 2 points for phonological irregularity
    - 3 points for morphological deviation (e.g. different noun classes)
    - 4 points for identical (=regularly corresponding) forms
- Comparison of the phonological and lexical index values gives indication of historical scenario.
Bantu expansion

- Content analysis: semantics allow reconstruction of some aspects of history, as Bantu e.g. shows
- Glottochronology
- Least moves principle: Origin in Cameroon/Nigeria border
- Ancestral Bantu had words for fishing, yam & palm oil but not cereals, (larger) livestock; so root crop cultivators, fishers

Bantu expansion: WUSTL.edu

- One group goes across northern forests towards Great lakes of E Afr; probably in Great Lakes area by 1000 BC; this branch picks up cereal cultivation, cattle and sheep herding from Nilo-Saharan speakers
Bantu expansion: WUSTL.edu

- Other come down thru Cameroun; develop words for goat-cattle-axe; develop metalworking (smithing); arch sites show their spread to lower Congo by 400 BC

Historians’ views

- Lamphear & Falola
  - Relatively peaceful interaction between populations
  - Iron-working evidence: Urewe. (<Meroe?)
  - Livestock <Nilotic speakers?, even earlier <Cushitic speakers?
- Vansina contests idea of actual population spread, arguing rather for movement of language and cultural traits “like ripples in a pond”:
  - […] no longer a cogent argument in favor of a Bantu migration […] In the case of language shift the indigines usually became first bilingual and then lost their original language only several generations later […] implying] either a decisive demographic advantage or [that they] enjoyed a huge prestige over the indigenous language.
  - […] The technological differential was small. Did Bantu-speakers have a monopoly on trade? Were they conquerors, or somehow superior in religious matters? There is no evidence so far…
What do the arrows stand for?

- To some extent, migration
- To a larger extent, distributions of linguistic similarities (loanwords, other features)
- (Datable) distributions of archeological evidence (iron sites)

Sir H.H. Johnston
Malcolm Guthrie

- Strongly influenced by structuralism, intends a strict division between synchronic analysis and diachronic interpretation: Common Bantu constructs (numbered starred series C.S.)
- He observes highest similarity/number of represented roots in the area SE of the rainforest: Bemba
- Assumes this as the historical nucleus
- Problem: geographical gap between his Bantu-nucleus and W Africa.

Later work

- Heine, Hoff, Vossen 1977
  - Lexicostatistics
  - Sequence of nuclei
- Möhlig 1981
  - Stratificational model
  - Dialectology, language geography, sound history
  - Primary distinction between Rainforest and Savanna (not necessarily to be understood as an early split!)
- Bastin (Tervuren group), various publications
  - Doubts primary branching RF vs Sav.
  - Instead, Congo basin settled along two different pathways
  - Recent evidence:
    - Bostoen & Wotzka unpubl. Linguistics and archaeology
    - Leitch 2009, VH systems in group C