Toponymy course

9. Languages

Tjeerd Tichelaar
Toponomy and language

- Geographical names belong to languages
- Grammatical and orthographic rules apply to the origin of names
- Toponymists should be acquainted with the language of the names
Names vs. language

- Like all utterances of language, toponyms consist of *words* and *phonemes*.
- The *grammar* of a language provides the structure through which the language and elements thereof (like names) convey *meaning*.
- Syntactic rules determine the ways in which the elements of a language may be combined.
Stratford-upon-Avon =

1. (settlement by the) ford where the Roman road crosses the river Avon

2. The birthplace of William Shakespeare.

stræt ‘paved (Roman) road’
ford
upon
Avon river name meaning ‘water’
Languages of the world

- SIL counts 7,000 different languages
- Languages are grouped into *families / phyla*
- Ca. 100 families and phyla
- More than 75% of the known languages belong to 10% of the families
- 2/3 of the world population speak languages belonging to only 2 families: Indo-European and Sino-Tibetan
- For toponymy, numbers of speakers are not the sole criterium of ‘importance’.
Toponymic importance of a language

- Area occupied – now and in the past
- Lifestyle
- Geographic attachment
- Historic rights
Toponymic importance of linguistic status

- An official language is ‘a language expressly adopted by the government of a country … and employed as a language of administration’

- A non-official language is ‘a language that lacks official status in a particular legally constituted entity’

- A dialect is ‘a variety of language which is distinguished by phonological and/or morphological characteristics that give it a distinctive identity’.

- A literary language is a ‘written form of language regarded as the desirable standard for works of literature’
- A *national language* is a ‘language in widespread current use throughout a given country or in part of its territory …’, and it ‘… may have or may not have the status of an official language’.

- A *minority language* is ‘any language not used by a significantly large part of the country’s population’

- A *principal language* is ‘in a linguistic community where more than one language is in use, that language which has greatest currency’.
• A *living* resp. *dead language* is ‘any language spoken today, resp. not longer spoken’.

• A *lingua franca* is a medium of communication for people who speak different first languages.

• A *pidgin* is a language with a reduced range of structure and use, with no native speakers.

• A *creole* is a pidgin that has become the mother tongue of a speech community.
Language vs. dialect

- Analytical linguistic definition:
  - lexical correspondence
  - grammatical similarity

- Politico-linguistic definition:
  - official status
  - literary standard
  - dialect chains cut by ‘language’ names:
    ethno-linguistic identification
● Toponymic definition: a language is ‘a system providing a means by which the members of a community can communicate orally and/or graphically’

● The majority of toponyms has once been created by the (local) community.

● Those terms traditionally close to the communities’ daily experience – like geographical generics – are the first to be distinctive.
Official vs. non-official languages

- In large parts of the world, the official language is not even the language actually spoken today in the largest part of the country.

- Where the official language is a foreign language or a relatively new language developed from a *lingua franca*, the official language is the language of just a few toponyms, or even none at all.

- Here the toponyms belong to the local vernacular, whether this is considered a language, a dialect, or a sub-dialect.
Dead and disappeared languages

- Toponyms generally outlive their creators, thus locally disappeared and even ‘dead’ languages are not necessarily deprived of their importance from a toponymic point of view.

- Dead languages may come to life in toponyms.

- *Substrata*: residues of local predecessor languages in unrelated immigrant successor languages.

- To understand toponymy or language in general, knowledge of settlement history is indispensable.
The Khoisan family

- An ancient family of a few scores of languages currently still spoken in the Kalahari and Namib Desert areas of southern Africa, as well as in some isolated areas in Tanzania.
- Characteristic ‘click’-sounds, represented by !.
- In earlier days native to a larger area than they are at present.
- Nama (150,000 in Namibia) and Sandawe (70,000 in Tanzania) most numerous.
The Niger-Congo family

- Most prominent language family of sub-Saharan Africa: up to almost 1,500 separate languages distinguished.
- A hypothetical common ‘proto-Niger-Congo’ ancestor thought to have ceased to exist 5,000 years ago.
- Exact hierarchical subdivision still under debate.
- The largest of the sub-families is made up by the *Bantu* languages, native to Central and the largest part of Southern Africa.
Most spoken languages:
- Swahili (5 M first language speakers in the East African countries, plus 30 million using it as a second language *lingua franca*)
- Yoruba (20 M) in Nigeria and the eastern part of West Africa; Igbo (17 M) of Nigeria
- Fulani (13 M, including second language speakers) in West Africa
- Wolof (2.7 M in Senegal, plus 7 million second language speakers)
- Zulu (9.5 M) of South Africa, Lesotho, Swaziland and Mozambique
- Rwanda (9.5 M) of Rwanda and adjacent countries
The Nilo-Saharan family

- 200 languages spoken in the southern fringe of the Sahara and around the upper course of the Nile, from Mali in the west to Eritrea and Tanzania in the east.

- Subdivision under fierce debate.
The Afro-Asiatic family

- Almost 400 languages spoken in Northern Africa and Southwest Asia.
- Semitic subfamily in Asia and Ethiopia/Eritrea (Arabic, Hebrew Neo-Aramaic languages, Ethiopic languages)
  - Berber subfamily in N.W.-Africa
  - Chadic subfamily in Chad and Nigeria (Hausa)
  - Cushitic subfamily in N.E. Africa
- Egyptian subfamily (Coptic)
- Omotic subfamily (Ethiopia)
The Indo-European family

- Spread since the 3rd Millennium B.C. all over Europe and well into S.W. and S.-Asia.

- Branches:
  - Italic (Latin and Romance languages)
  - Celtic (Gaelic, Welsh, Breton)
  - Germanic (German, English, Dutch, ...)
  - Baltic (Lithuanian, Latvian)
  - Slavic (Russian, Polish, Czech, ...)
  - Hellenic (Greek)
  - Indo-Iranian (Indian and Iranian languages)
  - Armenian
  - Albanian
The Dravidian family

- Mainly spoken in the southern part of South Asia.

- Native to the whole Indian subcontinent by the time the Indo-European languages spread into the Indus and Ganges valleys in the late 3rd Millennium B.C.

- Most important:
  - Telugu (75 M in Andhra Pradesh, India)
  - Tamil (70 M in S.E. India and N. Sri Lanka)
  - Kannada (45 M in Karnataka, India)
  - Malayalam (35 M in Kerala, India)
The Caucasian family

- Spoken in the Caucasus area.
- Two major branches, North Caucasian and South Caucasian, might not be related.
- Official language of Georgia, national language of some autonomous republics in the Russian Federation (Chechnya, Ingushia etc.). Also in Abkhazia.
The Uralic family

- N. Europe, Hungary and W. Siberia.
- Presumed common proto-Uralic ancestor as early as the 6th Millennium B.C.
- Two branches: Finno-Ugric and Samoyedic.
- Most important:
  - Hungarian
  - Finnish
  - Estonian
  Also Sami and some national languages in autonomous republics of Russian federation: Mordwin, Komi, Mari, Udmurt etc.)
The Altaic family

- Spread through several waves of migration during the 1st Millennium A.D. from (probably) an area to the east or northeast of Central Asia in a western and southwestern direction.

- Three branches: Turkic, Mongolian and Manchu.

- Turkic languages official in Turkey, Azerbaijan, Turkmenistan, Kazakhstan, Uzbekistan, Kyrgyzstan.

- Mongolian languages spoken in Mongolia, China, Russia (Kalmyk, Buryat).
The Sino-Tibetan family

- 1.3 billion speakers of Mandarin Chinese (900 million) and other Sinitic languages share the same written language.

- Tibeto-Burman languages spoken in Myanmar, Tibet and the Himalayan area (Dzongkha in Bhutan, etc.).

- Although Chinese emigrants brought their language with them all over the world, it never spread to non-ethnic Chinese populations elsewhere.
The Austro-Asiatic family

- Between Southern China and the Malay Archipelago.

- Mon-Khmer languages of Indo-China include Vietnamese (disputed), Khmer (in Cambodia), and ancient literary language Mon (Myanmar).

- Munda languages in India.
The Daic family

- Including official languages of Thailand (Thai) and Laos (Lao).
- Attempts are made to link it to the Sino-Tibetan or the Austro-Asiatic family.
The Austronesian family

- Largest geographical span: from Madagascar to Easter Island (Chili).
- Historic heartland in the Malay Archipelago, including Taiwan.
- Official languages of Indonesia (Bahasa Indonesia) and Malaysia (Bahasa Malaysia), the Philippines (Pilipino), Madagascar (Malagasi), and several island states in the Pacific.
- Almost 1,300 different languages.
The Indo-Pacific phylum

- Groups the languages of the Southwestern Pacific and Northeastern Indian Ocean area not belonging to the Austronesian family
- More than 1,000 different languages spoken on the island of New Guinea, grouped in 13 possibly unrelated ‘stocks’.
- Also in Solomon Islands, Eastern Indonesian provinces, and Indian Andaman and Nicobar Islands.
The Australian family

- Perhaps the oldest living languages on earth (Australian aboriginals have been living where they are for maybe 40,000 years)
- In the 18th Century still 500 languages (est.), nowadays 250 survive.
- Divided in 28 different subfamilies, 27 of which spoken in the northern 1/8 of the country.
The Paleo-Siberian languages

- Grouping of 4 most probably unrelated language families in Eastern Siberia.
- Less than 30,000 speakers, but covering a large area.
- Chukchi and Koryaks have autonomy within Russian Federation.
Amerindian languages

- Languages of the population native to the Americas before the ‘discovery’ of Columbus.
- At least 50 separate language families.
- Most prominent:
  - Oto-Manguean family (170, in Mexico)
  - Arawakan family (75, Honduras to Argentina)
  - Tupi family (70, Paraguay to French Guyana)
  - Uto-Aztecan family (60, Western USA to El Salvador)
  - Quechuan family (47, Andean area)
  - Na-Dene family (42 from Alaska to S.W. USA)
  - Algic family (33, Canada and USA)
Language isolates

- Japanese (125 M speakers).
- Korean (75 M speakers)
- Basque (minority in French-Spanish border area).
- Famous extinct languages:
  - Sumerian
  - Etruscan
  - Language of Indus Valley civilization
  - Pre-Indo-European languages of Europe (Iberian, Ligurian)
When collecting geographical names, we need to \textit{identify} the language the name belongs to in order to know which phonological, grammatical and syntactic rules apply to it.