GENETIC CLASSIFICATION OF LANGUAGES

A. Classification (dictionary definition): It is a division or category in a system which divides things into groups or types.

B. Taxonomic Linguistics:
-- The science concerned with the identification and grouping of concepts, organisms, texts, etc. is called Taxonomy.
-- This is done in every field of study.
-- In linguistics, this approach to things (i.e. linguistic data) is called Taxonomic linguistics.
-- Traditional linguistics is primarily taxonomic in its approach – the goal of taxonomic linguistics is to group individual grammatical items (sounds, words, phrases) into categories, or to group even entire languages into types or families.
-- Taxonomy or classification runs through everything we will be doing in this class.
-- Today I want us to focus on language classification, specifically the classification of languages into families.

C. Explanation of concept: Genetic
-- languages of the world are classified based on the hypothesis of common origin. (Dzameshie, LNGS 224)
-- the term 'genetic' is NOT associated with genes or genealogy (which would mean that a language has properties of at least two languages);
-- the term 'genetic' is associated rather with 'genesis' or 'origin'

D. Genetic classification of languages
-- When languages are known to have developed out of a common ancestral language they are said to be genetically related.
-- Genetic relationships therefore have to do with the linguistic characteristics that are inherited by one generation of speakers from another (as opposed to those which are acquired from other sources).
-- That is, “all languages of the world are classified into families. All languages belonging to a particular family are believed to have the same origin, that is, they originated from the same ancestor language” (Dakubu 1988)
-- Let’s read this quotation from Welmers (1973) to understand this better:

Assume that a linguistically homogenous community splits into two groups, through a process such as migration or invasion that creates a geographical separation between them. As long as neither group completely gives up its own language to adopt the language of some other people, there will now be two separate generation-to-generation continua. Linguistic changes will take place within each continuum, but many or all of the changes will be different for the two. After a few generations, members of the two groups, if reunited, will still be able to understand each other, though they will note peculiarities in each other’s speech. But after several centuries, enough diverse changes will have accumulated so that members of the two groups will no longer be able to communicate with each other. They may then be said to speak different languages, each of which is equally a ‘direct’ descendant of the original common language. Some of the changes that have taken place in each are internal – that is, languages change even apart from contact with other languages. Some of the changes, on the other hand, may be, and usually are, the result of external influence; the most conspicuous of such changes is the adoption of foreign words. But no matter how extensive the external influences have been, each resultant language has had a continuous history from its point of origin, the common parent language. (Welmers 1973: 3 – 4)

-- Languages are related if they are divergent continuations of a single language spoken at an earlier time (page 77).
-- This implies that if we were to trace each language back in time, we would find increasing similarities until finally we would be dealing with a single language.
E. Method of classifying languages genetically, or of reconstructing a proto-language:

Two complementary techniques have been devised for accomplishing such a task—i.e. there are procedures for determining the relationship among languages: the comparative method and internal reconstruction. (That is, by these methods, an ancestral language can be reconstructed. LING 224.)

F. The comparative method: Application

-- Linguists generally agree on this basic steps in constructing language genetically:

(a) Assemble cognate lists

-- What we need is linguistic evidence—in the form of similarities in lexicon and linguistic structure to establish language inter-relatedness.

-- the analyst then seeks and makes a list of, shared cognates (i.e. words meaning the same) across two or more possibly related languages. (Childs 32). A cautionary note:

-- shared-cognates (or similarities) must not be due to cross-language adoption ('borrowing') or attributable to coincidence or chance; shared similarities—or items shared should be the types unlikely to be involved in cross-language adoption ('borrowing')

-- How do you tell whether or not forms were borrowed? You need to have an idea about what is usually borrowed, etc. This problem can usually be overcome by using basic vocabulary (such as kinship terms, numbers, body parts, pronouns, and other basic terms). Nonetheless, even basic vocabulary can be borrowed.

(b) Establish regular/systematic sound correspondence in the vocabulary that was shared by the putative ancestral language: -- Once cognate list are established as in (a), the next step is to determine the regular sound correspondence they exhibit.

(c) Reconstructing proto-phonemes

(d) Interpretation:

(i) How different two languages are from a common ancestor are assumed to be the historical changes

---

1 They share linguistic properties and items which cannot have been borrowed, thereby demonstrating that they have a common ‘parent’. Genetic classification of languages is non-arbitrary in that it reflects history (descent through common ancestor) and thus it is the favorite classification system of linguists.

2 Adapted from LING 224: One main method that has been employed to establish genetic relatedness of languages is Standard Word List. That is, it would be best to compare speech forms of the earlier forms of the languages. So a standard list of words in English is compiled and people are asked to give translations of the words in their own speech forms. The shortest list has about 100 words. The list is made up of words that are not likely to be borrowed. Such words include those that have to do with body parts, kinship terms, and words dealing with common activities like eating, cooking, buying, selling and common nouns as fire, water, salt, house, etc. The probability of having identical speech forms in two linguistic communities will be very negligible, unless the two languages are related. A similarity of about 80% indicates that the two speech forms are dialects of the same language.

Two of the principles that guide the comparative reconstruction of an ancestor languages are: (a) The majority Principle, and (b) The Most natural Development Principle. (a) According to the majority principle, if, in a cognate set, three forms begin with a [b] sound and one begins with a [p] sound, then our best guess is that the original sound in [b], which the majority have retained. (b) The Most Natural Development principle is based on the idea/belief that certain types of sound change are very common, whereas others are extremely unlikely.

---
that they have undergone.
(ii) Shared innovations (Newman 2000a) – sharing common changes and innovations means the languages concerned are more closely related than are other daughter languages (Furthermore, languages that share an innovation of any kind are closer than those that do not)

(iii) Languages that share a greater number of cognates are genetically closer than those that share a lesser number (Higher degrees of similarity reflect closer relationships; lower degrees of similarities reflect more distant relationships.)
(iv) Completely unrelated languages show only a very little random similarity, attributable to coincidence or possibly to adoption of occasional lexical items from one language to another.

(e) Representation of genetic relatedness: Degree of relatedness
(i) The tree model: represented by an upside down tree diagram; daughter languages branch out from a protolanguage; and then each daughter splits into further daughter languages. (Criticism against this representation: change is not abrupt, and the tree model does not reflect this reality.)
(ii) The wave model: it is used to show that change is not abrupt, and that the varieties concerned did not split quickly as the tree model would have us to believe.

G. Conclusion on Method: -- Very often we have the above method of genetic classification supplemented by the internal reconstruction method, and non-genetic methods namely areal/geographic classification, race, mutual intelligibility, and typology.
GENETIC CLASSIFICATION OF AFRICAN AND GHANAIAN LANGUAGES

GENETIC CLASSIFICATION OF LANGUAGES INDIGENOUS TO AFRICA

1. Greenberg’s Grouping of African Languages

-- (a) Africanist linguists including Greenberg (1966), have classified African languages genetically into four large families/phyla:
   - Nilo-Saharan with 196 languages;
   - Khoisan with 35 languages;
   - Afroasiatic/Afrosian/Afrasian with 371 languages; and
   - Niger-Congo (Niger-Kordofanian) with 1,436 languages. (Heine and Nurse, 2000).

-- (b) This classification excludes languages such as Malagasy (which is Austronesian), Afrikaans, English, French, Spanish, and Portuguese.

-- (c) The existing viewpoint is that these language families descended from Proto-Africa.

-- (d) From a classificatory standpoint all phyla are equal, despite differences in size.

2. More information on the four major language families

(2a) Language family 1: Khoi-san

Subfamilies: (a) languages of the San (Bushmen) and Khoikhoi (parts of sub-Saharan Africa); (b) Sandawe (E. Africa/Tanzania); (c) Harts (Hadzane or Hadzapi) (in E Africa/Tanzania).

Geography: Tanzania, Namibia, Botswana, and Angola

(2b) Language family 2: Nilo-Saharan

Subfamilies:

(I) (i) Komuz languages, (ii) Saharan languages (including Kanuri language), (iii) Songhay languages, (iv) Fur languages (including Fur language); (v) Maban languages; (vi) Central Sudanic languages; (vii) Kunama language; (viii) Berta language;

Geography: Sudan, Ethiopia, Uganda, Kenya, and Northern Tanzania; Some languages in Central and West African

Languages: Luo (Sudan); Acholi, Lango, Dholuo (Uganda, Kenya, Tanzania); Ateker, Teso, Karamojong and Turkana (Uganda/Kenya); Maasai (Kenya and Tanzania); Kanuri (Nigeria), Songhay (Malī).

(II) Eastern Sudanic languages –

(i) Nubian languages, and

(ii) Nilotic Languages: Eastern Nilotic languages, Southern Nilotic languages and Western Nilotic languages

(2c) Language family 3: Afro-asia/Afrosian/Afrasian

Subfamilies: Egyptian, Semitic, Berber, Cushitic, Omotic, and Chadic

Geography: N Africa; much of the Sahara; parts of E, central, and W Africa
(2d) Language family 4: Niger-Congo and Kordofanian:
-- there are three major branches of the Niger-Congo: Kordofanian, Mande and Atlantic-Congo (LING 224)
-- The joining of the Kordofanian group of languages to the Niger-Congo group of languages to form a single phyla was proposed in 1950s by Joseph Greenberg
Subfamilies (and geography): Kordofanian (Sudan) and Niger-Congo (Southern and Central Africa and in most of West Africa below the Sahara);
Sub-families and language names
(a) Kordofanian (langs.: Koalib, Tegali, Talodi, Tumtum, and Katla) – languages spoken in Sudan.
(b) Niger-Congo (6 groups)
   (i) West Atlantic (Wolof/Senegal; Temne/Sierra Leone; Fulani/Senegal to East of Lake Chad)
   (ii) Mande – Niger Valley, Sierra Leone and Liberia – Mende (Liberia); Malinke/Maninka (Mali); Bambara (Mali)
   (iii) Gur, or Voltaic (Mossi/Burkina Faso; Dagomba, Mamprusi/N.Ghana)
   (iv) Kwa – Ghana, Ivory Coast, Benin, Nigeria, Liberia (Akan, Ewe, Nupe, Bini and possibly Ijo; *Yoruba, *Igbo (Atlantic-Congo by LING 224,)
   (v) Benue-Congo (or Atlantic-Congo?): Bantu languages: Swahili (Tanzanian, Kenya), Shona (Zimbabwe and Zambia), KinyaRwanda (Rwandan, Uganda and DRC/Zaire, Zulu (South Africa and Lesotho), Xhosa (South Africa); and non-Bantu spoken in Nigeria and Cameroon-- Tiv, Jukun, and Efik;
   (vi) Adamawa-Eastern – Nigeria, Cameroon, North of Bantu Territory (Banda, Zande, and Sango)

3. On the Niger-Congo Language Phyla
4. More on the Niger-Congo Language Phyla

![Diagram of Niger-Congo Language Phyla]

*Proto-Niger-Congo
  Kordofanian

*Proto-Mande-Atlantic-Congo

Atlantic — Mande

*Proto-Ijo-Congo

Ijoid

*Proto-Dogon-Congo

Dogon — Ijo — Defaka

*Proto-Volta-Congo

*West Volta-Congo — *East Volta-Congo = Proto-Benue-Kwa

Kru

*Pre

?Senufo

?Gur-Adamawa — Kwa — West Benue-Congo — East Benue-Congo

Central Gur

Peripheral Gur (not a unified group) + Bikwin

Adamawa

1, 7, 9, 10*

2, 4, 5, 12

6, 13, 14 + Day

Ba

Fali

Kam

Ubangi

Bantoid

Bantoid-Cross

Cross River

* numbers refer to Greenberg’s Adamawa groups, amended by Boyd (1989).

Fig. 2.1. The internal structure of Niger-Congo.
Genetic Classification of Languages Indigenous to GHANA
(Or Language Families in Ghana)

-- (a) most of the languages spoken in Sub-Saharan Africa belong to the Niger-Congo phyla.
-- (b) “the languages of Ghana belong to the Niger-Congo phyla” (LING 224)
-- (c) this means that all languages indigenous to Ghana are ultimately related (i.e. because they all belong to branches of the Niger-Congo family.)
-- (d) Ghanaian languages fall into three major sub-groups, namely “the Kwa group…, the Gur group …, with two tiny pockets of Mande languages.” Amonoo 1989. The diagram below illustrates the subgrouping.

![Genetic Classification of Languages Indigenous to GHANA](https://example.com/diagram.png)

-- (e) “Most of the languages in the north belong to the sub-family of Gur and those in the south belong to Kwa. These two comprise most of the languages of Ghana.” (LING 224, Dakubu 1988).
THE MАНDE LANGUAGES OF GHANA

1. Introduction
-- Mande is the smallest of the three main language families in Ghana.
-- It comprises only two languages, namely Bisa (i.e. Mande language I) and Ligbi (i.e. Mande language II).
-- a considerable proportion of the speakers of both are in adjoining territories of neighboring countries.
-- in Ghana, there is little if any contact between the two (i.e. Bisa and Ligbi).
-- In Ghana neither language occupies a discrete area of its own but is spoken by settled communities interspersed among those of another language. (Dakubu and Naden 1988.)

2. The Bisa Language
(a) Names: the people (i.e. native speakers of the language) are sometimes called Busanga, Busansi or Bussansi

(b) Geography:
-- (i) Bisa is spoken in the far northeast of Ghana (and also in Burkina Faso) 
-- (ii) Many Bisa people have settled further south, in the territories of the Kusaasi and Mamprusi where they form “colonies.” They can be found in a village called Buguri (in Kusaasi) and Nabogo south of Nalerigu.

(c) Varieties/Dialects:
-- (i) There are very considerable variations of dialect within the Bisa language, although all speakers feel themselves to be part of a single ethnic group, and can usually understand the dialects which differ from their own.
-- (ii) the major cleavage is between Eastern and Western dialects.
-- (iii) the dividing line between the two dialects (in the northeastern part of Ghana which is home to this language) runs roughly along the White Volta river.
-- (iv) the Western form/dialect is called Lebir – Note that: the western dialect is spoken in some villages to the East of the White Volta river
-- (v) the Eastern dialect has no generally-accepted name -- a possible term, used by some, is Baraka.

(d) linguistic repertoire/multilingualism:
-- (i) In Kusaasiland and Mamprusiland where the Bisa people have formed colonies, they tend to be multilingual.
-- (ii) They speak either Kusaal or Mampruli or hausa, a number of them speak French and/or English.

3. The Ligbi Language
(a) Name: There is no single name for either the language of the people, although the unity of the varieties is recognized by all.

(b) Geography: Ligbi is spoken in north-western Brong-Ahafo

(c) Sub-grouping
-- (i) Muslim Ligbi of Banda-Ahenkro and Manji
-- the people are called Ligbi, and the term Gyogoh is applied to the language-variety (i.e. of Banda-Ahenkro and Manji); the term used in Akan to refer to their variety (i.e. Gyogoh) is Dwera (i.e. similar to the related Jula).

3 It closest relatives are the Marka or Samo of Burkina Faso and the Busa of the Niger-Bening-Nigeria frontier; and others in the Ivory Coast and Liberia.
4 -- their homeland is mainly in (i) Burkina Faso (i.e. South of Ouagadougou and West of Tenkodogo, on both sides of the upper course of the White Volta.)
5 It is also spoken in the Bondoukou area of the Ivory Coast. (Dakubu and Naden 1988)
(ii) (*Numu) Non-Muslim of Brawhani and Kwametintini:

- the people are called Numu (Noumou) or Ton (which are names applied to them by the Muslim Ligbi).

- The term used in Akan to refer to their variety/language is Ntoore, or Atumfuɔkasa “blacksmiths’ language.”

- The non-Muslim Ligbi-speakers themselves refer to their language by this term or a Ligbi expression of the same meaning, but they do not use the name “Numu” (i.e. a term applied to them by the Muslim Ligbi).

(iii) Muslims of Namasa

- the term Hwela (Huela or Wela) refers to the language and also the people.

- These are distinctions on social grounds but which coincide with the dialect divisions – the dialect differences seem to be minor – which means that the varieties (Gyogoh, Atumfuɔkasa and Hwela) are mutually intelligible.

(d) Language use/linguistic repertoire (of Ligbi people)

- (d-1) the villages in which Ligbi (in its slightly different forms, namely (i) Gyogoh/Dwera as spoken by Muslim Ligbi; (ii) Ntoore/Atumfuɔkasa spoken by non-Muslim Ligbi/*Numu; and (iii) Hwela by Hwela Muslims of Namasa) is spoken are interspersed among Nafaanra-speaking villages. (ii) For example, Many Ligbi-speaking towns, including Banda, Manji and Brawhani consists of a Ligbi-speaking sector and a Nafaanra-speaking one – the Ligbi sector being numerically dominant in the latter two. (iii) Despite this, the two linguistic groups tend not to speak each other’s languages.

- (d-2) the second language among the Ligbi speakers is the Asante dialect of Akan, which is spoken by virtually everyone, and Jula is also spoken by the majority of adults.
4. Diagram of Proto-Mande

Below is a diagram on the Proto-Mande language family

The languages we have been focusing on can be found under the West and East divide of the diagram: the West shows Ligbi with its varieties: (i) Jọgọ/Gyogoh, (ii) Atumfungkasə/Atumfour, (iii) Wela/Hwela; and the East shows Bisa – Bisa has the varieties Lebir (the Western dialect), and Baraka (the Eastern dialect).
Other methods of classification fall under: Areal Linguistics deals with the features of linguistic structure that tend to characterise a particular geographical area. Typological classification: Focuses on formal generalisations which apply to even genetically unrelated languages. That is, the classification of languages according to their structural characteristics. Typological studies group languages on the basis of similarities in their syntactic patterns, morphological structure and phonological systems. Phonological Typology: according to size of vowel or consonant systems; suprasegmental systems – tone, intonation, and stress; syllable structure; morphological typology – isolating, agglutinating, inflectional, polysynthetic.