

Tokyo University of Foreign Studies
Studies in Linguistics 2

Geographical Typology and Linguistic Areas

edited by Osamu Hieda,
Christa König and Hiroshi Nakagawa

John Benjamins Publishing Company

Geographical Typology and Linguistic Areas

With special reference to Africa

Edited by

Osamu Hieda

Tokyo University of Foreign Studies

Christa König

University of Cologne

Hiroshi Nakagawa

Tokyo University of Foreign Studies



John Benjamins Publishing Company

Amsterdam / Philadelphia



The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

Geographical typology and linguistic areas : with special reference to Africa / edited by Osamu Hieda, Christa König and Hiroshi Nakagawa.

p. cm. (Tokyo University of Foreign Studies (TUFS), studies in linguistics, ISSN 1877-6248 ; v. 2)

Includes bibliographical references and index.

1. Typology (Linguistics) 2. Africa--Languages. I. Hieda, Osamu, 1951- II. König, Christa. III. Nakagawa, Hiroshi, 1955-

P381.A3G46 2011

409.6--dc22

2010049191

ISBN 978 90 272 0769 2 (Hb ; alk. paper)

© 2011 – Tokyo University of Foreign Studies

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands
John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

Geographical Typology and Linguistic Areas

Tokyo University of Foreign Studies (TUFS)
Studies in Linguistics

Volume 2

Geographical Typology and Linguistic Areas. With special reference to Africa
Edited by Osamu Hieda, Christa König and Hiroshi Nakagawa

Contents

Message from the President

Ikuo KAMEYAMA (President, Tokyo University of Foreign Studies)..... 1

Center for Corpus-based Linguistics and Language Education

Makoto MINEGISHI (GCOE Project Leader) 3

Introduction

Christa KÖNIG 7

Section 1

Areal Features and Linguistic Areas: Contact-induced Change and Geographical Typology

Alexandra Y. AIKHENVALD 13

Areas of Grammaticalization and Geographical Typology

Bernd HEINE 41

Case Marking and Linguistic Geography

Christa KÖNIG 67

Can Ethiopian Languages be Considered Languages in the African Linguistic Area?

The Case of Highland East Cushitic, particularly Sidaama and Kambaata

Kazuhiro KAWACHI 91

Proto-Bantu and Proto-Niger-Congo: Macro-areal Typology and Linguistic Reconstruction

Tom GÜLDEMANN 109

Section 2

Explaining Convergence and the Formation of Linguistic Areas

Yaron MATRAS 143

Is Kumam a Creole Language?

—A Mechanism of Linguistic Convergence in the Southern Lwo Area

Osamu HIEDA 161

The Continuum of Languages in West Tanzania Bantu:

A Case Study of Gongwe, Bende, and Pimbwe

Yuko ABE 177

Patterns of Linguistic Convergence in the Khoe-speaking Area of Southern Africa

Rainer VOSSSEN 189

Tense and Aspect in Khoesan: The case of Ju/'hoansi <i>Budzani GABANAMOTSE-MOGARA</i>	201
--	-----

Section 3

Ritual Pathways: Contact in a Framework of Difference, Imitation and Alterity <i>Anne STORCH</i>	213
---	-----

The Eastern Kalahari Khoe: A Focus on Inter-Khoisan Ethno-language Dynamics around the Makgadikgadi Salt Pans of Botswana <i>Andy CHEBANNE</i>	233
--	-----

Language Contact and Social Change in North-central Namibia: Socialization via Singing and Dancing Activities among the !Xun San <i>Akira TAKADA</i>	251
--	-----

Two Types of Kinship Classification Found among the Khoe Languages —Relative and Absolute Calculations in Determining the Seniority among Classificatory Siblings <i>Hitomi ONO</i>	269
---	-----

A First Report on G ui Ideophones <i>Hirosi NAKAGAWA</i>	279
---	-----

Section 4

Noun-Modifier Order in Africa <i>Matthew S. DRYER</i>	287
--	-----

Index of Authors	313
------------------------	-----

Index of Languages, Language Families and Areas	315
---	-----

Index of Subjects.....	318
------------------------	-----

Contributors	321
--------------------	-----

Message from the President

Ikuro KAMEYAMA (*President, Tokyo University of Foreign Studies*)

It was a great honor for me to participate in the international symposium entitled “A Geographical Typology of African Languages”. This symposium was also attended by the eleven international scholars who have traveled to join us today from America, Australia, Botswana, Germany, and United Kingdom.

I will now dwell briefly on the Global COE Program, which began in April 2007. This program is an effort by the Japanese government to strengthen its support to research and educational institutions in which internationally renowned work is taking place. The program was developed to take advantage of world-class resources to help foster the development of creative researchers who can lead in their fields, and to strengthen research and education in Japan’s centers of graduate education. In 2007, proposals were solicited in five areas. The program submitted in the area of humanities by our university was one of 12 selected nationally. The humanities category encompasses fields as diverse as philosophy, art, psychology, education and archaeology. The submission from our university was the only one selected in the area of linguistics. We believe this reflects the high level of research and education at our institution.

Our submission, entitled “Corpus-based Linguistics and Language Education” emphasizes on a field of empirical linguistics based on the uses of corpora. The program’s goal is to foster the growth of advanced researchers with international perspectives. This program continues the research conducted under the “Usage-based Linguistic Informatics” 21st Century COE Program, which concluded March 2007. The new program will build on the international joint research framework that was created over the past five years to achieve two goals, with the support of the entire university:

1. To further develop a comprehensive education program for graduate students
2. To give graduate students opportunities to perform fieldwork, build and analyze corpora, and receive language education and training, both in Japan and overseas.

I am not an expert in linguistics nor do I have a deep scholarly understanding of corpus linguistics. However, as a scholar of literature, I have a keen interest in the possibilities inherent in the field. The corpus concept was introduced into my area of specialization, Russian literature, in the late 1980s. As far as I know, this resulted in the creation of corpora for the works of authors such as Fyodor Dostoevsky and Andrei Platonov. However, it is not yet clear

how effective the corpus concept will be in the development of the study of literature. In contrast, corpus-based linguistics seeks not to use linguistic data to understand the latent properties of a text as a close system, but to understand the linguistic structure and function of a language within a larger context. So, I believe that corpus linguistics provides us with higher objectivity and richer possibilities in the field of humanities.

Still, it is my opinion that the greatest hurdles for corpus-based linguistics are still to come. Humans are creatures that cannot help but seek out meaning and possibilities of systemizing matters. It is evident that corpus linguistics has not been a field that describes only the actual uses of languages, but one that finds ways to generalize creative discoveries and to extend its insights. Its value lies in its ability to push itself. For corpus-based linguistics to grow creatively as a human science, we must help young researchers to develop innovative and unique capabilities for analysis. I believe that this is where the real importance of the current G-COE Program lies.

In conclusion, I would like, as president of this institution, to express my sincere respect to all the leading researchers who attended this symposium, for their untiring efforts. More importantly, I hope that the young scholars who attended the symposium have imbibed some of the passion that was on display, and I hope that it will help them to grow internationally competitive researchers.

September 1, 2010

Center for Corpus-based Linguistics and Language Education

Makoto MINEGISHI (*GCOE Project Leader*)

The Center for Corpus-based Linguistics and Language Education (CbLLE) was established with the express target to build an education and research center with unique strengths in the study of linguistic diversity and also in usage-based research of linguistic structure and language education. This centre builds on the strengths of the nationally high-ranking Graduate School of Area and Culture Studies of the Tokyo University of Foreign Studies (TUFS) and of the Research Institute of Languages and Cultures of Asia and Africa (ILCAA). Its educational and research uniqueness is achieved by integrating the three core areas of activities: (a) collection and analysis of naturally occurring language use data through field research, (b) compilation and analysis of large-scale corpora of language use data from a wide range of languages, and (c) application of corpus-based linguistic analyses to language education and pedagogy. A few details of the work that is being done in the above core areas follows:

Field linguistics: The field linguistics program supports fieldwork-based research on a typologically diverse set of languages, including not only the world's major languages but also lesser-studied languages. They include languages of Africa, Eurasia, and North America. It also aims at advancing typological research on the basis of the primary data from a broad range of languages. It provides a solid training to students in the methodology of collecting, processing, and analysis of the field data. The project has undertaken fieldwork-based study of a diverse range of languages of the world (lesser-studied languages in particular) and typologically-informed description of these languages. Some of the projects under this category are: Compilation of a Word List for Field Research on Khwe Languages; Field-work based study of under-studied speeches of India; Collation of Spontaneous Conversational Data of Individual Languages such as: Swahili, Russian, French, Spanish, Turkish.

Corpus linguistics: The program in corpus linguistics supports analysis of a large amount of language use data and compilation of corpora, which feed into linguistic informatics research and also into descriptive and typological research. Some of the specific targets are: Building electronic corpora and developing analysis and processing tools in order to support new ways of analyzing language data and multipurposing of the data; Developing

multilingual and multifunctional integrative corpora of language use for major languages on the basis of language use data collected in language teaching classrooms, blogs, etc.; Conducting international collaborative research and providing support in development and utilization of tools for corpus creation, morphological analysis, electronic dictionary creation, text analysis. The projects undertaken here include: Development of Electronic Dictionaries for Russian as well as Thai (separately); corpus Compilation of Data from Medium/Minor Language Groups; Development of Utility Manuals for German Corpus; Preparation of Introductory Text-book on Lexicology based on Corpus Data; Research on Corpora for Minor Language Group in EU Countries.

Linguistic informatics: The linguistic informatics program builds on the research in field linguistics and corpus linguistics components to significantly advance research in language pedagogy. It seeks to make a significant contribution to the research in language pedagogy by taking into account the factors of linguistic and cultural diversity through analysis of corpora of language use in actual contexts of language instruction, including naturally occurring conversations and learners' language use. A few studies undertaken in this context are: Research on Lexicon/language-use based on Corpora for Various Fields; Language Processing/education Technology; POS Search Engine for Spoken French as well as Spoken Spanish; Basic Research on E-learning through Moodle; Corpora of Learners' Language Use (both as an internal project as well as an international collaborative project); Creation of Language Tests based on Error Analysis of Language Use of English Learners.

The GCOE trains researchers and educators who have a clear understanding of the nature and significance of linguistic and cultural diversity and can take a flexible research approach to language structure and language education. This project equips young researchers with a broad foundation for linguistic research by providing practical trainings in field research, corpus-based research and language education. These training programs support the integrative research on linguistic and cultural diversity and usage-based linguistics by connecting effectively field data collection, data analysis, and educational application of theoretical insights obtained from the analysis.

The specific projects and tasks listed above form part of the larger plan of building an international research and education center with more generalized targets described below. The Center seeks to build a world-leading research and education center in the study of linguistic diversity and in the usage-based research of linguistic structure. The national and international infrastructure for supporting the GCoE are being built through the following activities:

Formation of an international network of collaborative research:

Collaboration in corpus creation and in development of analysis tools (such as electronic dictionary systems); building a network of international collaboration and academic exchange in linguistic research and teaching within the framework of the ‘Consortium for Asian and African Studies’ which has its headquarters at the University.

Expansion of opportunities for academic interaction across institutions and across countries: Expanding opportunities for young researchers, as well as established scholars, within Japan and abroad to assemble and interact through visiting scholar programs and through employment.

Support program for young researchers: Providing young researchers with financial and technical support for linguistic field research, corpus creation, and education research in the field; and providing young researchers with financial support for professional development (including presentation at international conferences).

Active international dissemination research results: Building an information technological infrastructure that supports active electronic dissemination of research results; and Publishing the research results in a series of publications through international publishers that are specialized in publication in linguistics — the present volume being a contribution towards this aim.

Introduction

Christa KÖNIG

The present volume is the result of a conference held at the Tokyo University of Foreign Studies, Research Institute for Languages and Cultures of Asia and Africa Studies in May, 2009. We are very grateful to the Global COE Program, ‘Corpus-based Linguistics and Language Education (CbLLE)’ project members who provided the funds for this conference.

For roughly half a century, work on the reconstruction of African languages and their interrelationship has been based on the work of Joseph Greenberg (1963). What this work has established in particular are findings such as the following: (a) The most easily accessible way of describing the historical relationship of these languages is by reconstructing their genetic relationship patterns. (b) The multitude of African languages can be reduced to four genetically defined units, called families by Greenberg and phyla by others. These units are Niger-Congo (or Niger-Kordofanian), Nilo-Saharan, Afroasiatic, and Khoisan.

More recent research suggests that the typology of African languages appears to be strongly shaped by language contact and areal diffusion (Heine and Nurse 2008), and one main goal, for some authors in fact the only goal, is to show that the history and structure of the languages of Africa was shaped more by social and linguistic interaction than by any other factors.

Areal influence on languages, the topic of the present volume, is to be looked at from various angles. Africa is the primary region to be considered; however, findings from other regions of the world are also included, as in the chapter by Sasha Aikhenvald, who has vast knowledge e.g. about language contact phenomena in Amazonia, and Yaron Matras, whose contribution contains findings on European languages.

The contributions of the present volume are on the one hand theory-oriented, in particular those by Sasha Aikhenvald, Bernd Heine, Anne Storch, and Yaron Matras. On the other hand, they are data-oriented, like the chapters by Kazuhiro Kawachi, Christa König, Tom Güldemann, Osamu Hieda, Yuko Abe, Rainer Vossen, Budzani Gabanamtse, Hitomi Ono, Hiroshi Nakagawa, and Matthew Dryer. Sociolinguistic aspects are in the focus of the contributions by Yaron Matras and Andy Chebanne. That contact phenomena also influence the socio-cultural behavior is demonstrated in Hitomi Ono’s and Akira Takada’s contributions.

Geographical areas covered in particular are southern Africa with its Khoisan languages, as in the chapters by Rainer Vossen, Budzani Gabanamtse,

Andy Chebanne, Hitomi Ono, Hiroshi Nakagawa, and Akira Takada, and eastern Africa, see Hieda's, Kawachi's, and König's chapters.

The present volume consists of sixteen papers, highlighting a number of aspects of the linguistic geography of Africa. Sasha Aikhenvald demonstrates with an example from the Vaupés River Basin in north-west Amazonia how language contact of genetically unrelated languages may lead to structural isomorphism, whereby the grammar and semantics of one language are almost fully replicated in another. According to her, a profound analysis of areal contact involves at least two steps: The macro-scale analysis leading to macro-zones which are defined by the presence or absence of a particular linguistic feature. A detailed knowledge of the processes involved however is not possible without the second step, namely a micro-analysis of the macro-zones, which results in *low tier convergence zones*, "where languages share more than just features" but morpheme intertranslatability.

Bernd Heine highlights the potential of grammaticalization areas, that is, areal occurrences of a particular grammaticalization, for historical reconstruction in Africa. This can be seen in the following facts: Grammatical forms may be influenced by the socio-cultural environment in which they take place. He exemplifies this aspect with the following example: Cross-linguistically, the probability that a given unknown language takes 'body' as the source concept for a reflexive marker is the highest. However, the choice which any given unknown African language has made, whether 'body' or one of the other four common sources for reflexives in African languages, is really influenced. A particular language change may be the result of internal or external forces. Heine stresses that external and internal changes are by no means mutually exclusive, as he puts it.

Christa König presents a nearly salient feature of Africa: Marked-nominative, a type of case system with a nominative-accusative alignment but where the accusative, encoding the object, is the morphologically unmarked case. It occurs most frequently in East Africa. In the border area of Kenya, Uganda, and Sudan it cuts across the genetic boundaries of Nilo-Saharan languages, such as Nilotic and Surmic subgroups, and Afroasiatic languages with Omotic and Cushitic. Areal influence must have played a role. There is evidence that it originated in Afroasiatic before spreading to Nilo-Saharan languages. The areal situation between Dhaasanac (Afroasiatic) and Turkana (Nilo-Saharan) suggests the opposite direction, however. König suggests that the latter reflects a younger development triggered by contact.

Kazuhiro Kawachi questions whether Highland East Cushitic languages can be considered as languages of an African Linguistic Area. Kawachi measures the degree to which Highland East Cushitic languages are typical African languages according to Heine and Leyew's quantitative approach

(Heine and Leyew 2008). The latter define Africa as a linguistic area using eleven features. Kambaata and Sidaama, two Highland East Cushitic languages, get only very low score, that is, two of the eleven possible African features. And this holds even more for other Highland East Cushitic languages in general. Accordingly, he argues, Highland East Cushitic languages are only peripheral members of the African Linguistic Area. In the history of Highland East Cushitic it has already been speculated whether they have originated from outside Africa. Kawachi proposes that future research should investigate whether Highland East Cushitic languages show more similarities with other areas, such as southern Europe or southwestern Asia.

Tom Güldemann's main concern is with the question of what areal linguistic findings can tell us about history. According to him, the early Bantu languages more likely had a structure which is recurrent in the Macro-Sudan belt than in Narrow Bantu languages of today. Therefore, when trying to reconstruct early Niger-Congo, Narrow Bantu structures are less helpful than macro-areal typology taken from the Macro-Sudan belt.

Yaron Matras with his paper on explaining convergence and the formation of linguistic areas, takes a critical view of the concept linguistic area. He proposes a re-examination of the term. For him, "Linguistic areas are simply cases of convergence that catch our attention because of the density of shared isoglosses in a multiplicity of languages." Bilingualism, language loyalty, linguistic creativity and lax normative control, as he puts it, are preconditions for areal diffusion. There is evidence to suggest that "forms of clause organisation" are a prominent domain for convergence. Therefore it is necessary to compare the clause structures in the respective languages. His examples are primarily taken from well described non-African languages, such as a comparison of Macedonian Turkish with Standard Turkish, or Macedonian, Neo-Aramaic, and Kurdish, or English, Hebrew and German.

Osamu Hieda covers language contact phenomena in the verbal system of Southern Lwo languages (Nilo-Saharan). According to him, the Southern Lwo languages in general developed nominal prefixes and tense morphemes as a result of contact with the neighboring Bantu languages. Kumam, also Southern Lwo, however, is an exception. Kumam has a new verbal system, which might be the result of language shift possibly from Eastern Nilotic to Western Nilotic.

Yuko Abe presents the first data about an unknown Bantu language, Gongwe, spoken in West Tanzania. According to Abe, Gongwe is neither mentioned in Guthrie (1967-1971) nor in the Ethnologue (Lewis 2009). The native speakers of Gongwe describe Gongwe as a language intermediate to Bembe and Pimbwe. Abe's results show that Gongwe has a Bende-like phonology and a Pimbwe-like grammar. Abe's case is an example of language

contact which has led to a severe kind of language mixture.

Rainer Vossen illustrates patterns of linguistic convergence in the Khoe-speaking area of Southern Africa. He presents cases of phonological, morphological, and lexical convergence among Khoe languages of Southern Africa, including are Buga, Danisi, and Ts'ixa. Particular findings are e.g. the following: Languages and dialects of the NARO-IGANA subgroup of Khoe play a significant part in the complex contact relations. Within the Khoe family, their sound inventories are unique in so far as they contain pharyngealized oral and nasal vowels that do not exist in other subgroups.

Budzani Gabanamotse presents the tense aspect system of Ju/'hoansi, a !Xun language (former Northern Khoisan). The Ju/'hoansi variety considered in her paper is spoken in the Ghanzi and Northwest area of Botswana. She does so by exploiting generally the language situation of Botswana, giving some background information on the overall language situation of Botswana, and of the minority languages, like Ju/'hoansi in particular.

Anne Storch develops a framework in which particularly extra-linguistic features are considered, such as socio-history, culture, rituals, speakers' beliefs, etc. Central Nigeria covers about a hundred different languages which are still being kept apart, although there are many convergence phenomena among them. One reason for this lies in the fact that, according to Storch, the speakers' beliefs influences linguistic change. In Storch's example, the fact that the languages are spoken along river systems plays an important role and shapes the nature of language contact of e.g. Jukun (West Nigeria).

Andy Chebanne takes a socio-linguistic viewpoint when describing the language situation in the Makgadikgadi Salt Pans of Botswana. Language shift from Eastern Kalahari Khoe (former Central Khoisan) to Bantu languages, such as Setswana or Kalanga, are triggered by extra-linguistic factors. "The negative attitude", as Chebanne puts it, from one Kalahari Khoe speaker to another may trigger a shift to languages which are seen as more neutral, such as Setswana or Kalanga.

Akira Takada demonstrates that areal influence can affect the social behavior of people. He exemplifies how children of two forager groups, !Xun and Ju, change their behavior under the influence of Bantu neighboring groups.

Hitomi Ono describes the complex way in which the people speaking G|ui (a Khoe language, former Central Khoisan), classify the category "Sibling", showing that there are two systems at work, a relative and an absolute one, and looking into the question of how according to these systems seniority among siblings is determined.

Hiroshi Nakagawa's topic is ideophones in G|ui. According to him, G|ui shows *inter alia* morphologically derived ideophones which are facilitated by contact with another Khoisan language, !Xóo (Taa family, former