

Language and Identity
Theories and experiences
in lexicography
and linguistic policies
in a global world

Edited by
Ilaria Micheli,
Flavia Aiello,
Maddalena Toscano,
Amelia Pensabene

EUT

ATrA 7

Aree di transizione linguistiche e culturali in Africa



ATrA

Aree di transizione linguistiche e culturali in Africa

7



This book has been edited and published in the framework of the RICFVG 2018 programme funded by Regione Autonoma Friuli Venezia Giulia



Opera sottoposta a *peer review* secondo
il protocollo UPI – University Press Italiane



EUT Edizioni Università di Trieste, Trieste 2021.

ISBN 978-88-5511-267-3 (print)
ISBN 978-88-5511-268-0 (online)

EUT Edizioni Università di Trieste,
via Weiss 21 – 34128 Trieste
<http://eut.units.it>
<https://www.facebook.com/EUTEdizioniUniversitaTrieste>

Language and Identity
Theories and
experiences
in lexicography
and linguistic policies
in a global world

Edited by
Ilaria Micheli
Flavia Aiello
Maddalena Toscano
Amelia Pensabene

Table of contents

Introduction - Ilaria Micheli	7
PART 1: GENERAL ISSUES	17
Linguistic identity in and out of Africa - Karsten Legère	18
Multilingualism in India and the significance of multilingual dictionaries - Suhnu Ram Sharma	35
The natural supremacy of spoken language. Orality and writing in Africa - Ilaria Micheli	43
From IPA to wildcards: a critical look at some African Latin orthographies - Mauro Tosco	56
PART 2: LEXICOGRAPHY AND DIDACTICS OF L2	85
L'italiano come seconda lingua nei CPIA (Centri Provinciali per l'istruzione degli Adulti): questioni e Prospettive - Emilio Porcaro	86

A bottom-up experience: the DiM project - Angela Mormone, Mariastella Battista, Amelia (Lia) Pensabene	98
Lexicography and language learning of Swahili L2 at UNIOR: the Swahili-Italian online dictionary project - Flavia Aiello, Maddalena Toscano, Rosanna Tramutoli	113
Unveiling oral and writing skills of low-literate learners of L2 Italian: from research to teaching - Marta Maffia	133
PART 3: RELATIONSHIPS BETWEEN ORALITY AND WRITING IN MULTILINGUAL CONTEXTS, LINGUISTIC POLICIES AND CASE STUDIES	145
Farsi language and Iranian immigration in Europe - Adriano Valerio Rossi	146
Lingua e storia in Africa: considerazioni sul caso del tigrino (Eritrea ed Etiopia) - Gianfrancesco Lusini	157
Transcription and orthography in two endangered languages of Ethiopia: Ts'amakko and Ongota - Graziano Savà	174
Wolof language and literature: an introduction - Emiliano Minerba	202
The future of minority languages in Nigeria - Gian Claudio Batic	217
Linguistic ambiguity in Timor-Leste: local languages between pride and shame - Carolina Boldoni	231

Introduction

The spark that gave rise to the first reflections, which were then channelled into this volume, was lit on 7 November 2020, when the CPIA (Centro Provinciale per l'Istruzione degli Adulti), the Provincial Centre for Adult Education of Avellino celebrated the closure of the DiM Project, an international project involving, in addition to Italy, also Greece, Scotland and the islands of Malta and Cyprus.

Carried out in the framework of the *Erasmus plus KA204 - Strategic Partnerships for Adult Education programme*, the DiM (Dizionario Multilingue) had certainly represented a challenge for many actors at the same time, who, in the process of implementing the activities had found themselves having to come to terms with various problems relating to the complexity of the relationships between the linguistic code(s), the culture to which the migrants attending the classes belonged and, more generally, their identity and values, especially in the context of the global village.

In fact, it is well known that in the contemporary world diasporas play a major role in forcing human groups to face each other, people who only a few decades ago would never have had to meet, and it is equally well known that this encounter represents a delicate moment for the balance and the very maintenance of the social fabric where it takes place.

Special attention must therefore be paid to language, or languages, as the privileged instruments of the encounter, since there are aspects imbricated in inter-linguistic and inter-cultural dynamics that go beyond and accompany the mere grammatical dimension and are equally important in building solid bridges between native and non-native communities.

These aspects range from the habit of Western speakers to lead an almost monolingual life and their consequent difficulty in fully grasping what a language code represents for those who were born and live in a strongly multi-ethnic, multi-linguistic and multi-cultural environment, to a whole series of other apparently trivial factors, but which, when considered in their natural habitat, that of lived life and not that of theoretical thinking, turn out to be so intertwined that they cannot be focused on or considered independently and autonomously.

Among these, there are certainly education, the prestige reserved to written languages as privileged instruments for “high” and institutional interactions, the deep feeling of homesickness for one’s mother tongue, often transmitted from one generation to the next only in oral form and spoken by minority groups that already in their homeland do not enjoy any consideration and that during migration end up being diluted and identified simplistically according to their geographical area of origin.

All this is accompanied by the fact that there are often few people in the country of arrival who show the slightest interest in learning more about the subject and its implications, even if only with a view to improving co-existence between citizens and encouraging mutual encounters, exchanges and the integral promotion of people in an environment that is at first sight unknown.

Unravelling these issues is no easy task and involves various scientific approaches, including at least the linguistic, cognitive, psycho-pedagogical, sociological and anthropological ones.

This book, far from offering answers, is intended as a contribution to the currently lively discussions on the theme of intercultural encounter and the teaching of L2 as a means of facilitating it.

The first to enthusiastically join the initiative were those colleagues who had participated as speakers in the closing day of the DiM project on 7 November 2020 and who, in most cases, had followed closely, as consultants and reviewers, the work of compiling and producing the multilingual and multimedia dictionary that can now be consulted free of charge online at www.dimproject.net.

They were then joined by other experts, who are also often personally involved in projects concerning diasporas, minority people and/or orally transmitted and endangered languages.

As the reader will see, the majority of the papers refers to Africa, since the concept note of the work developed from a brainstorming among africanist

friends, but we are sure that all the papers presented will be helpful to all those who are interested in the dynamics connecting language - identity - linguistic policies, above all in multilingual contexts all around the world.

Before starting our trip in the pages of the book, we would like to express our sincere thanks to all the contributors.

DESCRIPTION OF THE CONTENTS

This book is a collective work featuring three different types of contributions.

More theoretical chapters that introduce the reader to some thorny issues such as endangered languages, multilingualism and the eternal dichotomy between orality and writing - Legère on the UNESCO policies and the (ab) use of the (derogatory) term *indigenous* referring to minority languages and peoples; Sharma on Multilingualism in India; Micheli on the dynamics at play in the relations between oral and written languages in Sub-Saharan Africa and Tosco on the provocative issue of graphic systems for African national or minority languages that do not have a consolidated written tradition - alternate with more descriptive chapters focused on specific case studies - Rossi on the Farsi diaspora; Lusini on Tigrinya; Savà on two endangered languages of Southern Ethiopia (Ts'amakko and Ongota); Minerba on Wolof; Batic on the future of minority languages in Nigeria and Boldoni on the situation of local languages in Timor Leste -, all derived from fieldwork experiences.

Finally, in order to offer a glance at the present-day huge movements of people from Low and Middle Income Countries towards the western world, some applicative chapters relate more specifically to L2 teaching and learning, with a focus on inclusive projects for diasporas in Europe, and Italy in particular as the place of arrival of people coming from the Middle East, South East Asia and, mainly, Africa - the DiM project that inspired the whole volume (Mormone, Battista, Pensabene); the papers by Porcaro on the experience of Italian CPIAs; Aiello, Toscano and Tramutoli on the implementation of the UWAZO project, a tool designed to facilitate the learning of Swahili by Italian-speaking students and, last but not least Maffia on teaching techniques for courses aimed at non-literate adults.

The volume is thus organized into three distinct parts reflecting this kaleidoscopic amalgama:

Part 1 contains all the papers related to the more general issues, i. e. multilingualism, minority languages and the dichotomy between orality and writing;

Part 2 is a collection of all the papers focused on lexicography and L2 teaching;

Part 3 gathers all case studies on different linguistic policies adopted (or that we hope can be adopted) in particularly complex multilingual contexts in order to allow the full and correct educational and personal development of speakers of languages not recognised as official and excluded from institutional communication.

As it is evident from the table of contents, most of the essays relate to African case studies, but the editor's wish is that they may also serve as a tool for reflection or as comparative material for other scholars working in similar contexts in other parts of the world.

THE CONTRIBUTORS' PROFILES

The editors and the contributors to this volume are not all from the same background. Many of them are colleagues, lecturers and researchers from Italian and foreign universities, while others are school managers or teachers engaged in the front line of teaching L2 to adult students, often illiterate and sometimes struggling to navigate with ease in the educational system of the host country.

Below are their biographical profiles in alphabetical order according to their family name.

Flavia Aiello is Associate Professor at the University of Naples "L'Orientale"¹ (Department of Asian, African and Mediterranean Studies), where she teaches Swahili Language and Literature at BA and MA level. She is the author of numerous publications on Swahili language and contemporary literature (both oral and written genres, including children's literature) and of translations of Swahili literary works. Between 2015 -2019 she participated in the e-learning project "Progetto formazione a distanza dell'Università degli Studi di Napoli L'Orientale" as responsible and coordinator for the realisation of resources and activities for Swahili language (with M. Toscano, R. Tramutoli and other collaborators), available on the UNIOR Moodle platform (<https://elearning.unior.it/>). She is presently developing, with M. Toscano and R. Tramutoli, a new Swahili-Italian online dictionary, which will be freely accessible from the UNIOR website.

Gian Claudio Batic is Senior Researcher in African languages and literatures at UNIOR. His research focuses on the description and documentation of Chadic languages. He has been visiting scholar at LLACAN-CNRS (Langage, langue et

¹ From now on UNIOR.

cultures d’Afrique - Centre Nationale de la Recherche Scientifique - Villejuif), J. W. Goethe University (Frankfurt am Main), and Bayero University, Kano. He is currently documenting Kushi, a minority language of northeast Nigeria.

Maria Stella Battista has been School Principal of the CPIA Avellino (Provincial Centre for Adult Education) since 2017, previously from 2015 she was school manager of CPIA Avellino-Benevento. She holds a degree in Geological Sciences from the Università degli Studi di Napoli Federico II, awarded in 1989, and subsequently worked as a freelance geologist from 1990 to 2015. Alongside this, she taught mathematics, chemistry, physics and natural sciences in a number of secondary schools from 1996 to 2015. She has attended masters courses on the management function in schools and on the organisation and management of educational institutions in multicultural contexts.

Carolina Boldoni is a Postdoctoral Research Associate of the Universidade Nova de Lisboa, in Lisbon (Portugal). Her research interests are intangible cultural heritage and the national formation of East Timor, where she conducted a 15-months fieldwork between 2017 and 2018. She is currently involved in the formation of a museum in the area of Trafaria, Lisbon, where she is conducting an ethnographic analysis regarding the tangible and intangible memories of the area.

Karsten Legère is Emeritus Professor of African languages, University of Gothenburg, Sweden. He studied African and General Linguistics (M. A. courses) at the Karl Marx University (of Leipzig, then German Democratic Republic) 1964-1968. The same year Sept. 1 he started his academic career as a Lecturer (focus on [Ki-] Swahili and Bantu languages), teaching in Leipzig 1968-1975, 1979-1983 (Senior Lecturer) and 1986-1994 (Associate Professor of Bantu languages), at the University of Dar es Salaam, Kiswahili Dept. 1975-1979 (as Lecturer) and 1983-1986 (Senior Lecturer), University of Namibia, Dept. of African Languages 1994-2000 (Professor), in Gothenburg 2001-Nov. 2010 (Professor, African Languages). After retirement, in Sweden Honorary/ Visiting Professor at the University of Vienna until 2018. Academic degrees PhD. (1974), Dr sc. phil. (1982), Dr phil. habil (1991). He is honorary member of the Linguistic Association of SADC Universities (2001). His research interests include language policy and implementation, sociolinguistics, linguistic landscape, documentation of African languages, description of Kavango languages (Namibia), Tanzanian languages (Bantu, Cushitic, Southern Nilotic), orature and more.

Gianfrancesco Lusini (born 1962, PhD 1991), formerly Assistant (1995) and Associate Professor (2006), since 2019 is Professor for Gə‘əz and Amharic

languages and literatures at the Dipartimento Asia, Africa e Mediterraneo of UNIOR. He is the editor of the journal *Rassegna di Studi Etiopici* and of the series *Studi Africanistici. Serie Etiopica*. “Alexander von Humboldt” Fellow in 2001-2002 (Hamburg University), and Visiting Professor at the Addis Ababa University (since 2014), he is the director of the “Centro di Studi sull’Africa”, and currently he leads the national research project CaNaMEI (Catalogo Nazionale dei manoscritti etiopici in Italia).

Marta Maffia is Junior Assistant Professor (RTD-A) at UNIOR, where she received her PhD in Linguistics in 2015. She has been an L2 Italian teacher in voluntary associations and in reception centres for refugees and asylum seekers. Her main current research interests encompass second language acquisition and second language teaching, particularly in the case of illiterate or low-literate learners; sociolinguistics of migration; prosodic aspects of emotional and pathological speech in L1 and L2.

Ilaria Micheli is Associate Professor in African linguistics at the University of Trieste. Her main interests encompass descriptive linguistics, endangered and orally transmitted languages of sub-Saharan Africa, applied ethnolinguistics and ethnomedicine for cooperation. Among her publications, worthy of attention are the books *Profilo Grammaticale e Vocabolario della Lingua Kulango (Côte d’Ivoire)* (Ed. Università degli Studi di Napoli “L’Orientale” - Dissertations vol. VI - 2007), *Figlio della Radice: Djedoua Yao Kouman - Guaritore e Cacciatore Kulango* (EUT - 2011 and 2017 in English) and *Grammatical Sketch and Short Vocabulary of the Ogiek Language of Mariashoni* (EUT - 2019).

Emiliano Minerba graduated at UNIOR. He is currently working at his PhD research, a comparative historical analysis of Swahili and Wolof metrical systems. Besides Swahili and Wolof prosodies and classical literatures, one of his other research interests is modern Swahili literature, particularly theatre.

Angela Mormone holds a degree in Business Economics from the Università degli Studi di Napoli Federico II, and has enriched her CV with a master degree and advanced courses in the management of complex organisations. Currently working at the Direzione Generale dell’USR (Ufficio Scolastico Regionale) per la Campania (General Directorate of theUSR for Regione Campania), pursuant to Law 448/1998. Angela is the regional contact person for adult education, and is engaged in implementing practices to tackle educational poverty with particular reference to the inclusion of foreign schoolchildren and those with socio-economic difficulties. She has participated in numerous round tables and steering committees at both regional and national level on issues relating to social hardship, unaccompanied foreign minors, adult education, system evaluation and guidance. She coordinates innova-

tion and educational research initiatives for USR Campania, paying particular attention to the implementation and application of project management techniques. She is a trainer and lecturer employed by universities as well as private institutions.

Amelia (Lia) Pensabene is a teacher at the CPIA Avellino where she teaches English at the prison school of the Casa Circondariale in Ariano Irpino and Italian L2 within asylum seekers' host structures. She has previously taught adults in the CTPs (Centri Territoriali Permanenti) both in a prison setting and in evening schools. Lia has a degree in Modern Languages and Literatures (Hindi, Urdu, Sanskrit, English) from the Istituto Universitario Orientale Napoli (IUO)² in Naples, with a thesis on Indian Linguistics. Lia spent several years in India, first in 1989 with a scholarship from the IUO, where she carried out research at the Delhi University and Banaras Hindu University Libraries and field research on the Bhojpuri language. She was subsequently awarded a three-year scholarship by the Indian Government through the Ministry of Foreign Affairs and from 1990 to 1993 she was enrolled at the Deccan College Postgraduate and Research Institute Pune, where she carried out postgraduate research studies in sociolinguistics and lexicography. Her current interests are in language teaching in multilingual and multicultural contexts. She was responsible for the design and overall coordination of the DiM Project and is currently the contact person for CPIA Avellino within the CRRSeS (Centro Regionale di Ricerca, Sperimentazione e Sviluppo sull'Istruzione degli Adulti) for Regione Campania.

Emilio Porcaro, head teacher of the CPIA in Bologna, has devoted the last twenty-five years to various spheres of work ranging from adult education and adult learning to issues relating to teaching Italian to foreigners, to focusing on competency evaluation and certification, and schooling in a prison setting. Since 2012, he has been president of RIDAP (Rete Italiana Educazione degli Adulti), the Italian adult education network. His published works include: *Il riconoscimento dei crediti per l'Istruzione degli Adulti. Metodi, procedure e strumenti* (with R. Sibilio e P. Buonanno, Loescher, Torino, 2020); *Il ruolo del Centro provinciale per l'istruzione degli adulti (CPIA) dentro e fuori dal carcere nella formazione scolastica degli adulti*, in "Adultità fragili, fine pena e percorsi inclusivi. Teorie e pratiche di reinserimento sociale" (edited by Luca Decembrotto), Franco Angeli, Milano, 2020; *Istruzione degli Adulti e religioni*, in "Diritti Doveri Solidarietà" - *Religioni per la cittadinanza*", Garante delle persone sottoposte a misure restrittive della libertà personale, Emilia Romagna, 2019; *Il ruolo dei CPIA nel processo di integrazione lin-*

² IUO (Istituto Universitario Orientale) is the former name of the current UNIOR (Università degli Studi di Napoli "L'Orientale").

guistica dei migranti con un focus sui MSNA in Sguardi Simmetrici, Collana “I quaderni dell’Ufficio Scolastico Regionale per l’Emilia Romagna”, n. 41, February 2018, Tecnodid Editrice); *Minori stranieri non accompagnati* (in Atti del Convegno “Le attuali emergenze pedagogiche: i minori stranieri non accompagnati nelle scuole della Toscana”, Regione Toscana, 2015).

Adriano V. Rossi is Emeritus Professor of Iranian Philology and History of Ancient and Late-Ancient Iran at UNIOR. He studied languages, history and civilizations of the ancient and modern Near East at University of Sapienza (Rome), and directed for decades the Center for Asian Lexicography, (IsIAO - Istituto Italiano per l’Africa e l’Oriente, now IsMEO - Associazione Internazionale di Studi sul Mediterraneo e L’Oriente). He was director of the Department of Asian Studies (1987-1988), Dean of the Faculty of Letters and Philosophy (1990-1992), Pro-Rector (1987-1988) and Rector (1992-1998) at UNIOR. He is a member of scientific councils as well as numerous academies/societies, including Accademia Nazionale dei Lincei and Balochi Academy, Quetta (Honorary). He represented Italy in the Asia-Europe Foundation (1997-2004); since 2016 he is president of IsMEO-The International Association for Mediterranean and Oriental Studies, Rome.

Graziano Savà. Italian (Sicilian) linguist, Graziano Savà is specialised in the documentation and description of minority and endangered languages of Ethiopia. He is the author of the only grammar of the Cushitic endangered language Ts’amakko (Savà 2005) and has documented other little-known languages of Ethiopia such as Ongota, Bayso and Haro. He is also the author of studies concerning historical linguistics and code-switching. He is presently a Postdoctoral Fellow at UNIOR, Italy, carrying a research on the Nilo-Saharan language Nara (Eritrea).

Suhnu Ram Sharma had his MA and PhD in Linguistics from the university of Kurukshetra, (1970, 1975 respectively). Joined Anthropological Survey of India in 1974 as a Research Associate and worked on the tribal languages of Arunachal Pradesh, Andaman Islands, Rajasthan, Gujarat, Madhya Pradesh, Uttarakhand, and Himachal Pradesh. In 1980 Dr Sharma joined Deccan College, Post-graduate and Research Institute, Deemed to be university, Pune as Associate Professor and retired in 2006 as Professor of Tibeto-Burman Linguistics. He has published more than thirty research papers and three books on the description of tribal languages. He taught MA M. Phil courses in sociolinguistics, morphology, semantics and second language teaching methods. He guided 12 PhD students in different subjects related to language descriptions. Dr Sharma has been associated with the Himalayan Languages Project first at Leiden university, Netherlands and now in Bern, Switzerland. He attended several national and international conferences and seminars in

Europe, USA and China. Dr Sharma was a Gonda Fellow at IIAS (International Institute for Asian Studies) Leiden twice during 2005 and 2007 and Senior Fellow at the Himalayan Languages Project in 1997, 1995. Currently he is working towards the completion of Manchad Grammar.

Maddalena Toscano is a former researcher and teacher at UNIOR. She worked in the field of Bantu languages and literatures, mainly dealing with Swahili and Zulu languages, with special attention to applications of computational linguistics to digitalized text analysis. She wrote about Swahili lexicography and Zulu morphological structures. Results of her research activities were also included in her advanced course about the use of application of corpus linguistics in the teaching of Swahili language.

She had experiences in various Socrates projects. She coordinated the ‘CAMEEL’ (Computer Applications to Modern Extra European Languages) project, the results of which were included in the chapter ‘European studies on computing for non-European languages’ (<http://korpus.uib.no/hum-fak/AcoHum/book/NEL-chapter-final.html>) of the ACO-HUM (Advanced Computing in HUmanities) publication. Among the results of these activities is UWAZO, a Swahili-Italian online dictionary, which is now being updated, thanks to the fundings from IsMEO and UNIOR. She was also the main coordinator of the Comenius project ‘TIME for Teachers’ (Tools in Multicultural Education), the final result of which is available in pdf format from the UNIOR Open Archives (<http://opar.unior.it/179>). She was the local coordinator of the ‘EVLang’ project (EVeils aux Langues), which aims at introducing a multicultural and multilingual awareness in the language teaching in primary school.

Mauro Tosco is Professor of African Linguistics at the University of Turin. His main area of research is the Horn of Africa, where he has been working on the analysis and description of Cushitic languages in an areal and typological perspective. Among his books: *A Grammatical Sketch of Dahalo, including texts and a glossary* (Hamburg, 1991), *Tunni: Grammar, Texts and Vocabulary of a Southern Somali Dialect* (Köln, 1997), *The Dhaasanac Language* (Köln, 2001); *A Grammar of Gawwada* and *A Gawwada Dictionary* (Köln, forthcoming).

A native speaker of Piedmontese, an endangered language of NW Italy, he works on the expansion and revitalization of minority languages, language policy and ideology.

Pidgins, creoles and language contact (*Pidgin and Creole Languages: A Basic Introduction*; München, 2001; with Alan S. Kaye) are his third main area of research.

Rosanna Tramutoli is a Postdoctoral Research Fellow at UNIOR. Her research focus is Bantu lexical analysis and a comparison of Swahili and Zulu body terminology. She holds a PhD in Swahili linguistics (2018) from UNIOR and

the University of Bayreuth. From 2013 to 2017 she was Swahili lecturer at the University of Bayreuth. Among her main research interests are cognitive linguistics, anthropological linguistics, semantic analysis, and lexicography. She has published contributions on international journals, such as *RAL (Research in African Literatures)*, *Kervan (International Journal of Afro-Asiatic Studies)*, *Africa and Swahili Forum*. She has recently published a monograph entitled *Encoding Emotions in Swahili. A Cognitive Linguistic Analysis with a Consideration of the Socio-cultural Context* (Köppe, 2020).

Part 1

General issues:
multilingualism, minority
languages and the eternal
dichotomy between
orality and writing

Linguistic identity in and out of Africa

KARSTEN LEGÈRE

Professor Emeritus, University of Gothenburg

ABSTRACT

The paper discusses at least two approaches to determine linguistic identity. In so doing, particular attention is paid to the preparation and implementation of UNESCO's IYIL (International Year of Indigenous Languages) 2019 initiative. As known, given UNESCO's international prestige, institutions and speech communities felt stimulated by this IYIL2019 initiative. As a consequence, their focus was on dealing with those national languages of their countries which in a linguistic hierarchy are not in a top position (like e.g. English, French, Spanish and more), but are rated somehow less important by their speakers or officials. It turned out in the data analysis process for this paper that UNESCO's conceptualization deficits have hampered a productive grassroots response such as evidenced in Namibia. With regard to the development and dissemination of a unified identity concept worldwide a prominent African colleague points out that in Europe, North America, China, in many African countries south of the Equator, etc. own umbrella terms are well established. This implies that UNESCO's identity related activities have not so far much contributed to feasible changes.

KEYWORDS

Identity; grass roots self-identification; top-down identification; African and Scandinavian examples; variety of results; terminological diversity versus UNESCO's global umbrella term.

1. INTRODUCTION - SOME BACKGROUND INFORMATION AND BASICS

On 22 November 2016 the Third Committee (Social, Humanitarian, Cultural) of the UN General Assembly proclaimed IYIL 2019. In this regard, the press release of 8 December 2016¹ announced that the resolution on the “Rights of indigenous peoples” stated the following: “The resolution stresses the urgent need to preserve, promote and revitalize endangered languages [...]”, inviting UNESCO to “serve as the lead agency for the Year”. [Emphasis added]

Shortly thereafter, the same text draws “[...] attention to the critical loss of indigenous languages and the urgent need to preserve, revitalize and promote indigenous languages at the national and international levels”. [Emphasis added]

The above extract shows a terminological contradiction in that endangered languages are equated with *indigenous languages* (henceforth IL) and vice versa. This cannot be explained as a simple oversight, because the UN/UNESCO/ILO, etc. approach to (non-) defining *indigenous peoples* includes also the language of each people. Thus, for assessing language endangerment there are criteria in UNESCO (2003) or *Ethnologue*. This discrepancy is glaring if, for example, one compares the glossonyms in the endangered languages collection at the Max Planck Institute of Psycholinguistics in Nijmegen, The Netherlands,² or of the Endangered Languages Archive at the University of London, with the ethnonyms that are listed e.g. in the IWGIA Yearbook, by Wikipedia, etc.³

It is completely incomprehensible why for the terminological inconsistencies outlined above, in preparation of IYIL2019, UNESCO as a global player was not willing to get down to basics regarding the definition dilemma. Here the definition approach of the Council of Europe traced in the “European Charter for Regional or Minority Languages” is an example that is worthy of being imitated:

“Article 1 – Definitions,

For the purposes of this Charter: a) “regional or minority languages” means languages that are: i traditionally used within a given territory of a State by nationals of that State

¹ http://www.unesco.org/new/en/unesco-liaison-office-in-new-york/about-this-office/single-view/news/united_nations_general_assembly_proclaims_2019_as_the_intern/; accessed 27 January 2019, recent search was unsuccessful, but available (accessed 21 August 2021) is another relevant link, i.e. Permanent Forum on Indigenous Issues, Seventeenth session, New York, 16–27 April 2018, Item 3 of the provisional agenda, Follow-up to the recommendations of the Permanent Forum, Action plan for organizing the 2019 International Year of Indigenous Languages that is a useful overview of the IYIL 2019 focus and content. Link <https://undocs.org/E/C.19/2018/8>

² The results of the Documentation of Endangered Languages (DoBeS) initiative of the Volkswagen Foundation, the DoBeS Archive were accepted by UNESCO as “Memory of the World”.

³ See The International Work Group for Indigenous Affairs (IWGIA), 2020. https://iwg-ia.org/images/yearbook/2020/IWGIA_The_Indigenous_World_2020.pdf as well as https://en.wikipedia.org/wiki/List_of_indigenous_peoples accessed 17 August 2021.

*who form a group numerically smaller than the rest of the State's population, ii different from the official language(s) of that State [...]"*⁴

As an alternative for the English description of 'indigenous' the Merriam-Webster dictionary may be helpful, as pointing out that this lexical item means either:

- 1 a) *produced, growing, living, or occurring natively or naturally in a particular region or environment, or*
- 1 b) *Indigenous or less commonly indigenous, of or relating to the earliest known inhabitants of a place and especially of a place that was colonized by a now-dominant group.*⁵ [Emphasis added]

This Webster entry lists synonyms such as aboriginal, autochthonous, born, domestic, endemic, and native; antonyms are non-indigenous, non-native etc.

A check of UN/UNESCO websites to obtain an up-to-date list of ILs which are conforming to the UN/UNESCO conceptualization has been unsuccessful. There were info bites, such as the following data overview on the IYIL 2019 website:

Indigenous Languages

At present, 96 per cent of the world's approximately 6,700 languages are spoken by only 3 per cent of the world's population. Although indigenous peoples make up less than 6% of the global population, they speak more than 4,000 of the world's languages.

*Conservative estimates suggest that more than half of the world's languages will become extinct by 2100. Other calculations predict that up to 95 per cent of the world's languages may become extinct or seriously endangered by the end of this century. The majority of the languages that are under threat are indigenous languages. It is estimated that one indigenous language dies every two weeks.*⁶

Although having been responsible for the implementation of IYIL2019, UNESCO has been reluctant to submit a list of those ILs the endangerment of which is advanced.

As reported above, the groundwork for the IYIL2019 proclamation was done by Bolivia and Ecuador that also co-authored the associated text. The

⁴ <https://www.coe.int/en/web/conventions/full-list/-/conventions/rms/0900001680695175> accessed 17 August 2021.

⁵ <https://www.merriam-webster.com/dictionary/indigenous>, accessed 17 August 2021. An enquiry about "indigenous languages" is rejected as "The word you've entered isn't in the dictionary".

⁶ <https://www.un.org/development/desa/indigenouspeoples/wp-content/uploads/sites/19/2018/04/Indigenous-Languages.pdf>. [emphasis added]. This means that when annually 26 languages die until 2100, 80 years ahead approximately 2150 ILs (and not 4000) will be lost.

background against which these two countries propagated for a global initiative an umbrella term which is deeply rooted in their colonial past, as well as a critical review of terminological problems and inadequacies not only in the South American neighbourhood will be focused on in the following sections. In this respect, it should be borne in mind that the proclamation text offers various interpretations for its ambiguity and lack of adequate definition, as already pointed out above.

2. REGIONAL LINGUISTIC PROFILES IN THE LIGHT OF IYIL2019

2.1. BOLIVIA AND ECUADOR

Both countries support efforts to protect, empower and promote ILs back home and worldwide. However, in their constitutions, for example, the linguistic terminology involves catchphrases that have nothing to do with IL, because the term indigenous is only traced in the context of the word ‘people’, i.e. *pueblos indígena* ‘indigenous peoples’, which does not refer to any Bolivian or Ecuadorian language. This is illustrated below.

Quoted from Bolivia’s 2009 Constitution: “Son idiomas oficiales ... el castellano y todos los idiomas de las naciones y pueblos indígena originario campesinos ...” [emphasis added].

An exhaustive list of all languages other than Spanish follows. Here, *idiomas*, i.e. ‘languages’ applies to Spanish as well as to all other co-official languages.⁷

From Ecuador’s 2008 Constitution, Article 2: “... el castellano, el kichwa y el shuar son idiomas oficiales ... 1. Los demás idiomas ancestrales son de uso oficial para los pueblos indígenas ...”.

Notable is the catchphrase *idiomas ancestrales* (“ancestral languages”). Furthermore, Article 28 of Ecuador’s *Ley de Educación* (Education Law) speaks of “lenguas aborígenes” (“aboriginal languages”).

It is strange to observe above that, although both Bolivia and Ecuador have been IYIL 2019 protagonists, these countries use modified linguistic umbrella terms for the languages spoken by their populations back home. In view of this discrepancy, for comprehensive country profiles one has to consult *Ethnologue* (Simons & Fennig 2018). *Ethnologue* publishes statistical data on what it defines as IL (i.e. belonging to a specified country) vs. non-indigenous languages (Spanish plus immigrants’ languages). In addition, *Ethnologue*

⁷ A quotation from a Spanish digital dictionary - “We usually say *idioma* when we talk about languages (human tongue) while we use *lenguaje* as language like a programming “language” it would sound awkward to say *el lenguaje Inglés* instead of *el idioma Inglés*...”

(as a reliable reference source for determining the status of a language according to the Expanded Graded Intergenerational Disruption Scale/EGIDS) lists 13 languages in Bolivia that are in trouble and 18 as dying, while eight are in trouble in Ecuador, and three are dying.

Further, it may be interesting to note here how, for example, Bolivian languages and peoples suffered from discrimination under Spanish colonial rule as well as under post-colonial governments after independence from Spain in 1820 (source: Leclerc 2020, Bolivia):

a) *Les langues indigènes furent interdites dans toutes les manifestations officielles de l'État espagnol, mais malgré tout les autochtones résistèrent à la tentative d'assimilation (castillanisation).*

b) *Quant aux populations autochtones, elles furent considérées comme «inférieures» et plus ou moins dépouillées de tous leurs droits civils, politiques, sociaux et linguistiques.*

Moving away from South America, the next region is sub-Saharan Africa with its strong focus on ethnic and linguistic self-identification⁸ away from the colonial legacy.

2.2 SUB-SAHARAN AFRICA

During a review of official language use in Africa the constitutions or other relevant official documents in 23 African countries mostly South of the Equator were reviewed. Below is a summary of the results.

As enshrined in the Constitutions of Benin, Cameroon, DRC, and Gabon among other African countries, all languages of African origin are identified as *national languages* since they belong to a defined nation-state. In the Congo Republic and in Mozambique, the terms *langues nationales véhiculaires* and *línguas veiculares* (in French and Portuguese respectively, the equivalent of *lingua francas*) are stipulated. More umbrella terms are published in Legère (2017), for an update see Legère (2021: 181-182).

Out of the 23 countries reviewed only Kenya and RSA use the term *indigenous* as follows:

In Kenya, referring to all languages other than the official ones (English and Swahili, the latter being also identified as Kenya's national language) the country's 2010 Constitution (Republic of Kenya 2010:14) states in its Article 7(3): "*The State shall (a) promote and protect the diversity of language of the peo-*

⁸ Remember the United Nations Declaration on the Rights of Indigenous Peoples - A/RES/61/295, p. 24, Art. 33; 2.: "Indigenous peoples have the right to determine their own identity or membership in accordance with their customs and traditions." Source: https://www.un.org/development/desa/indigenouspeoples/wp-content/uploads/sites/19/2018/11/UNDRIP_E_web.pdf last accessed 21 August 2021

ple of Kenya; and (b) promote the development and use of indigenous (i.e. Kenyan, K.L.) languages [...]” [emphasis added].⁹

In the South African constitution (= Act 108 of 1996, Founding Provisions, Languages) the text reads as follows: “6. (1) *The official languages of the Republic are Sepedi, Sesotho, Setswana, siSwati, Tshivenda, Xitsonga, Afrikaans, English, isiNdebele, isiXhosa and isiZulu.* (2) *Recognizing the historically diminished use and status of the indigenous languages of our people [...]*

Here “indigenous” clearly means the official African languages of the RSA i.e., belonging to a particular place/country, as described in Webster 1) a).

Further, (5) (a) (ii) refers to South African heritage languages such as Khoi, Nama, and those of the San communities, which are **not** classified as ILs in the RSA constitution, although being related to Webster 1, b).¹⁰

In the 23 African countries studied, the language of the former foreign colonisers/administrators has everywhere been stipulated as the official language, even when it is co-official, i.e. sharing this status with another language, like e.g. Swahili in Kenya, Tanzania, Rwanda and Uganda.

It is strange that in designing and organizing IYIL 2019, UNESCO was not prepared to study or to care about the official terminology that was coined by African governments or other official institutions¹¹ and the resulting linguistic implications at the national level.

Instead, IYIL protagonists made it easy for themselves by putting aside the principle of self-identification (see above), identifying all languages spoken by formerly colonially oppressed peoples worldwide with the umbrella term “indigenous languages”. This kind of strange conceptualization was the focal point that was discussed with senior colleagues and African language experts in 2019/20 against the African background.

The lack of a solid, globally acceptable UNESCO position on the implementation of IYIL 2019 became apparent at the annual meeting of the Namibian UNESCO Commission in Windhoek April 5, 2019 (not recorded by UNESCO Paris in its IYIL2019 event list). At this meeting, the Namibian organizers

⁹ For a comprehensive overview that covers the constitutional dispensations in all African countries, see https://en.wikipedia.org/wiki/List_of_linguistic_rights_in_African_constitutions; accessed 21 August 2021.

¹⁰ Further, in Act No. 6 of 2019: Protection, Promotion, Development and Management of Indigenous Knowledge Act, 2019, Chapter 1, page 8 see definitions of “indigenous community”, “indigenous cultural expression”. “indigenous knowledge” and “indigenous knowledge practitioner” Mind the terminological contradictions which are also evidenced in the SABC recording “Elevating indigenous languages remains a challenge” - <https://m.youtube.com/watch?v=3C9bYYzYdx8>. compared to https://iwgia.org/images/yearbook/2020/IWGIA_The_Indigenous_World_2020.pdf p. 161-PART 1 – Region and country reports – South Africa (web pages last accessed on 21 August 2021).

¹¹ Remember that earlier African governments rejected the term „indigenous peoples“ for its negative connotations in Africa, as it was used in derogatory ways during European colonialism (web sources Legère 2019 4).

came forward with their own interpretation of how to deal with this complex issue. In no way did languages of Namibia's First Peoples, such as those spoken by San communities, play a role. Instead, the invited speakers (divided into grassroots and academic) dealt with problems of African languages in Namibia in general (i.e. the Webster 1a interpretation) in terms of social prestige, language maintenance, promotion and development.

By the way, among the Namibian marginalised San, Ovatie and Ovatjimba the latter speak Otjiherero. Thus, if these communities are identified as "indigenous peoples" by UNESCO and others, the language that Ovatjimba speak is not endangered.¹²

Next, what follows here are some observations regarding the position of the Sámi communities in Sweden and Norway.

2.3 SAMI - AN EUROPEAN MINORITY LANGUAGE GROUP (DIALECT CONTINUUM)

The Sami call themselves *Urfolk* (First People). Being *Urfolk*, the Sámi communities are frequently cited, when the current situation among and the future of First Peoples and minorities in Europe and the Arctic Region are dealt with.

In Norway the legal status of Sámi is defined in "The Sámi Act" (Act of 12 June 1987 No. 56) "§ 1-5. *Sami languages. Sami and Norwegian are languages of equal worth* (emphasis added). *They shall be accorded equal status pursuant to the provisions of Chapter 3*".¹³

During a visit to Karasjok (Norwegian Sámi administrative centre) and its Sámi Park some years ago i.e. the multitude of publications in the Sámi languages was noted with great attention and interest. By comparison, even prominent African languages do not come up by far with such a large number and wide range of publications as those of the Sámi. Another important achievement is the existence of the Sámi University of Applied Sciences (website <https://samas.no/en>) in Kautokeino in Norway which even offers language courses in various Sámi languages. The following Norwegian Sami languages are recognised by the EU authorities which at the right hand side of the table also determine the extent of this recognition.

¹² Even in e.g. Cameroon First peoples have given up their heritage language in favour of that of their neighbours thus speaking a noun class language, the endangerment of which is not known.

¹³ <https://www.regjeringen.no/en/dokumenter/the-sami-act-/id449701/> further Chapter 3. The Sami language. § 3-1. Definitions. etc.; last accessed 21 August 2021.

Lule Sami	Part II (Article 7) and Part III (Articles 8-14)
North Sami	Part II (Article 7) and Part III (Articles 8-14)
South Sami	Part II (Article 7) and Part III (Articles 8-14)

article 7 - recognition and support, article 8 - education and 9 rights, media, etc.¹⁴

As for the Sami languages in Sweden, a Swedish website¹⁵ summarises that

“In 2000, Sami was recognised as an official minority language¹⁶ in Sweden, and the central government has since given the Sami Parliament greater influence and financial resources to preserve the Sami languages, which are rich in variation. Just imagine more than 300 different ways of saying snow – from powder to slush.”

Another Sámi website states that *“Alla samiska språk är klassade som hotade språk [...]”*¹⁷ in so doing accepting UNESCO’s endangered language version as portrayed in the “Atlas of the World’s Languages in Danger of Disappearing“ (1996 > 2001, 3rd edition Mosley 2010). Further, the Sámi website records IYIL 2019 in Swedish as *urfolksprakaret*.

Similar to Norway, Sámi studies and research are important and well established e.g. at the Umeå universitet, Várdduo - Centrum för samisk forskning.¹⁸

The list of Swedish Sami languages and their status as European minority languages is identical with the Norwegian list above.

As a matter of fact, the endangerment process among the Sámi languages is advanced. The estimate of Sámi numbers (i.e. 80.000 persons spread over 4 countries, 20,000 speakers of the most widespread North Sámi) worries the Sami communities.¹⁹ What is highly interesting to note is the community commitment as reflected e. g. in the *Förslag till handlingsprogram för bevarande av de samiska språken*²⁰ which is a comprehensive action plan for language maintenance and empowerment. Its implementation is certainly not easy, in particular at the national level, where Swedish, Norwegian or Finnish

¹⁴ <https://rm.coe.int/languages-covered-en-rev2804/16809e4301> which is a 10 page listing up all recognised minority languages updated on: 28 April 2020, entry Norway, pp. 2-3; accessed 17 august 2021.

¹⁵ <https://sweden.se/society/sami-in-sweden/>.

¹⁶ See also *Lag om nationella minoriteter och minoritetsspråk* (SFS 2009:724): *Samiska är ett officiellt minoritetsspråk i Sverige*. source: <http://www.notisum.se/rnp/sls/sfs/20090724.PDF>, further <https://www.sametinget.se/1079> - web pages last accessed 21 August 2021. Similarly, for Finland see “Sámi Language Act” (1086/2003) Link. Saamen kielilaki.PDF (finlex.fi).

¹⁷ <http://www.samer.se/2739>; last accessed 21 August 2021.

¹⁸ <https://www.umu.se/vardduo-centrum-for-samisk-forskning>; last accessed 21 August 2021.

¹⁹ <https://sweden.se/society/sami-in-sweden>; last accessed 21 August 2021.

²⁰ <https://www.sametinget.se/151550>; last accessed 21 August 2021.

dominate as the national languages with the highest communicative prestige country-wide. It is evident that even if the minority languages are *de jure* equal to other national languages, the distribution and domains where the first-mentioned languages are used contribute to their privileged status. Outside the autonomous Sapmi area in the north of Scandinavia the use of the Saami language is rather restricted. Of course, people can speak Sámi in Helsinki, Oslo or Stockholm, but the number of those who understand this language there is small. The problems the Sámi communities are exposed to were well summarised and assessed in the context of IYIL2019 by Aili Keskitalo, Sametingspresident in Norway (see below), and Lars Miguel Utsi, Deputy Chairman of the Swedish Sametinget.

The Sámi language has been dealt with here quite extensively, because, from a global point of view the linguistic situation in Sápmi and the Sami endangerment process are very typical.

In this respect, judging from the Tanzanian experience,²¹ special attention should be paid to the question - why should someone speak or preserve a language that is compared with other national languages for him/her less prominent in official, formal and even informal domains, as well as less widespread at the country level? As a consequence, the focus of any initiative should be to win the hearts of the young generation who should be prepared to take over the lead in language policy and implementation in due time. Obviously, neither the Sámi, First Peoples of both Americas and in other regions, nor ethnic communities and minorities such as in Africa, Australia or Asia are satisfied to see how their languages are getting lost, because they are no longer used.

At the end of IYIL 2019 the President of the Sami Parliament, Ms. Aili Keskitalo was contacted per email.²² The message drew attention to the fact that IL is a stigma term in larger parts of Africa where it is associated with primitive, second-class or underdeveloped languages. Here is her response:

From: Keskitalo, Aili <aili.keskitalo@samediggi.no >

Sent: Saturday, November 23, 2019 1:33:55 PM

To: Karsten Legère <karsten.legere@african.gu.se >

Re: IYIL 2019

“[...] Sami refer to ourselves as Indigenous,²³ and we do not consider this a derogatory term.²⁴ We are well aware of the International Year of Indigenous Languages, and have engaged substantially in the celebration [...]”

²¹ In Tanzania and elsewhere the older generation esp. in rural areas is deeply concerned about this situation.

²² She was a member of the IYIL2019 Organizing Committee.

²³ Who is non-indigenous in Scandinavia, according to the Sámi understanding?

²⁴ Which it is, even if this is not recognised by Sámi authorities, while people not only in Africa, but also elsewhere avoid it.

As a comment to this message - the question is why do the Sami call themselves *Urfolk*, but reject any self-identification which is similar to that in Canada and the USA as First People or First Nation? Instead, the Sami identify themselves with an English umbrella term that, as shown above in the example of Namibia, may be understood as negating the First People's concept.

As for Sámi and IYIL2019, mainly the Norwegian Sámi communities organised the Arctic Conference, where e.g. the following statement was made regarding priorities:

“Our fundamental tenet is that the Norwegian and Sámi languages shall be considered to be of equal stature and value”.²⁵ This calls for a robust language campaign throughout society, where all players can contribute to promoting the Sámi languages.²⁶

This means, the Sámi authorities both in Norway and in Sweden keep the ball of language maintenance and promotion rolling. However, even the achievements so far are worth to be reported (maybe as a Sami language profile on the UNESCO website).

Identity problems have come up at the end of IYIL 2019, when UNESCO presented an overview of the events during this year 2019. This overview has been analysed with regard to linguistic umbrella terms which were traced in the events list. The results are summarised in the following section.

3. IDENTITY CONCEPTUALIZATION AND ITS TERMINOLOGY

The lexical items below are arranged according to the Merriam-Webster dictionary approach, as outlined in 1. above

a) * national languages, *línguas nacionais*, *nationale Sprachen* (other than the national language, e.g. *Kiswahili*, *Kirundi* or *ikiNyarwanda*) belonging to a particular nation, * minority language, *langue minoritaire*,²⁷ * *lugha za jamii* in Tanzania, Kenya and more;

b) relating to the earliest known inhabitants and their language

* ancestral languages;

* heritage languages,²⁸ treasure language;

²⁵ KL: Has this not been made clear in the Sami Act quoted above?

²⁶ <https://en.iyil2019.org/events/sami-language-conference-iyil2019-launching-in-arctic-region/>; last accessed 21 August 2021.

²⁷ E.g. <https://en.iyil2019.org/events/langues-changements-et-adaptations-ethnographies-et-ecolinguistiques-des-communautés-cotières-nord-européennes-a-laune-du-xxième-siècle/>; last accessed 21 August 2021.

²⁸ LINGUIST List: Vol-32-1080. Wed Mar 24 2021. ISSN: 1069 - 4875 announcing publications in Southern Unami which is “the heritage language of the Delaware Tribe of

- * *urfolkssprak* in Scandinavia;
- * languages of *Orang Asli* ‘First People’ in Malaysia;
- * *lenguas originarias* (e.g. Mexico);²⁹
- * First Nations/First Peoples languages (USA, Canada),
- * First languages (Australia);³⁰
- * *языки коренного первородного населения*, in Russia.

Suffice it to note here again that for UNESCO indigenous languages are also supposed to mean endangered languages, *langues en danger*, ~ languages in danger (of disappearing), since the whole IYIL2019 initiative has strongly argued for the maintenance, promotion and empowerment of small and not so small languages (like the Sami cluster)³¹ that are in the process of being given up.

4. IDENTITY FROM THE GLOBAL PERSPECTIVE - ALTERNATIVE VIEWS

The following are initially comments from a senior African colleague (a participant of the Regional Addis Ababa IYIL2019 conference organized by UNESCO end July 2019) who is concerned about UNESCO’s global imposition of a terminology that e.g. in Africa (but also elsewhere) is controversial and unpopular in particular in former Portuguese colonies:

a) Is the term “indigenous” pejorative? Yes, it certainly is, not only in terms of its origins but in its virtually exclusive use for non-European languages. Even in its practical application as proposed for the IYIL 2019, it is muddled up in its restriction to museum piece endangered languages to the exclusion of many languages which are in need of promotion and empowerment.

b) Can the term “indigenous” be redefined to make it less offensive? For example, can it be made to refer to autochthonous languages found in a given area? If this can be done, all languages including English, French and German will be called indigenous languages in countries where they are natively spoken. I doubt if those who have the messianic mission of rescuing dying languages will agree to relegate their languages to an inferior status.

c) Can the term “indigenous” be abolished? I doubt if this is possible. This is a term that has become part of the international discussion, especially in UNESCO circles.

Indians (Bartlesville, Okla.) and the Delaware Nation of Western Oklahoma (Anadarko)” (emphasis KL).

²⁹ <https://en.iyil2019.org/events/exhibicion-y-venta-de-libros-relacionados-con-las-lenguas-originarias/>, last accessed 21 August 2021.

³⁰ The name of the First Languages Australia organization, see: <https://www.firstlanguages.org.au>; last accessed 21 August 2021.

³¹ Dialects should be added as suggested already in the 2018 Yuelu Proclamation.

d) What can be done in the circumstances? Rejecting the use of “indigenous languages” in the sense of limiting its scope to just endangered languages. In the African context, the lesson to be drawn from this is that initiatives to adjust and reinterpret the term “indigenous” according to situation and reality may well be the way out.

This is a perfect summary that, together with the fundamental Robillard/Bahuchet (2012) discussion of the terminological escapades that the authors have traced among Central African forest dwellers should be highly relevant for UNESCO and other institutions which believe that there is no alternative to the global umbrella term “indigenous languages”.

Based mainly on the review of UNESCO documents and other sources that deemed to be important for the discussion of terminological issues related to IYIL 2019 and beyond the following is suggested here:

- It is time to stop the supremacist interference of organizations, institutions, NGO’s, ‘expats’ and persons from outside who feel being authorized to tell others - communities, ethnic groups, individuals - who they are.
- African States which have been studied in the context of identity matters should be encouraged to pursue their will and way of deciding by themselves what is appropriate (or not) in the process of linguistic self-identification.

It should be borne in mind that in view of the colonial past or racist connotations, quite recently terminological changes that are supportive to a constructive identity approach were made at the national/regional level as follows:

- E-word being replaced by “Inuit”;
- P-word in Central Africa being banned by officials;³²
- partial self-identification as “First Peoples/Nations” instead of “Indian”;
- in Germany/Austria and Switzerland dropping words like *Mohr*, *Eingeborene* (already many years before *Eingeborenensprachen*), but Duden and some media still maintain stigma terms like “indigene Sprachen” und “indigene Menschen/Indigene”³³ for denoting languages and peoples in Africa, Latin America, USA, Canada, Asia, Australia (on top of Aborigines).

³² Robillard&Bahuchet (2012).

³³ A typical example is the way how in German media D. Jur. Deb Haaland’s ethnic origin was described, such as *Indigene wird Minister*, ‘Indigenous becomes Minister’, also Wikipedia’s https://de.wikipedia.org/wiki/Deb_Haaland_-_erste_indigene_US-Ministerin ‘First indigenous Minister’, whereas the English Wikipedia version writes that the Secretary of the Interior is “[...] an enrolled member of the Laguna Pueblo [...]” <https://www.doi.gov/secretary-deb-haaland>: “[...] a member of the Pueblo of Laguna [...]” - links last accessed 21 August 2021.

The rejection of the stigma terms above is a solid argument against UNESCO's reluctance of dropping its favorite *indigenous languages*. Similarly, after World War 2 the French colonial stigma term *peuples indigènes, langues indigènes* was replaced by *peuples/langues autochthones*.³⁴ Also the N-word had been replaced a long time ago by *Afroamerican*.³⁵ Another terminological re-orientations happened in Mozambique, Angola, Cap Verde and Guinea Bissau, where the stigma term *línguas indígenas* was rejected being replaced with *línguas nacionais*. In addition, as earlier summarised - out of 23 African countries, only two have included the expression *indigenous languages* in official documents, all others have stipulated other umbrella terms.

Above the question was asked whether it is acceptable to call (the national languages) English, French, German (also Spanish, Italian, Russian, Hungarian, Polish, etc.) *indigenous languages*, subsequently being answered - of course, not...

5. RÉSUMÉ AND OUTLOOK

This paper deals with issues related to linguistic identity. The starting point is the ambiguous use of the keyword *indigenous language* stimulated by UNESCO and its IYIL2019 profile. In this context, reference is made to two aspects, namely, on the one hand the self-identification by the language community on the grassroots level or likewise by qualified institutions on the national level. On the other hand, a linguistic identification is made from outside, the result of which is in contradiction to the national, regional or grassroots identification. Both cases produce a conflicting linguistic identity, since the second approach doesn't care for an identity which goes back to the linguistic self-identification process.

A number of examples from different countries (with a focus on South America, 23 African countries, Scandinavia) demonstrate how differently this global IYIL2019 initiative has been implemented in view of UNESCO's reluctance to define the exact profile of IYIL2019. Given UNESCO's international prestige, little opposition to the former authoritarian way, that has imposed a strange identity on a large group of languages worldwide, has been observed.

This is problematic, because UNESCO is now planning a Decade of Indigenous Languages, which, as a consequence, gives rise to a renewed divi-

³⁴ See also Bouchareb, Rachid. 2006.

³⁵ But mind the use of its diminutive in the Philippines which is as pejorative as the N-word. BBC reports that the N-word has also been used by white actors in the Australian TV serie Neighbours, <https://www.bbc.co.uk/news/world-australia-56652191>; last accessed 21 August 2021.

sion of the world into regions with languages the identity of which is on the one hand, determined from outside.

Accordingly, in Asia, Africa, Latin America and Australia/Oceania, as in the times of colonialism, there are indigenous languages. In contrast, Europe and North America are excluded from a global initiative to promote small languages and dialects, because there are hardly any languages whose speakers accept an identity imposed by UNESCO.

REFERENCES

- ACHPR/African Commission on Human and Peoples' Rights. 2006. *Indigenous Peoples in Africa: The Forgotten Peoples? The African Commission's work on indigenous peoples in Africa*. Banjul and Copenhagen: ACHPR and IWGIA (International Work Group for Indigenous Affairs).
- Calvet, Louis-Jean. 1978. *Die Sprachenfresser: Ein Versuch über Linguistik und Kolonialismus*. Berlin: Arsenal (German translation of Calvet's 1974 book in French linguistique et colonialism, petit traité de glottophagie. Paris: Edition Payot).
- CHAGS 12. 2018. Programme. Penang: School of Social Sciences.
- Economist (The). 2018, 7 July. "Aboriginal rights: 230 years later", pp. 45-46.
- Legère, Karsten. 2017. "Empowering national languages: The Bantu language area" in: Raija Kramer & Roland Kießling (eds.) *Mechthildian Approaches to Afrikanistik*, pp. 243-260. Köln: Köppe Publ.
- Legère, Karsten. 2021. "Official and co-official languages in sub-Saharan Africa: What about other languages?" in: Anderson Chebanne & Amani Lusekelo (eds.) 2021. *African Languages. Linguistic, Literary and Social Issues. A Festschrift in honour of Prof. Herman Batibo (CASAS BOOK SERIES*. No. 131), pp. 177-193 (Chapter 11). Cape Town: UWC/CASAS.
- Legère, Karsten, Richard Trewby & Mariana van Graan. 2001. *The implementation of the Namibian language policy in education: Lower primary grades and pre-service teacher education. An ERNESA [Educational Research Network of Eastern and Southern Africa] Report*. Windhoek: Namibia Educational Research Association.
- Newitt, Malyn. 1995. *A history of Mozambique*. Bloomington/Indianapolis: Indiana University Press.
- Robillard, Marine & Bahuchet, Serge (2012). "Les Pygmées et les autres: terminologie, catégorisation et politique" in: *Journal des africanistes*. 82 (1/2): 15-51.
- Skutnabb-Kangas, Tove. & Robert Phillipson (editors). 1995. *Linguistic Human Rights. Overcoming Linguistic Discrimination*. Berlin, New York: de Gruyter.
- Skutnabb-Kangas, Tove. 2000. *Linguistic genocide in education or world-wide diversity and human rights?* Mahwah, NJ: Lawrence Erlbaum associate

Web resources

- Bouchareb, Rachid (Director). 2006. Days of Glory (Indigènes). accessed <https://www.imdb.com/title/tt0444182/>; last accessed 17 August 2021.
- CIA/Central Intelligence Agency. 2019. World Factbook. Available at <https://www.cia.gov/the-world-factbook/>; last accessed 21 August 2021.
- Duden internet version, entry *Eingeborene*. <https://www.duden.de/rechtschreibung/> accessed 21 August 2021.
- Council of Europe. 1992. *European Charter for Regional or Minority Languages*. Strasbourg: CoE. <https://www.coe.int/en/web/conventions/full-list/-/conventions/rms/0900001680695175>, accessed 21 August 2021.
- ILO & ACHPR/International Labour Organization and African Commission on Human and Peoples' Rights. 2009. Overview Report of the Research Project by the International Labour Organization and the African Commission on Human and Peoples' Rights on the constitutional and legislative protection of the rights of indigenous peoples in 24 African countries. Geneva: ILO. Available at https://www.ilo.org/global/topics/indigenous-tribal/publications/WCMS_115929/lang-en/index.htm; last accessed 21 August 2021.
- IWGIA (= International Work Group for Indigenous Affairs). 2018. 370 million indigenous people. <https://www.iwgia.org/en/indigenous-world> last accessed 21 August 2021.¹
- Leclerc, Jacques. 2020. *L'aménagement linguistique dans le monde*. Ottawa: Laval University <http://www.axl.cefan.ulaval.ca/afrique/afraacc.htm> last accessed 21 August 2021).
- Legère, Karsten 2002. "The Languages of Tanzania project: Background, sources and problems" in: *Africa & Asia*, 2: 163-186. https://www.ipd.gu.se/digitalAssets/1324/1324043_the-languages-of-tanzania-project.pdf (for this paper) accessed 17 August 2021;
- Legère, Karsten 2019. OL > "Open letter to the Secretary General of the United Nations, 1 February 2019", available at <https://osf.io/a6n7p/quickfiles>.
- Legère, Karsten. 2020. "Deutsch als „indigene Sprache“?" in: Sprachnachrichten. Das weltweite Netz der deutschen Sprache 87 (III), 15. https://www.dialog-mb.com/files/2020_Sprachnachrichten_2020-03_Sprache_der_Gewalt.pdf.
- Legère, Karsten 2021. "IYIL and a forthcoming decade", unpublished document, <https://osf.io/h83fc/>.
- Marks, Don. 2014. "What's in a name: Indian, Native, Aboriginal or Indigenous?" *CBC News*, 2 October 2014. Available at <https://www.cbc.ca/news/canada/manitoba/what-s-in-a-name-indian-native-aboriginal-or-indigenous-1.2784518>; last accessed 21 August 2021.
- Merriam-Webster's Dictionary online. <https://www.merriam-webster.com/dictionary/indigenous#synonyms>, last accessed 17 August 2021.
- Republic of Kenya. 2010. *Constitution of Kenya, 2010*. Nairobi: The National Council for Law Reporting. <http://extwprlegs1.fao.org/docs/pdf/ken127322.pdf>; last accessed 21 August 2021.
- Simons, Gary F & Charles D Fennig (eds.). 2018. *Ethnologue: Languages of the World*, Twenty-first edition. Dallas, TX: SIL International. Available at <http://www>.

¹ Biased reports such as the Tanzanian one in: <https://www.iwgia.org/en/tanzania> (last accessed 21 August 2021), which claims "Tanzania does not recognize the existence of indigenous peoples", not aware at all that there is "Sera ya utamaduni" 'Cultural policy' which focuses on the existence of many national languages other than Swahili spoken by ethnic communities.

- ethnologue.com; last accessed 21 August 2021.
- United Nations, General Assembly. Human Rights Council. 2012. *Expert Mechanism on the Rights of Indigenous Peoples: Study on the role of languages and culture in the promotion and protection of the rights and identity of indigenous peoples* <https://www.un.org/esa/socdev/unpfii/documents/2016/egm/Expert-mechanism-language.pdf>, last accessed 21 August 2021.
- UN Human Rights. 2012. *The Role of Languages and Culture in the Promotion and Protection of the Rights and Identity of Indigenous Peoples*. <https://www.ohchr.org/en/issues/ipeoples/emrip/pages/studylanguages.aspx> last accessed 21 August 2021.
- UNESCO. 2003. *Language Vitality and Endangerment. Document submitted to the International Expert Meeting on UNESCO Programme Safeguarding of Endangered Languages* Paris, 10–12 March 2003 http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CLT/pdf/Language_vitality_and_endangerment_EN.pdf, last accessed 21 August 2021.
- UNESCO. 2019. *UNESCO Atlas of the World's Languages in Danger*. Paris: UNESCO. <http://www.unesco.org/languages-atlas/> accessed 21 August 2021.

Wikipedia. 2021. *List of linguistic rights in African constitutions*; https://en.wikipedia.org/wiki/List_of_linguistic_rights_in_African_constitutions; accessed 21 August 2021.

Multilingualism in India and the significance of multilingual dictionaries

SUHNU RAM SHARMA

(retired) Deccan College, deemed to be University Pune

ABSTRACT

The Indian subcontinent has four major language families, e.g., Indo-Aryan, Dravidian, Austroasiatic (Munda) and Tibeto-Burman. These language families have co-existed together for many centuries and have resulted in widespread multilingualism. The multiplicity of languages does not impede communication as large populations habitually learn more than one language for daily communicative needs. With the increase in the movement of people in recent years for business, travel and educational needs, there is an urgent demand to teach and learn more languages. India, with more than 22 major official languages, requires the use of multilingual dictionaries. It will help in translation work that would be required at a large scale. With the help of modern technology it is now possible to make use of multilingual dictionaries for many purposes in daily life as we make use of the spoken knowledge of different languages in our daily life. The present paper will attempt to provide some facets of multilingualism in India and the need for multilingual dictionaries.

KEYWORDS

Bilingualism/ multilingualism; Language/ dialect; communicative networks.

The Indian sub-continent has been the home of diverse language families for many centuries. It may be due to the movement of different populations at different points of time over some millennia. Recently, based on some archaeological findings, it has been hypothesised that the diverse populations are indigenous to this sub-continent that may have thrived over many centuries one after the other. Therefore, India has four major families of languages e.g., Indo-Aryan, Dravidian, Austroasiatic (Munda sub-family) and Tibeto-Burman. There are two minor language families e.g., Tai Khamti and the Andaman group of languages, like Great Andamanese, Jarwa, Onge and Sentinelese.

In addition to this diversity we have isolate languages like Nahali, (Central India) Burushaski (Hunza Valley, Pakistan) and Kusunda (Nepal) that belong to no other language families in the world. In order to account for this diversity we have to imagine that these distinct language groups must have been language families in their own right but were reduced to small isolates due to the dominant populations that thrived in Indian sub-continent. Otherwise it is difficult to imagine how these small populations belong to none of the families of languages that we know today in the world. Maybe their roots were here back in time and their populations decreased over a period of time for some reasons.

It has been well documented that the Indo-Aryan, Dravidian and Austroasiatic (Munda sub-family) and Tibeto-Burman families of languages must have had sustained and long term contact that has resulted in what Emeneau, M.B. (1956) called 'India as a linguistic area', where some linguistic features spread cutting across the genetic boundaries of language families all over in different degrees. Needless to say that long term contact must have resulted in sustained and stable bilingualism leading to borrowing and the formation of linguistic area. Dravidian languages like Malayalam, Telugu, Kannada and Tamil have a good number of Sanskrit borrowings. Even the existence of retroflex sounds in Indo-Aryan, Munda and some Tibeto-Burman languages is ascribed to Dravidian contact. The SOV word order, a pan Indian norm, must have resulted due to long term contact. This clearly shows that language families are not pigeon holes or islands but are deeply interconnected with widespread multilingualism for centuries. The linguistic communities in India habitually acquire and use languages other than their own for immediate oral communication. Multilingualism does not mean that people have good command over all the skills, like speaking, listening, reading and writing of more than two languages; but they have some competence to fulfil the communicative needs required. One native language may be used at home among the family members, the other dominant language may be used outside the home at work place and still another used at school as the medium of education. This is true of many small communities whose languages have no writing system developed so far. For example, a Byangsi speaker in

Uttarakhand is using his/her mother tongue at home and the dominant language Kumauni or Nepali just outside the home at market place, then Hindi and English at school as languages of education. Same is the case with many other tribal languages in North-East India and some parts of central India.

In modern times, the movement of populations has increased within India and abroad for education, work and business. This accounts for learning of other languages for daily use for different purposes, like education, business, entertainment and other electronic media like films and television, Facebook, WhatsApp, Twitter and others. This means that the populations are interacting in different languages and need translation and the use of multilingual dictionaries for daily use.

The census of India has recorded a steady increase in bilinguals from 9.7 percent in 1961 to 24.79 percent in the 2001 census. (Bhattacharya S.S., 2018) Pandit, P.B. (1972) pointed out that in India bilingualism is a norm and monolingualism is an exception. This means that basically multilingualism is widespread and very much required for daily communicative needs of the Indian populations. Before I explain the bilingual situation in the Indian context it will be worthwhile to provide some ideal definitions of bilingualism. Here, the terms bilingualism and multilingualism are used interchangeably where they mean the use of more than one language in various social contexts. In some locations there may be two languages and in other places there may be three to four languages/dialects. The terms, languages and dialects are also not used in a well-defined way and are used interchangeably in the present paper.

Bloomfield (1933:56) defines bilingualism as “native-like control of two languages” —a very high goal indeed. For Haugen (1953:6) a bilingual produces “complete and meaningful utterances in other languages.” Weinreich (1953:1) considers bilingualism “the practice of alternately using two languages.”

There are many other definitions of bilingualism that have been provided by the scholars depending on the age of acquisition of second language (Butler, 2013), context of acquisition (Hoffmann,1991), order and consequent acquisition (Hoffmann 1991), cognitive organisation (Hoffmann 1991), relative competence (Hamers & Blanc, 2000), functional ability (Butler, 2013), exogeneity (Hamers & Blanc, 2000), cultural identity (Hamers & Blanc, 2000) and social and cultural status of the languages (Hamers & Blanc, 2000). But for Indian contexts we need to redefine bilingualism.

First of all, in most social contexts, where oral communication takes place, more than one language may be used in particular situations like home, outside home, in the market place, in schools and colleges where the language and medium of education may be different. Therefore, the functional ability is much more important where the second or third language may be used in defined contexts. Moreover, the degree of bilingualism may differ from one

state to another. In a given situation one may use one language at home and the other in the market place just to buy goods and complete the transaction. S/he may not have full command of the second language with regard to skills like speaking, listening, reading and writing. So we can say that some minimal oral competence that is required in a given context of language use may be the ideal definition in Indian bilingualism rather than the definitions given above. Sure enough, some members of a community may have equal and effective command of two or more languages. So Indian bilingualism may be defined in one way or the other depending on the regions and social contexts. Some speakers may have equal and effective command over all the skills in more than one language whereas, in other cases they may master minimal functional abilities to carry out the communication in given social contexts. Oral communication networks are much more predominant in Indian bilingualism. For example, a native language like Byangsi (Uttarakhand) is used in most intimate and home domains, Nepali and Kumauni in outside home with friends and Hindi and English in school domains.

While looking at the global scene we find that many small languages are shifted in favour of dominant languages. But in the Indian sub-continent most languages are maintained even in the face of dominant languages. It is simply a matter of allocation of domains of language use for different purposes at different places without giving up one's language in the face of the other languages. Therefore, multilingualism helps maintaining languages over the generations. Moreover, for many small tribal populations the native language serves as a feature of their ethnic identity; therefore, they can't afford to give up their identity by adopting a different identity based on a different language. This is particularly true of Tibeto-Burman languages of North-East India like, Mizo, Ao, Thadou, Angami, Lota, Wanchoo, Nockte, Tangsa, Adi, Mishmi, Apatani, Tagin, Galo, Monpa, Manipuri, Bodo, Garo and many others. However, there are still some languages with less than ten thousand speakers with no writing system developed. These communities have shown a steady decrease in their number of speakers over the generations. But with modern technology the folk literature of these languages are being documented. In western Himalayas, where I have carried out extensive research on the Tibeto-Burman languages, I found a renewed interest among the people to write something in their languages. These languages are Byangsi, Rongpo, Manchad, (Patani) Kinnauri and a few others. After the spread of education they are becoming increasingly conscious of their languages.

Soon after the independence in 1947 India adopted a multilingual language policy by placing major languages in the Eighth Schedule of the Indian Constitution.

Presently the Eighth Schedule of the Indian Constitution recognises 22 major languages, e.g., Assamese, Bengali, Gujarati, Hindi, Kannada, Kashmiri, Konkani, Malayalam, Manipuri, Marathi, Nepali, Oriya, Punjabi, Sanskrit,

Sindhi, Tamil, Telugu, Urdu, Bodo, Santhali, Maithili and Dogri to be used at State level for all purposes. In addition to these languages, Hindi and English are used for interstate and international official communications. Over the years more and more languages are being encouraged to be used in school education and other official use.

There are 121 languages spoken by more than ten thousand people in India according to the Census of India 2011, out of which 22 are Scheduled languages and 99 are placed under Non-Scheduled languages. However, the census of India further adds that there are 270 identifiable mother tongues. Many scholars wonder why illiteracy is not removed from India? This is due to the fact that large populations speak very different languages that are not the medium of education. Many languages don't even have writing systems. For example, a speaker of Manchad in Himachal Pradesh or an Ao Naga speaker in Nagaland has to face Hindi and/or English at school in order to be literate. These speakers also have to face a dominant language of the area as well. Therefore, there will be various levels of literacy depending on the language to be used in education. Hence, many tribal populations end up learning more than three to four languages simultaneously for their daily communicative needs. However, they may not be literate in any of the languages. Therefore, literacy is not a measure of bilingualism as such. But large populations are bilingual or trilingual for their communication needs. Literacy is defined as the knowledge of the written form of a given language which may not be the mother tongue of some tribal populations. But their communicative needs are met with the help of multilingualism.

In the linguistic history of India, Persian, Arabic and English were also added to the Indian communication network. That is the reason why all Indian languages borrowed extensively from these languages.

Today English is mostly used for higher education, science and technology and other international communication. The use of English has even increased since independence. It is assumed that an average educated Indian makes use of two or more languages in various walks of life. Mother tongue is used in the family and with friends, whereas the second language may be the medium of education; still the third language, like English, is added at school. In India when some populations move from one part of the country to another they learn the dominant language of the place even without formal learning. For example, Gujarati and Rajasthani shopkeepers settled in South India learn Tamil, Malayalam, Kannada or Telugu, whereas people from South living in North India learn Hindi. This apart in various state borders like Gujarat and Maharashtra, people command both Marathi and Gujarati; between Maharashtra and Karnataka, people learn both Marathi and Kannada side by side. The same type of bilingualism is found at all state boundaries that connect the populations with bilingual networks. This includes Hindi-Panjabi, Hindi-Marathi, Kannada-Tamil, Tamil-Malayalam,

Marathi-Telugu, Gujarati-Rajasthani, Assamese-Bangali, Bengali-Odia, Odia-Telugu and so on. The entire country is connected with bilingual networks of one type or the other.

Hindi is widely used in popular TV programs, films and other social media network communication and also functions as the medium of communication at informal level almost all over India.

The story of multilingualism becomes all the more interesting when we come to the tribal populations. They command three to four languages simultaneously.

For example Santhali-- a Munda language - at home; Bengali or Hindi, or Oriya outside home. At school, English and/or Hindi are added for education.

The same is true for many speakers of Tibeto-Burman languages spoken in North- East and North-West of India where native languages, plus regional languages, plus Hindi and/or English are further added to the discourse. One of the outstanding features of Indian multilingualism is the fact that, even if a person is somewhat deficient in speaking his/her second/third language(s), s/he is encouraged to speak it rather than discouraged. No one cares for the so-called standard prescribed or correct form of the language, as the immediate aim is to put one's idea across, that is, oral communication. The communicating oral skills in various domains of language use are considered more important, without caring for any standard or written mode of communication. That is why, many people don't feel the urgency to learn written form of some languages since the communication goes on for their daily needs. This kind of communication takes place in informal situations rather than formal occasions. It is a very interesting feature of Indian multilingualism that when there is an official directive to introduce some major language for official purpose, there is a resistance to learn/adopt that particular language, however, in given informal situations people learn and use the dominant language that one may confront. This means that the Indian people in general may resist learning a particular language if imposed, but would not mind learning a language that may be needed in certain domains of language use. It is a matter of freedom, choice and need that is more important. It is the linguistic diversity that holds rather than the uniformity of languages and culture.

With the increase in travel and business among the different states and populations, India needs more and more translations and formal learning and teaching of different languages. This means multilingual dictionaries will be very useful for a multilingual country like India, where more and more languages are being taught and learnt on a daily basis. It is important to note that most Indian populations don't give up their mother tongues in the face of dominant languages, as different languages are learnt and used in different domains of usage. That does not mean that small languages are not endangered as such. Some languages without any writing systems are likely to be lost over the generations and their use is highly restricted. Moreover, they are

moving from rural settings to urban settings where different languages are spoken and used for educational purposes.

Finally, with increasing social mobility, education and literacy in a multilingual country like India, more and more populations are moving from monolingual to multilingual communities. Therefore, the multilingual dictionaries will be required for multiple purposes, like translations, online learning and teaching languages and for other electronic media like blogs, twitter, facebook and WhatsApp usages. Over the years, more and more languages have been recognised and included in the Eighth Schedule of the Indian Constitution. The number of languages that are being written and used on a daily basis are on the rise due to linguistic awareness. The increase in multilingualism is also helping to maintain endangered languages as educational resources are being upgraded and updated with the growing communicative needs of the people.

REFERENCES

- Bhattacharya, S. S. 2018. *Languages, Mother Tongues and Bilingualism in India: An Overview*. A paper presented at the Workshop on Field Linguistics, Department of English and Foreign languages, Tezpur University, Assam. India.
- Bloomfield, L. 1933. *Language*. New York: Holt Rinehart and Winston. (published in 1935 by George Allen and Unwin, London).
- Butler, Y. G. 2013. "Bilingualism/ Multilingualism and Second-Language Acquisition". In Bhatia, T. K. & Ritchie, W.C. (Ed.), *The Handbook of Bilingualism and Multilingualism: Second Edition*, UK: Blackwell Publishing, pp: 109-136.
- Census of India 2011 *Paper 1 of 2018. Language. India, States and Union Territories (Table C-16)*. Office of the Registrar General, India 2A Man Singh Road, New Delhi-110 001. India.
- Emeneau, M. B. 1956. "India as a Linguistic Area" in: *Language* Vol. 32, No. 1 (Jan. - Mar., 1956), pp. 3-16.
- Hamers J. F. & Blanc M. H. A. 2000. *Bilinguality and Bilingualism* (2nd edition). Cambridge: Cambridge University Press.
- Haugen, E. 1953. *The Norwegian language in America: A study in bilingual behavior*. Philadelphia: Univ. of Pennsylvania Press.
- Hoffmann, C. 1991. *An Introduction to Bilingualism*. London: Longman.
- Pandit, P. B. 1972. *India as a Sociolinguistic Area, Professor P. D. Gune memorial lectures*; 3. Poona: Poona University, India.
- Weinreich, U. 1968. *Languages in Contact*. The Hague: Mouton. (first published in 1953).

The natural supremacy of spoken language. Orality and writing in Africa

ILARIA MICHELI
University of Trieste

ABSTRACT

This contribution aims to introduce the reader to the African socio-linguistic panorama and to the main issues bound to the often oversimplified dichotomy which sees, in abstract terms, verbal communication as opposed to written production. In the present article I will try to bring to light all different aspects which contribute to make the question a very complex and multilayered one, at least in Africa, above all because the vast majority of African languages to date do not even have a written form.

In such conditions, the issue is not whether, when or how, writing is used or allowed, but why the lack of adequate written varieties is so pervasive and whether and how things could be possibly changed for the better, in order to really value, protect and promote the use of minority and endangered African spoken languages.

In brief, I will discuss the following 5 points: 1) the number of living languages in the continent (and the difference between pluri- and multilingualism); 2) the presence of indigenous African scripts and imported alphabets from the 3rd millennium b.C. to present day; 3) the supremacy of orality over writing with specific references to traditional literature and artistic production; 4) the unfair distribution of literacy and formal education in the continent; 5) the importance of linguistic policies for the right to education and the protection of endangered minority languages and communities.

KEYWORDS

African languages; Literacy; Writing systems; Endangered languages; Language policies.

“Asking a community to choose which of its languages should receive institutional support is like asking a mother to choose which of her children should be given new clothes”

Jufferman & Abdelhay 2016.

INTRODUCTION

Far from being an exhaustive dissertation about the oral and written varieties of the more than 2000 different living African languages, which would, indeed, represent a too ambitious endeavor even for linguists much more capable than me, my simple goal in this paper is to outline at least the main features of communication and use of languages in the continent, as well as the characteristics of their appearance in verbal and/or written contexts, or the position reserved to minority, spoken languages in present day Sub-Saharan African school systems.

In the following paragraphs I will try my best to shed some light at least on five points which could help the reader to grasp the underlying complexity of the whole question of language *at large* in a continent which is everything but a monolith.

Thus, § 1 will be devoted to a brief commentary on the numbers registered for African languages and on the effective usability of data derived from a simple enumeration of languages and linguistic families in the continent.

§ 2 will offer a simple overview (with some personal reflections) on the dynamics underlying the development of scripts for local languages, considering that, far from being a continent without writing, Africa has been home of some of the most well known ancient scripts, which, in some cases (e.g. the Egyptian hieroglyph) date back even to the 3rd millennium b.C.¹.

In § 3 I will try to discuss what I mean when I speak of the supremacy of spoken over written language, making specific references to the African traditional oral literature(s) and artistic production(s).

§ 4 will be devoted to a brief discussion of the unfair distribution of literacy in the continent, while in § 5 I will focus on the importance of linguistic policies for the right to education and the protection of endangered minority languages and communities.

¹ This paragraph could be seen also as an introduction to the chapter by Mauro Tosco in this volume.

1. AFRICAN LANGUAGES: NOT A MERE QUESTION OF NUMBERS

Africa is home to 54 different countries, to hundreds of different cultures and to thousands of different indigenous languages. According to the most recent estimates of Ethnologue, in the world, around 7.000 languages are currently spoken, and Africa, together with South East Asia represents the region with the highest variety. The continent hosts, in fact, more than 2000 living languages. This means, in other words, that almost 1/3 of the world's languages are African.

The presence or absence of a writing system is not a *condicio sine qua non* for the identification of a variety as a language or a dialect, which is indeed a malicious belief we usually inherit from obsolete Western school memories. In order to be defined a language, a verbal variety must have a peculiar structure, with peculiar phonology, morphology, syntax as well as specific vocabulary and it must represent a *pivot* around which other slightly different but inter-comprehensible varieties (dialects) turn around.

Unfortunately, despite some noticeable exceptions (e.g. the Bantu family), we lack reliable data and complete descriptive materials on most African languages and their varieties².

However, it can be accepted as a given that this is a feature that Africa shares with all those areas of the world which were subjected to colonialism (Latin America and South East Asia in the first place), or that have not been so much influenced by the Western education systems or by the true passion for description and scientific classification that the Darwinian experience boosted all around the world since the era of Enlightenment.

However, the complexity of the linguistic mosaic of the continent is not only due to the huge numbers of languages currently spoken. Understanding the numbers, in fact, does not simply mean to understand their geographic distribution, fancying that each language/community of speakers occupies a specific area, which is in itself monolingual, monocultural, and mono-ethnic. On the contrary, an attentive reading of these numbers tells us that the African reality is made, also and above all, of something much more complex, and that the cohabitation of multilingual, multicultural and multiethnic groups in the very same area, often creates highly variable socio-linguistic networks and dynamics that are very hardly perceivable by outsiders (see Batibo 2005, Lüpke 2010, Jufferman & Abdelhay 2016).

Researches on plurilingualism in themselves, referring only to the numerical dimension, can tell us much on the problems that African modern States should face, if and when they seriously intend to tackle the delicate issue of language policies, since, except for a very few cases in Mediterranean Africa or in the Horn (Somalia) where the number of languages spoken is

² cf Micheli 2010.

very low, usually the density degree of linguistic diversity is uncomfortably high, above all in the equatorial sub-saharan regions, where in a country like Côte d'Ivoire, which has the same extension of Italy, more than 70 different languages belonging to 4 different families are spoken or, in a country like Nigeria, which is three times larger than Italy, there are more than 500 languages. However, numbers cannot tell us anything about how, when and by whom all these languages are really spoken.

Researches on multilingualism, on the contrary, being focused on the speakers' behaviour and attitudes rather than on numbers, tell us more about the status of the different languages that co-habit the same geographical area.

According to Batibo 2005, the more typical socio-linguistic model in the continent is represented by triglossia. An average African speaker has in fact in his/her linguistic toolbox at least three different languages, which are characterised by different statuses. The higher position is occupied by the official language of the country, generally the one inherited from the colonial period, which is used as a technical medium in official domains. The second position is generally reserved to the *lingua franca* of the macro area, typically an L2 for most of the speakers, who use it in their daily activities outside the family, at the market, when they travel to town or the like. The last position is represented by the local language, that is usually the speaker's mother tongue and is reserved to the domestic domain and to the verbal exchanges among close relatives or people living in the same village or in its neighborhoods in rural areas.

Many times, this basic triglossic model gets expanded with two or more other languages, due to different cultural traditions and marriage preferences. It is so for example in the Kulango sub-prefecture of Nassian in Côte d'Ivoire, where the Kulango (Gur) community represents a buffer group between the Akan (Kwa) speaking clans of the southern regions and the Lobi (Gur) located in the northerner areas. Kulango are traditionally open to intermarriages both with Akan and with Lobi groups and, therefore, a Kulango child can easily be exposed at the same time to his/her parents' language (Kulango), to his/her grandparents' one (possibly Lobi, Lorhon, Akan or even all of them), to Dyula (Mande), representing the *lingua franca* spoken at the market or in the shops and finally French (Romance) used in the school.

It should be self evident, but it is maybe important to underline, that in this case, the very same child is exposed not only to four (or five, or six) languages belonging to the same linguistic family, but even to four (or five, or six) languages belonging to different language families and/or linguistic phyla.

Two last points must be added (or anticipated) here in order to have a more complete idea of the whole picture:

1) of all the languages involved in the example just reported, the only one that is really used in daily written contexts is French. With the exception of Lohron, which has even not been fully described until now (at least as far as I know), the other four (Kulango, Lobi, Akan and Dyula) could actually be

written according to specific scripts based usually on the Latin alphabet with the addition of some diacritics and/or IPA symbols³, realised mainly by missionaries for their Bible translation, but certainly not universally known and, above all, very little, if ever, used for writing, being their usefulness limited just to the need of reading the Gospel or the Bible during the Mass or in other occasions reserved to the religious domain;

2) the very threat for true minority or endangered languages in Africa is not represented by ex-colonial languages, as it is often believed; rather their endangerment is much more due to other African languages enjoying a higher status, i.e. neighboring languages with a definitely higher number of speakers (in our example Kulango is suffocating Lorhon in the region due simply to their demographic majority) or the interethnic *lingua franca* (Dyula in our example), which allows its speakers to communicate with a larger network of people.

2. AFRICAN AUTOCHTHONOUS SCRIPTS

Despite the common view according to which Africa is a continent without writing, quite the opposite is true, since “The development of literacy in Africa seen as a whole certainly predates the histories of European colonialism and Islamic conquest” (Abdelhay; Asfaha & Juffermans 2014: 5).

One of the most ancient and prestigious scripts in the world, the Egyptian hieroglyph, was, indeed, born in Africa. The Nile valley, together with the Horn of Africa have been home to other very important ancient scripts, some of which are now extinct since long (e.g. Meroitic, Hieratic, Coptic, Old Nubian), while others still resist and continue to be used (e.g. Amharic, Ge’ez).

In addition to these more widely known written traditions, there are different other autochthonous scripts that were spontaneously developed by local communities in various parts of Sub-Saharan Africa, e.g. Nsibidi in southeastern Nigeria and Tifinagh in the Maghreb (now standardised in the Neo-Tifinagh form, which has recently become one of the three official scripts of Morocco, alongside with the Arabic and the Latin alphabets).

Of course, the local ancient scripts were reserved, known, used and passed down among specific groups of initiated peoples, mainly religious chiefs, without reaching and impacting on the lay population, but this is no wonder,

³ The case of Dyula offers indeed a nice example of a Mande language in which many *ajami* texts have been produced. We call *ajami* all those texts written with the Arabic alphabet but in local African languages. For writing Dyula, people could also use an autochthonous script, namely the N’ko alphabet, invented in 1949 by S. Kanté, which will be cited also in the next § and in Tosco’s chapter in this book.

since the same is true for the western writing tradition as well, which started to spread among the lower classes only when school attendance became compulsory.

Sometimes the invention of autochthonous scripts derived from the independent initiative of enlightened political chiefs, as it happened for example in Cameroon, where King Njoya himself, around 1885, created a special script for his language, Bamun, to be used in the court⁴.

Be it as it may, when the Arabic and Latin alphabets reached the continent and started to be used as mediums of “civilization”, first of all as a result of the Islamic and Christian missionary activities, or as a fruit of the commercial networks derived from the contacts with Arab merchants or the colonial administrations, they certainly “made an impact on Africa’s language and literary ecology” (Abdelhay, Asfaha & Juffermans 2014: 6).

In a way, indeed, these two imported, so powerful and flexible tools, worked as a time bomb for the creativity of many local leaders, who used them as a basis for the development of new peculiar scripts for writing down the local languages, mainly as an expression of cultural and ethnic identity.

As pointed out by Dalby (1967, 1968 and 1969) and Cooper (1991), dozens of scripts were created between the end of the 19th and the first half of the 20th centuries and they were usually (but not always) presented to the local population as a divine revelation accorded to religious leaders directly from God. Consequently, as it happens for the myriad of alternative religious movements which are born in the same way nowadays and which last until the charismatic leader is alive or in business, most of these scripts had a lifespan limited to that of their inventors and their public was limited only to a small group of proselytes.

Also in this case, however, some experiments had better outcomes with respect to others, and a couple of these scripts have survived till now, despite the fact that they remain limited to a small network of users. Two examples that owe being mentioned here are: 1) the Vai script of Liberia, “revealed” to Momulu Duwalu Bukele around 1830⁵, which is not an alphabet, but rather a syllabary; 2) the N’ko alphabet developed by the Guinean writer and educator Solomana Kanté in 1949 on the basis of the Arabic script and thought as a tool for writing down all the languages belonging to the Manding continuum of West-Africa⁶.

⁴ The script seems to have ceased to be used a few years after Njoya’s death in 1931.

⁵ On this see also Scribner & Cole 1981.

⁶ On this see also Vydrin 2017 and Micheli 2021 (forthcoming).

3. THE NATURAL SUPREMACY OF SPOKEN LANGUAGES

That spoken language is a priority with respect to its written variety is for linguists an evidence that does not even need being commented.

However, for a larger public, it is probably worth underlining that this claim is easily ascertainable when considered under different perspectives: anthropological, ontogenetic and phylogenetic.

Anthropologically, it is evident that all languages are spoken, while not all of them are written. Indeed, only few hundreds of the 7000 languages spoken in the world at present, are written.

Ontogenetically, it is under everyone's eye that every human being learns first of all to speak and then, maybe, and only under certain specific conditions, he/she can also learn to write. Learning to speak is a natural process, which spontaneously starts in babies, even without specific training in this sense from their parents, through the imitating mechanism of lallation. Parents, on their part, can only help accelerating the process, if they constantly interact verbally with their children.

On the contrary, writing is a skill that can be acquired only much later and which needs already a quite complex ability on the part of the child in terms of movement coordination (learning to keep a pen in the right way in order to sketch small signs on a paper is already difficult). Thus, drastically simplifying, we can just say that learning to write is a cultural rather than a natural process and, therefore, it needs to be passed down to the younger generation by experts spending their time with this specific objective in their minds.

Phylogenetically, it is a fact that, in human history, spoken languages appeared ages before someone had the idea of writing them down. Between the moment in which *homo sapiens* evolved morphologically so to develop an articulatory apparatus suitable for the production of verbal sounds and the moment in which the first human community invented the first script, at least one million years had to pass.

Still today, the vast majority of human beings produce huge amounts of knowledge, literature, memories and cultural traditions even though they are not able to write.

Indeed, writing is not a *condicio sine qua non* culture grows and peoples have always developed and expressed themselves through poetry, theatre, music or storytelling even without writing anything down.

From this point of view, Africa is a continent extremely rich and the value of oral traditions is immense and is characterised by specific features. Every summary I could propose here would be too poor to be considered interesting, therefore, I simply refer the reader to Ruth Finnegan (2012, 2014, 2018) and the therein cited bibliography for an exhaustive description of the richest oral traditions of the continent.

4. LITERACY AND EDUCATION IN AFRICA

Another stereotype on Africa is that it is a continent with no education and with a very low literacy rate. Unfortunately, under this point of view, the stereotype is not that far from reality, since, despite the average literacy rate in sub-saharan African countries has risen very quickly in the last five years⁷, it still remains attested, in general terms, at 65% and there are still cases where illiteracy represents a tragic problem.

In Chad, for example, only 22% of the population can read and write.

The best rate is 96%, registered in the Seychelles, but as a matter of fact, sub-saharan countries show a literacy rate of more or less 50%.

In addition to this already dramatic figure, there are evident disparities in the real distribution of literacy and quality education among the population. Usually, in fact, the literacy rate of women is lower than that of men, and in rural areas there are still many communities where the indicator drastically decreases.

For example it is so even in a country like Ethiopia, where the national literacy rate is attested at 52% and where in the capital city, Addis Ababa, there are high quality universities: in the Hamar and Daasanach districts of South Omo Region, in fact, the literacy rate is as low as 1 and 2% respectively (see Micheli 2019).

I will come back on the issue of education in the next §, but let me just stress here that, at present, primary education, even where it is quite well widespread, is generally carried out in the official language of the country, i. e., with very few exceptions (e.g. Somalia as the best example), in the European language inherited from the colonial administration; a language which, despite its high presence in daily use in capital cities and the big towns, at least for what concern upper class people, is instead very poorly known and understood in rural areas as well as by lower class people in towns (see for example Micheli forthcoming about the usability of French in Bobo Dioulasso, Burkina Faso).

In addition to the low literacy rate, it is worth underlining that the access to libraries or the possibility to buy books is not at all obvious in many areas of the continent.

This fact tells us at least two things about African literature: 1) orality still represents the more widespread and most commonly used medium for literary production and expression, and 2) written literature, usually produced in the ex-colonial languages, is still reserved to a very small minority of the African population and African writers are well aware that their readers are mainly people living in other parts of the world, rather than in Africa.

⁷ World Bank data 2019 - <https://data.worldbank.org/indicator/SE.ADT.LITR.ZS?locations=ZG>

However, writing in a European language can be an asset for them, if we think that in this way African writers can at least make black voices and instances be heard in the global village.

5. LINGUISTIC POLICIES

As we have already seen in §1, Africa is a continent socio-linguistically characterised by multilingualism.

Everyone speaks usually at least three languages: his/her mother tongue, the regional dominant African language and the official language of the country.

The prestige and the contexts of use of these different languages reflect socio-economic features and relegate minority languages to the lowest status. Therefore, for the vast majority of these languages the risk of disappearing in the next few generations is not only very high, but also, unfortunately, quite probable.

Considering the possible evolution of the vitality of spoken languages all around the world, due not only to socio-economic reasons, but also to the effect of mass communication in social media and the like, Michael Krauss warned that “By the turn of the century only 600 languages will remain on the face of the earth, meaning that 90% of the world’s languages will have perished”⁸.

Of course, this phenomenon will have a much higher impact on the local living languages of sub-saharan Africa than the one it will have in the Western world.

In fact, in the Western world, almost all minority languages have nowadays at least been described and tools have been created for their preservation.

Already in 2001, according to the UNESCO *Atlas of the World’s Languages in Danger of Disappearing*, in Africa 97 languages were on the verge of dying. This means that already 20 years (or one generation) ago, one hundred African languages had only a few dozen speakers, all of whom in their old and very old age.

Today linguists usually consider a language as being severely endangered if it:

- has lesser than 5000 speakers;
- has no intergenerational transmission;
- has no social prestige;
- has mother tongue speakers who are bilingual in the regional dominant language and prefer to use the latter in their daily activities;
- does not respond in a creative way to new domains.

Looking at things under this light, it results self evident that the situation of African indigenous languages is dramatic and that the number of endangered

⁸ Michael Krauss in Batibo 2005:VII.

languages at present easily reaches and even probably surpasses four or five hundreds.

There is no way to discuss the issue at length here, but, as Batibo (2005: 114) puts it:

“Only if there is a strong political will associated with a chain of activities such as the sensitization of speakers, documentation of the minority languages, their introduction in school systems and promotion to wider public use will language revival succeed”.

Unfortunately until now, despite the true commitment of some enlightened rulers (as a single name I mention here Mwalimu Julius Nyerere, the first and beloved President of Tanzania), linguistic policies in Sub-Saharan Africa result in many cases inefficient and in some cases even non existent.

It is true that innovative programs, developed with the best intentions, have been proposed here and there and from time to time in different countries, aimed not only at providing primary school education to every child, but also to do so including local African languages in some experimental projects for primary schools (see for example Micheli 2021 - forthcoming - Sawadogo 2004 e Sanogo 2011 about Burkina Faso).

Unfortunately most of these experiments, which imply the use of local major languages alongside the official ones in different percentages during the different years of the first cycle of primary school, remained in most cases only on paper and were actually never realised or realised only in a very small number of schools with very few pupils involved.

It must be added that these kind of programmes, despite their positive intentions, are potentially dangerous for the ecology of local linguistic dynamics, because, if on the one hand it is already difficult to choose which of the many languages of the region shall actually be used in the schools (on this topic see for example Jufferman & Abdelhay 2016), on the other hand the use of a local language in the school is usually not perceived as an asset by the majority of parents, who rather prefer their children to start learning the official language in a “proper” way from the beginning, given that, in their view, only a proficient knowledge of the official language represents a real possibility for their children to get good jobs and reach better life conditions.

6. CONCLUSIONS (ADVOCATING FOR A BETTER APPRECIATION OF SPOKEN AND MINORITY LANGUAGES OF AFRICA)

With this short contribution, I hope to have reached the goal of outlining the 5 most important features characterising the extremely varied and highly complex linguistic panorama of the African continent.

I also hope to have been able to argue why neither an accurate enumeration of living, endangered or extinct languages, nor a detailed projection of that enumeration on topographical maps are appropriate ways for discussing all the delicate questions rotating around linguistic diversity, linguistic rights and linguistic agency of the different African peoples.

As reported in § 5, by the end of this century, our world will probably lose 90% of its languages and, of course, the first languages to disappear will be those languages whose prestige in terms of number of speakers and dignity of use in formal domains are low, i.e. the quasi totality of African languages.

In this perspective, only a huge effort in terms of seriously planned linguistic and educational policies adopted by States could help to partially hinder or (at least) relent the process.

Of course, we all know that Sub-Saharan low and middle income countries face a huge quantities of problems, ranging from health, economy, civil and political tensions, the effects of climate change as well as many others in addition, but I think that the question of good quality education and the protection of minority peoples and their languages should at least appear on their agenda as a true priority.

As a matter of fact, linguists, experts in cognitive studies as well as psychologists and educators all around the world underline since long now, that pupils learn more quickly and in a better way when they are taught in their mother tongue or in a language they truly manage.

In Africa, as we have seen in §§ 4 and 5, most primary education is taught in the country official language, which is usually the one inherited from the colonial administration and which is sometimes completely obscure to young children.

In addition to this, we have seen how, even when experimental programmes have been proposed, until now they have too often failed.

Therefore, much is still to be done (and must be done) to grant all pupils the right of learning at least during the first two years of primary school in a language they are really at ease with.

Moreover, it is important to consider new pedagogic strategies, more focused on the African rich tangible and intangible heritage, oral traditions, and artistic / literary production in order to trigger a more positive attitude of the youth towards their past and their cultural / ethnic identities.

Only reinforcing their education systems will African governments help their population to become more conscious of who they are and of the possibility they have, in order to be ready to compete and cooperate on an even footing at all levels with the outside world.

REFERENCES

- Batibo, Herman M. 2005. *Language Decline And Death In Africa: Causes, Consequences And Challenges*. Multilingual Matters.
- Blommaert, Jan. 2008. *Grassroots Literacy: Writing Identity and Voice in Central Africa*. London: Routledge.
- Cooper, Robert L. 1991. "Dreams of scripts: Writing systems as gifts of God. The Influence of Language on Culture and Thought" in: Cooper, Robert L. & Spolsky, Bernard (eds.) *Essays in Honor of Joshua A. Fishman's Sixty-Fifth Birthday*, Berlin: Mouton de Gruyter, pp. 219–226.
- Dalby; David. 1967. "A survey of the indigenous scripts of Liberia and Sierra Leone: Vai, Mende, Kpelle, and Bassa" in: *African Language Studies*, vol. 8, pp. 1–51.
- Dalby, David. 1968. "The indigenous scripts of West Africa and Surinam: their inspiration and design" in: *African Language Studies* vol. 9, pp. 156–197.
- Dalby, David. 1969. "Further indigenous scripts of West Africa: Manding, Wolof, and Fula alphabets and Yoruba holy-writing" in: *African Language Studies* vol. 10, pp. 161–191.
- Dimmendaal, Gerrit J. 2008. "Language ecology and linguistic diversity on the African continent" in: *Language and linguistics compass*, 2 (5): 840–58.
- Finnegan, Ruth. 2018. *Oral Poetry: Its Nature, Significance and Social Context*. Wipe and Stock Publishers.
- Finnegan, Ruth 2014. *Where is Language?: An Anthropologist's Questions on Language, Literature and Performance*. London: Routledge.
- Finnegan, Ruth. 2012. *Oral Literature in Africa*. Open Book Publishers.
- Hinton, Leanna (ed.) 2018. *The Routledge Handbook of Language Revitalization*. London: Routledge.
- Juffermans, Kasper; Asfaha, Yonas, M. & Abdelhay, Ashraf (eds) 2014. *African Literacies: Ideologies, Scripts, Education*. Newcastle-upon-Tyne: Cambridge Scholars Publishing.
- Juffermans, Kasper & Abdelahay, Ashraf. 2016. "Literacy and Multilingualism in Africa" in: Street, Brian, V. & May, S. (Eds.) *Literacies and Language Education*. Springer International Publishing, pp. 1-12.
- Lüpke, Friederike. 2010. "Multilingualism and Language Contact in West Africa: towards a Holistic Perspective" in: *Journal of language contact - THEMA* 3, pp. 1-11.
- Lüpke, Friederike. 2018a. "46. Supporting vital repertoires, not revitalizing languages"

- in: Hinton, Leanna (ed.) *The Routledge Handbook of Language Revitalization*. London: Routledge, pp. 475-483
- Lüpke, Friederike. 2018b. "Escaping the Tyranny of Writing: West African Regimes of Writing as a Model for Multilingual Literacy" in: Weh, Constanze & Jufferman, Kasper. *Tyranny of Writing. Ideologies of the Written World*. Bloomsbury Academic Publishing, pp. 129-148.
- Micheli, Ilaria. 2010. "Le lingue dell'Africa" in: *AiOn, Annali dell'Istituto Universitario "L'Orientale" di Napoli*, vol. 30, pp. 1-30.
- Micheli, Ilaria. 2019. "Considering the gender narrative in the mirror of reality: how women consider and manage their right to health in some pastoralist tribes of South Omo Zone (Ethiopia)" in: *Ricerca Folklorica*, vol. 74, pp. 229-243.
- Micheli, Ilaria. 2021 (forthcoming) "Limits and potential of Dyula in Burkina Faso: observations for use in cooperation", in: Baggio, Serenella (ed.) *Lingua Franca, Lingue franche*. Edizioni dall'Orso.
- Sanogo, M. L. 2011. *L'influence de la fonction sur la forme interne de la langue. Le cas du dioula à Bobo-Dioulasso*. Ouagadougou, INSS-CNRST (texte de communication).
- Sawadogo, G. 2004 "Les langues nationales à l'école burkinabé: enjeux d'une innovation pédagogique Majeure" in: *Repères*, N°29, 2004, pp. 251-260.
- Scribner, Sylvia & Cole, Michael. 1981. *The psychology of Literacy*. Cambridge: Cambridge University Press.
- Tosco, Mauro. Forthcoming. "MEASURING LANGUAGENESS: Fact-checking and debunking a few common myths" in: *DIVE-IN*.
- UNESCO. 2001. *Atlas of the World's Language in Danger of Disappearing*.
- Vydrin, Valentin. 2017. "5. New Electronic Resources for Texts in Manding Languages" in: *Searching for Sharing. Heritage and Multimedia in Africa*, in: Merolla, Daniela & Turin, Mark (eds.) Cambridge: Open Book Publishers (https://open-bookpublishers.com/html-reader/978-1-78374-318-6/contents.xhtml#_idTextAnchor006)
- World Bank Organization - <https://data.worldbank.org/indicator/SE.ADT.LITR.ZS?locations=ZG>
- Weh, Constanze & Jufferman, Kasper. 2018. *Tyranny of Writing. Ideologies of the Written World*. Bloomsbury Academic Publishing.

From IPA to wildcards: A critical look at some African Latin orthographies

MAURO TOSCO
University of Turin

ABSTRACT

The article presents and discusses a few African Latin orthographies. The scope of the work is set out in section 1, while section 2 discusses a few orthographies featuring IPA symbols and diacritics. They were often the work of linguists and missionaries and were conceived for mother-tongue alphabetization and in order to translate and publish religious literature. They are scarcely useful in everyday casual writing, and especially so on a keyboard (where only a restricted set of symbols is to all practical purposes available). They are contrasted in section 3 with the use of digraphs and, most of all, with “wildcards:” symbols of the Latin, basic (unmodified) alphabet that are taken to use, often in an idiosyncratic manner, in order to represent phonemes that do not have a direct, built-in representation. The discussion is wrapped up in section 4, where the limits on the use of wildcards are evidenced and the practical limitations of many African orthographies reiterated.

KEYWORDS

Orthographies; Latin script; IPA; diacritics; digraphs.

1. A SHORT INTRODUCTION WITH MANY CAVEATS AND UNSOLICITED EXCUSES

Why are many African orthographies so bad? Is it because so many African languages are phonologically complex? Indeed, they are. But is this the only reason? And, for that matter, in what sense can we say that an orthography is “bad”?

This article will delve into an analysis along language-internal criteria, and the appropriateness of an orthography (which inversely correlates with deviation from a one-to-one correspondence between graphemes and phonemes; Sgall 1987) will not be called into question. Rather, attention will be focused on effectiveness (Cahill and Karan 2008; Cahill 2015). Granting that graphization has been, in Africa and elsewhere, a top-down process and that data on the approval of an orthography on the part of its real stakeholders – all of its potential users – are difficult to get, the analysis will concentrate on the internal characteristics of an orthography that are in all likelihood bound to facilitate or hamper its use.

Just as all languages are equal but some of them are more equal than others because they have been ausbauized (Tosco 2008) into written, official languages, also not all orthographies are equal: “big” languages may go along fine for centuries with awful, impractical and obsolete orthographies. Minority languages cannot: often their speakers are illiterate and must face the usual hurdles of acquiring literacy; in many cases reading and writing skills are offered to the minority language speakers in a locally or internationally big language. Both possibilities are widespread in Africa (and elsewhere). Or the minority language speakers already know the local majority language and its orthography and have thereby access to a sizable amount of material and information available in that language; the additional skills required in learning and using the minority language orthography are mostly justified in ideological and moral terms. Any material produced in the minority language will face very high hurdles in dissemination and will not be likely to be a winning competitor, in both qualitative and quantitative terms.

In Africa and elsewhere many orthographies were designed by foreign linguists, policy makers or missionaries – often with some input from local language consultants – and basically with the aim of publishing in the local language. That the publication itself is a Holy Book or the word of the leading party and its sacred leaders does not change much: the speakers themselves will mostly be readers – i.e., passive users. A technically very accurate orthography will certainly be easy to learn.

Not the same can be said about actively using the orthography – i.e., as writers, and especially so when handwriting is replaced by the use of keyboards (of typewriters in the past and of computers or cellphones nowadays).

It is certainly true that most of these orthographies have never been tested in extensive, daily usage by speakers, but this use has probably never even been considered.

In short, one feels that not all the needs of the potential users were taken into consideration: the degree to which an orthography will be user-friendly will be much different if reading or writing are taken into consideration.

As the title implies, orthographies based upon the Latin alphabet only will be presented and critically discussed, and actually a tiny minority of them. Specifically, I will not be concerned with:

- the vigorous use of the Arabic alphabet, not only in the past but still today (for which Mumin and Versteegh 2014 is not only a must, but also a fascinating reading);
- the use and present spread of indigenous scripts (*in primis* the Ethiopic syllabary in Ethiopia and Eritrea and the recently revived Tifinagh Berber alphasyllabary);
- finally – and regretfully (and simply out of lack of competence and data) – the rise and to a certain extent spread of many “new” alphabets of Africa – such as the Vai syllabary, the N’Ko alphabet or the Mandombe script. A good overview is provided in Kootz and Pasch (2010).

2. THE UNBEARABLE WEIGHT OF LINGUISTICS (AND LINGUISTS)

Both digraphs and IPA-based modifications of Latin letters are found in the orthography of Hausa (Chadic; ISO 639-3: hau¹), probably the African language with the highest number of native speakers. Digraphs are < sh >, and < ts > for /ʃ/ and /tʃ/, respectively; IPA symbols are < ɓ >, < ɗ > for implosives and < ɛ̃ > for ejective /k̟/. Ejectivization is instead not marked in the case of the affricate /tʃ/, while < ‘y > stands for a palatal glottal stop /ʔ/. Under representation is found in the case of the same symbol < r > used for both an alveolar trill (the latter often expressed by < ʀ > in linguistic works, with plain < r > being reserved to retroflex flap /ɽ/) and most of all in the case of vowel length, which goes unmarked. As in many, maybe most African orthographies, tones are left unmarked.

Not surprisingly, leaving aside the still widespread use of Arabic-based *Ajami*, everyday Hausa written in *boko* (/bōkò/), i.e., the Latin alphabet, often disregards the “hooked” letters altogether and ‘one still encounters publications, including newspapers, where the plain letters are used’ (Jaggar 2001: 698); ‘the letters are printed without the hooks’ (Newman 2000: 726).

¹ The ISO 639-3 code is provided for all and only African languages after their first mention in the text.

Nor are these problems limited to languages spoken in former British colonies: a cursory look at a few major languages of West Africa, suffices to show the extension of the IPA symbol < ŋ > for a velar nasal – e.g., in both Wolof (Atlantic; wol) and Bambara (Mande; bam). The latter adds to its inventory of graphemes marked by IPA symbols < ŋ > for the palatal nasal as well as < ε > and < ɔ > for the open-mid vowels; the same array of graphemes is also used in Dyula (Mande; dyu), a major lingua franca of West Africa (whose use and potentialities also as a written language are discussed in Micheli forthcoming).

Other languages face more and more difficult problems in their graphization.

In the case of many languages of South Sudan, troubles started at least in 1928, at the time of the Rejaf Language Conference (Tucker 1929), for which Abdelhay, Makoni and Makoni (2016) provide a useful overview. Their attention mostly goes to the ideological aspects of the conference, and their stark critiques – framed in the new orthodoxy of postmodernism and postcolonial studies – focus on the “Orientalist” attitude of the participants (where “Orientalist” is of course an abusive term) and the alleged invention of discrete ethnic groups and languages with the aim of reinforcing and maintaining colonial power. Still, their analysis of Dietrich Westermann’s (1875-1956) orthographic proposals has some merit.

Following his long research in Sudan and West Africa and serving as director of the *International Institute of African Languages and Cultures* (later the *International African Institute*) from 1926 until 1939, Westermann proposed a *Practical Orthography of African Languages* (1928, 1930).² Proposed IPA symbols to be used in these “practical orthographies” include among others < ŋ >, < ʃ >, < ʒ >, < ʏ >. Further recommendations include the use of apostrophes for ejectives and/or implosives, of umlaut (diaeresis) for “central vowels.” Also, the notation of dental stops with the digraphs < dh >, < th > is recommended.

An early example of these orthographical choices is Heasty’s (1937) Shilluk dictionary, but the same solutions lie at the basis of the alphabets still used for many languages of South Sudan of different genetic affiliation and to some extent used in education (English only being the official language of independent South Sudan).

Among the main languages, Dinka and Nuer are cases in point: Dinka (din) and other West Nilotic languages have both modal and breathy vowels. Breathiness is marked in IPA by a subscript umlaut (diaeresis), as /a̱/, /e̱/, etc. The IPA notation is apparently the source of the Dinka superscript umlaut above the vowel, as in < ạ̈ >, < ẹ̈ >, etc.

² Available online at <http://www.bisharat.net/Documents/poal30.htm> (last accessed June 2, 2021).

A a	Ä ä	B b	C c	D d	Dh dh	E e	È è	É é	Ë ë	G g	Y y	I i	Ï ï	J j	K k	L l	M m
[a]	[ä]	[b]	[c]	[d]	[d̥]	[e]	[ẽ]	[ē]	[e̝]	[g]	[y]	[i]	[ĩ]	[j]	[k]	[l]	[m]
N n	Nh nh	Ny ny	D̥ d̥	O o	Ö ö	Ɔ ɔ	Ǿ ǿ	P p	R r	T t	Th th	U u	W w	Y y			
[n]	[nh]	[ny]	[d̥]	[o]	[ø]	[ɔ]	[ɔ̃]	[p]	[r]	[t]	[t̥]	[u]	[w]	[j]			

Figure 1. The Dinka alphabet (<http://www.languagesgulper.com/eng/Dinka.html>)

Although a few digraphs are used for consonants, no generalization of such a solution has been attempted. In stops, a breathy release is marked in IPA by a raised symbol for a voiced glottal fricative, as in /bʰ/: one could imagine a transcription * < ah > for /ḁ/ (instead of < ä >), etc. As vowel length is not phonological, one could even use * < aa >. This leaves the problem of vowel quality: seven vowel qualities are phonemic in Dinka, with open-mid /ɛ/ and /ɔ/ opposed to close-mid /e/ and /o/. Accents are often used in European languages to the effect of marking openness, but other solutions are conceivable, such as digraphs (* < ae >?). The Dinka orthography simply keeps the IPA symbols, and not only sports < ε > and < ɔ >, but also breathy < ë > and < ö >.

Other IPA signs are found in the case of < ɣ >, < ɲ > and < ŋ >.

The Dinka orthography is interesting for the use of digraphs whose second element is consistently < h >: < dh >, < th > and < nh > for the dental counterparts /d̥/, /t̥/, /n̥/ to alveolars /d/, /t/, /n/.

At the same time, signs provided by the Latin alphabet and available in any standard keyboard but left unused in the Dinka orthography abound: < f >, < h >, < q >, < s >, < v >, < x >, and < z >. One can easily argue that at least < s > and < z > could be good solutions for the dental stops. As < ny > is used for the palatal nasal stop /ɲ/, the absence of the perhaps even more common digraph < ng > for the velar stop /ŋ/ is puzzling.

2. Bër ë dët

Ruëëth aabi yök anyiköl yic ë yän juëc yiic anyiköl yic.

Añuï kek Awan

Wäthæε yen aci Añuï lo tënë wën ë nyankënë cäl Awan ku lëk yeen, “Manh ë nyankäi, awiëc ba pïr yïn yaköl.”

Go Awan dhuök yeen, “Yeñö yinër ye yïn pïr yæn? Na wεεñdïit tö tēën cä ke bi la dök yïin, bi yïn ke pïr!” Ku kat Awan bi wεεñdïit wën nyuäth la kuaath wïir tiök yic bi weñ döt ebën. Ku le riñ tënë nërë Añuï ku lëk yeen, “Yen kan yen miëthdu, yïn abi cuet yaköl tē cït tē wïc.”

Añuï aci puöu jal mist ku riñ wïir bi yök la cam. Na le yëët tiök yic, kë döt aya cëmën yök, go yic löny bi dhiau.

Go Awan lëk yeen, “Duk go dhiau kë wun yök këc gua bën. Na le yëët lä, ka aba kuany yic apath yeñö këne diëc yäkke yiic.”

Thiëc

1. Yeñö ci Awan nyuöth Añuï bi ya miëthde?
2. Ci Awan poth nërë thok këdë?

Figure 2. An excerpt from a page in Dinka (Caguor 2003: 4)³

The same system is basically followed for Nuer (West Nilotic; nus), with the umlaut being replaced for breathiness by underscore:

³ This and all the following specimens of South Sudanese orthographies were obtained in Juba in 2013.

Man Dεεη κενε Nyaluaak cike wä ruup.
Mëë cike cop ruup, cuke wä göör ke tɔɔŋ.
Mëë görke, cu Man Dεεη nyääl nen rey
butbutä. Cue Nyaluaak cɔl. Cue wee, "Nen
nyääl emɔ. ɛ nyääl in la cam naath κενε leey."

Cukε riŋ wic. Mëë cike cop thaar wec, cuke
gaaç εlɔŋ ke dual mi ci dual ke ke. /Ken nyääl
emɔ ben κɔɔriɛn. Cue duoθh gaaθhde. Cu
Man Dεεη cieŋ nɔŋikä ke tɔɔŋ kä cue thaŋŋ
tɔɔŋni banɣ ruup. Nyaluaak ce dual εlɔŋ a thile
tɔɔŋ ti ce nööŋ dhɔr man.

Mëë ci Man Dεεη cop cieŋde, cue lar Guan
Dεεη i "Cakɔ nyääl nen rey butbutä ruup."

Cu Guan Dεεη wee i "ɛ jɛn nyääl in la deɣ
yɔɔk a cam en wic emɛ."

Cu Guan Dεεη ε wä guil κενε Dak ke Man
Dεεη ke yöö ŋäce gaaθhde. /Ken κεν ε jek.
Cu Man Dεεη wee, "Cε jɛn. ɛ jɛn gaaθhde,
en butbut emɛ."

Mëë cike ben wic, cua ke thiec, i "Ci yen ε
jek?" Cukε wee /ken κεν ε jek.

Cu Man Bol wee, "Kä ce wä nikä, en
nyääl?"

Cu Man Dεεη wee, "/Cä dee ŋaç en gaaθh
ε ce wä thɪn ɔ."

Figure 3. An excerpt from a page in Nuer (Kuëne ke Thok Nath Bok ŋuaan 1994: 65)

Breathiness is not phonological in Murle (Surmic; mur), and this entails the absence of umlaut and underscore; on the other hand, the same IPA signs seen above are again used for both vowels and consonants:

ZOZOK CI ADUWONEK DEMEZOK

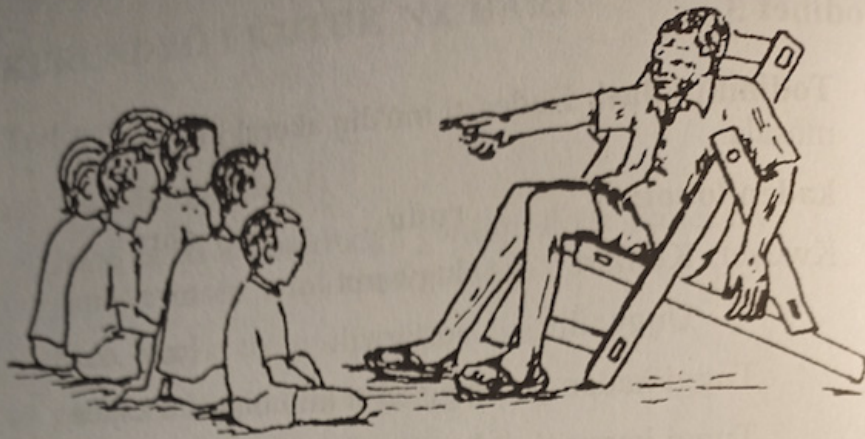
Omoloye ceen aməthəth thur ki ram o 52 obek deməzyak kəɛŋ ci (sura) ririthən ci avi ŋina noko. Zɛɛ enek nɔɔɔ nɛ, "Anyak zɔɔz ci colai ŋina. Agayu inoko zɔɔz nico kawo kazi nɛ? " Kanɛ Dokol." ŋina oovɔ zɔɔz ci ona kazi Dokol avi ŋina noko enek deməzyak nɛ, kazi zɔɔz nico nɛ, "Dokol" jɪnɔ bodo izi nɛ, kawo zɔɔz nico kazi nɛ? Azi deməzyawa nɛ, kawo kazi nɛ, "Dokol" ŋɛnɛn inoko ŋaathunonɔ ci ajowa zɔɔz ci kazi Dokol noko ŋaman ŋina? Azi deməzyaithi oman nɛ, nɔɔn ce. Izi zin niina demezoithi nɛ, "Abona ɔɔɔth". Anyek zin ŋina deməzyak wanɔi kathamanith kekebith golok ci zoozok niciko noko dook niigi doon kidicilim. Zɛɛ mazi icinu ijezith inoko niigi dook, Anyek zin nɔɔɔɔ dook kekebith moloye ci abɔbɔ. Zɛɛ mazi odothith anyek kebedozeke zoozok ci ɛɛl kaale ci aliyai ɛɛl ɔɔɔɔgi lootha noko kekebith niigi dook.

Figure 4. An excerpt from a page in Murle (Warage O Miliny Murleye 2004: 54)

The use of diacritics and IPA signs is quite widespread, being found in many, if not most, orthographies of the area.

Another major language of South Sudan, Bari (East Nilotic; bfa and others), replaces most IPA signs used for Dinka and Nuer with diacritics, with the exception of < ŋ > for the velar nasal. The implosives are marked by a preceding apostrophe: < 'b >, < 'd >, < 'y >. Umlaut is preserved in the case of < ö > and a single digraph < ny > is found. No breathiness contrast operates in Bari.

In this regard, Owen's (1908) Bari grammar, with its abundance of umlauts, accented letters and digraphs (among which... < ng >) but no special symbol, was certainly better.



Ŋutu modon̄ ko ŋwajik kanyit buker.

Ŋutu lele kata karen kanyit a Goke. Lepen̄ lo gwe a modon̄ parik. I diŋit nagon lepen̄ dene adi nye gwe adara ko kulya ti togolon a luŋundye ŋwajik kanyit liŋ kak. Kirut a lepen̄ kulyani ko lepen̄at adi "Ŋwajik kwe! Ta lo meddya nan ŋilo, nan gwe a modon̄ parik, riŋit nio kitayet gwe 'bayin, ko konyen kwe tine bulö meddya 'bura, swö kwe tine bulö yin̄ga 'bura. Nan yönö yin̄ga kasu i kutusen ko ta molu wulundyö wulundyö kitajin nagon a nyo."

Kirut a kilo ŋwajik kanyit poki i wu'yö na katajin kase gelej gelej. Togeleŋ adi, nye yönö kita. na katugwörönit lo toro'bo, tomurek adi nye yönö kita na tumunit, tomusala adi nye yönö kita na karyakanit lo kisuk ti ŋutu, a toiŋwan adi nye kita a kakurunita tobuker adi nye 'dekan kita na yuggu na kisuk.

I mukök nagon lepen̄at liŋ awu'yö kitajin kase, a monye lose kulyani ko lepen̄at adi, "Ana'but ŋwajik kwe, nan aje yin̄ kasu i kutusen. Ta lo meddi. Kogwon kulye kasu kilo

Figure 5. An excerpt from a page in Bari (Jujumbu Kendya ko Bari Buk Tomusala 1999: 52)

Mödö (or Jur; Bongo-Bagirmi; bex) is another language of South Sudan. The orthography used in Perrson and Perrson's (1991) dictionary and grammar resembles Bari in its use of < 'b > and < 'd >, to which < 'j > for a palatal implosive is added. Again, < η > marks a velar nasal and it further appears in the digraphs < ηg >, < ηb > and < ηm > for prenasalized phonemes.

Umlaut is used for < ï >, and < ë >, and a special sign for < ɔ >.

The extensive use of IPA symbols seems restricted to languages of South Sudan for the historical reasons outlined earlier in this section; digraphs, accents and apostrophes are rather used elsewhere. The orthography proposed for Rendille (East Cushitic; rel) of Kenya makes wide use of digraphs but also of an apostrophe preceding the sign for the modal stop in < 'b > and < 'd > for the implosives, and also, strangely enough, in < 'h > for the pharyngeal /ħ/. An apostrophe following a digraph is used in < ng' > for the velar stop (following the orthography of Swahili). The other digraphs are < ch > for the modal affricate /tʃ/, < kh > for the velar fricative /χ/, and < ny > for the palatal nasal. Acute accents mark a high tone. This is all the more disconcerting since Rendille belongs with Somali to the same sub-subgroup of East Cushitic (according to current classifications, they make up, together with Boni, the eastern branch of Omo-Tana, itself a major branching of East Cushitic). As we shall see in more detail below, the orthography of Somali could have provided a solution for a few phonemes, such as < dh > instead of < 'd >. Although in Somali the corresponding phoneme is postalveolar (/d/) rather than implosive (/d'), it could easily have been adopted and provide a model for * < bh > instead of < 'b > for bilabial /β/. While the use of the same pattern for * < hh > for the voiceless pharyngeal /ħ/ could have caused problems (in gemination), Somali offered an easy viable alternative in its use of < x >.

In Rendille, < x > is just one among a sizable number of unused signs of the Latin alphabet; the others are < c >, < p >, < q >, < w > and < z >.

Alohi Chiirnaan

(The Rendille Alphabet)

Chi' Lagaabicho (Short)	Chi' ladeeraacho (Long)
A, a..... abár	Aa, aaabaár
B, b..... haabáb	bb`dábbal
Ch, ch..... siicha	cchficcha
D, d..... disda	dddaddaáb
'D, 'd..... `dóo'd	'dd`dá'ddab
E, e..... ébel	Ee, ee..... eéra
F, f..... fálfí	ffáf far
G, g..... daaga	gghaggá
H, h..... á kaaha	hhá kaahha
'H, 'h..... á `hiira	('hhá a`hhiira)
I, i..... íbir	Ii, iiíid
J, j..... Jíirá	jjjíijjo
K, k..... kákahe	kkdíkkil
Kh, kh kháakhle	kkhíkkho
L, l..... léyley	lllallaáb
M, m..... maalím	mmmammáh
N, n..... naánah	nninnó
Ng', ng' ng'óoj	—
Ny, ny..... nyarnyáar	nnynyannyaame
O, o..... órro	Oo, ooóor
R, r..... Rén'dille	rraraárri
S, s..... síbbis	sssússukh
T, t..... tífto	ttháttto
U, u..... usú	Uu, uuúur
W, w..... Waáakh	wwhawwées
Y, y..... yaáy	yyyéyyah

Figure 6. The Rendille alphabet (Wori Haaggane MARKO Khore 1993)

Very similar is the alphabet devised by the same missionary body for Dhaasanac (East Cushitic; dsh). Here again we find no IPA symbols and the implosives are represented with an apostrophe preceding the sign for the

voiced stop; as in Rendille, acute accents (in Dhaasanac, on both moras of a long vowel) mark a high tone.

At the same time the alphabet has quite a few peculiarities: the digraph < dh > marks a laminal voiced fricative /ð/, and does not take into account that an alternative pronunciation with its apical counterpart /z/ is well attested in all positions (Tosco 2001: 19).

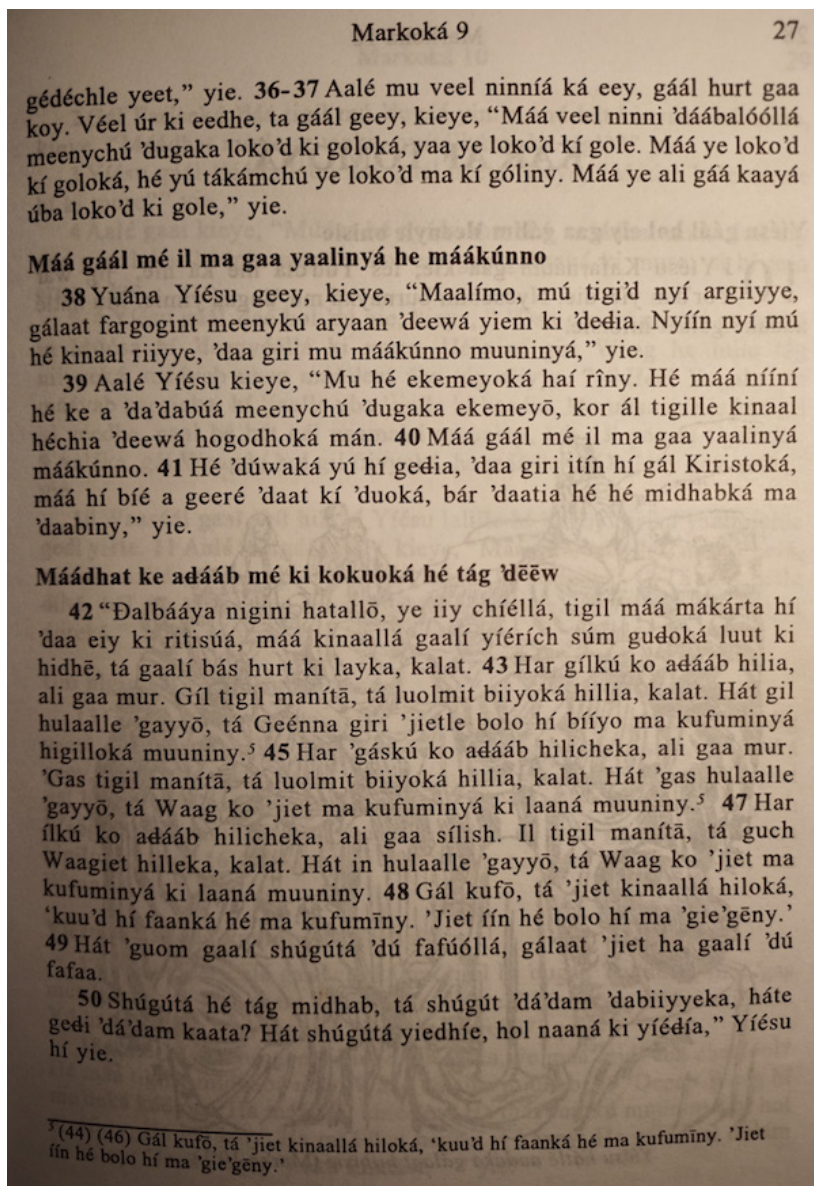


Figure 7. An excerpt from a page in Dhaasanac (War'gat Markoká 1997: 27)

Both diacritics and special symbols are found in the orthography of Gawwada (East Cushitic; gwd) proposed by the SIL International Literacy Department (with the additional complication that both the Latin alphabet and the Ethiopic syllabary are suggested). The apostrophe here follows the consonantal sign and marks an ejective (thus following the IPA conventions) in < c' >, < k' > and < t' >. IPA symbols are used for the pharyngeals: / ʕ / and / ħ /. Apart from < sh > and < ny >, the list comprises the use of < h > as second element for the implosives: < bh > and < dh >, but with no value in the case of < ch > for the affricate /tʃ/ (as in English) and of < qh > simply for /q/. The last two also imply that no “bare” < c > nor < q > are used, as well as no < z > nor < v >.




































IKKITO POKO ርALENA				ALE ALPHABET			
a  arapko	aa  aaxe	b  bahante	ch  chepʕe	c'  c'iwc'iwe	d  delle	e  telte	
ee  eero	f  fileye	g  gupo	h  harre	h  hisse	ɕ  ɕardo	i  ilge	
ii  iibaare	j  mirja	k  kocho	k'  k'alk'allo	l  leeʕo	m  maango	n  niiche	
ny  nyannya	o  orpo	oo  toomo	p  pasho	qh  qhole	r  rigte	s  siito	
sh  shenge	t  tawo	t'  t'ooqhanako	u  urrache	uu  uuffe	w  waalo	hx  hxaare	
	y  yiyako	lo'o  lo'o	bh  bhaqhasa	dh  dhilanko	q  quladi	z 3 izah	

Figure 8. A proposed Gawwada alphabet (<https://www.alepeople.org/sites/www.alepeople.org/files/ALPHABET%20of%20CA%95ALE.jpg>)

It is noteworthy that many of these orthographies, and in particular those of Kenya and Ethiopia, completely disregard the practical alphabets designed from the seventies for languages which are structurally and phonologically similar: just as no use of the solutions devised for Somali is made for the very similar Rendille, no attention is paid in the case of Dhaasanac to the contemporary Latin orthographies of Ethiopia (where a majority of the Dhaasanac live).

It is to these innovative orthographies that we turn our attention in the next section.

3. EXAPTATION, OR: LEARNING TO USE WHAT YOU HAVE

3.1. PLAYING WITH WILDCARDS

The Horn of Africa is home to at least two success stories among African Latin-based orthographies: two official or national languages in their respective countries and with many million speakers as well as potential users of written texts: these are Somali and Oromo (both East Cushitic; som, orm). Somali came first, with its orthography officialized in 1972.

The long, troubled history of the graphization of Somali and of the Somali language policy has been told many times and is the subject of whole monographs: Caney (1984) mainly deals with linguistic issues—history of the orthography and corpus planning; Labahn (1982) with the orthography as well as language policy in general. Laitin (1977, 1992 – the latter within the larger African context) explores the political side (although overtly biased in favor of state interventionism and nation building; for a critique cf. Tosco 2014). Short historical overviews are provided in Tosco (2010, 2015).

Predictably, Somali uses the digraphs < sh > for /ʃ/ and < kh > for /χ/, as well as < dh > for a postalveolar /dʒ/. Uvular /q/ is marked by < q > and vowel length by redoubling the sign for the vowel. Pitch is not marked (as well as vowel backing/advancement). Glottal stop is only marked when not in word-initial position by an apostrophe: < ' >.

The main problem was the absence of an established and practical way to mark the pharyngeals /ʕ/ and /ħ/.

The genial solution came with the use of unmodified Latin letters, namely < c > for voiced /ʕ/ and < x > for voiceless /ħ/ (< p >, < v >, and < z > remain unused).

Consciously or not, it was realised that any Latin alphabet is bound to have a few “wildcards”: symbols that simply come for free with the choice of using the Latin alphabet but have no clear phonemic value to start with, and are therefore available to get assigned, in principle, any value.

A few Latin letters are born as wildcards: e.g., < q > and < c > already in Latin marked allophones of /k/. In the conclusions we will argue that, nevertheless, < q > is worse than < c > as a wildcard. As for < x >, in Latin it was used since the beginning for the cluster /ks/ and has been put to many different uses in different orthographies around the world (its value as /ʃ/ in Maltese, Basque and many other languages is a major example).

Other letters become wildcards on a language-specific basis whenever a phoneme usually expressed by that letter does not exist. Of course, being language-specific, the value of a wildcard is also much “lighter” than the established value of another letter: it is therefore particularly prone to substitution.

When establishing an orthography for ‘Afar (Djibouti, Ethiopia, Eritrea; aar), the orthography of neighboring Somali was a possible choice. ‘Afar is spoken in Djibouti, Eritrea and Ethiopia, but in Djibouti only

competition with Somali and the ideological need to obfuscate similarities led instead to an orthography where the peculiar choices of Somali were shuffled: pharyngeals < c > and < x > of Somali became < q > and < c >, while the digraph < dh > for the postalveolar /dʒ/ (a rather obvious choice) became /x/. As ‘Afar (or, in the new orthography, Qafar) has no uvular stop, no new symbol for /q/ was needed. Remarkably, all the other signs of Somali were kept.

The “Djibouti” orthography is used in the ‘Afar regional state of Ethiopia alongside the Ethiopian syllabary.

In Eritrea, after independence (1991; *de jure* 1993), the languages of Eritrea have been provided with a unified, national Latin orthography from which they depart only for phonemes peculiar to single languages (Semitic languages Tigrinya, Tigre and, of course, Arabic are written, respectively, in the Ethiopic syllabary and in Arabic script). For the ‘Afar minority of Eritrea and the very similar Saho (East Cushitic; ssy) the Somali choices of the 1970’s have been implemented (plausibly in order to sever the links with the ‘Afar in other countries). In the end, three nation states have implemented two different orthographies for similar languages, with one and the same language (‘Afar becoming either Qafar in Djibouti or Cafar in Eritrea) having two different orthographies in different countries (three counting the Ethiopic syllabary).

phoneme	Somali		‘Afar (Djibouti)		‘Afar, Saho (Eritrea)		gloss of examples
ʃ	c	cad	q	qado	c	cado	“(to be) white)”
ħ	x	xaakin	c	caakim	x	xaakim	Somali “judge;” ‘Afar “governor;” Saho “doctor” (from Arabic ḥākim “ruler”)
ɖ	dh	dhal	x	xale	dh	dhale	“to give birth to”

Figure 9. Somali, ‘Afar and Saho: wildcards and political choices (adapted from Savà and Tosco 2008: 125)

The principle of using wildcards has been further implemented in Ethiopia in connection with the marking of ejectives in Oromo. A brilliant combination of the criteria of simplicity and frequency is used – supplemented in case by adherence to tradition. The overall picture is apparently puzzling but makes actually good sense:

modal		ejective	
IPA	orthography	IPA	orthography
/p/	< p >	/p'/	< ph >
/t/	< t >	/t'/	< x >
/k/	< k >	/k'/	< q >
/tʃ/	< ch >	/tʃ'/	< c >

Figure 10. Modal and ejectives in the Oromo orthography

For the rarely used (mostly in loans) bilabials, plain < p > stands for modal /p/ and the digraph < ph > for its ejective counterpart /p'/: an additional phonetic feature is paralleled by an additional graphic symbol. For the velars, < k > stands for the modal and < q > for the ejective (here following a long Orientalist and Ethiopianist tradition). Wildcards are instead used in alveolar stops and alveopalatal affricates. For the former, < t > stands for /t/ and < x > for ejective /t'/. For affricates, where < c > is a wildcard and the digraph < ch > a well-established solution for /tʃ/, frequency decides, and while < ch > is reserved to modal (and less common in Oromo) /tʃ/, simple < c > stands for its ejective counterpart /tʃ'/.

The web of motivations at play here is certainly complex; the present writer remembers that, when presented and discussed at the (first) International Symposium on Cushitic and Omotic Languages (Köln, 1986), “some modifications were recommended on the basis of phonetic consistency” (Heine 1988: 620). These recommendations – supported by most scholars (and a very young and naïve writer of these lines) – consisted basically in proposing the use of digraphs with < h > as second element for the ejectives.

Wisely, the Oromo did not pay attention to intellectuals and “experts” and stuck to their decisions. Since then, the Oromo solution has been highly influential in Ethiopia and has been followed in recent years by other Latin-based orthographies (cf. Savà and Tosco 2008). One could even say that the use of < x > for /t'/ has become a shibboleth of the new Ethiopian alphabets for a geographically and genetically diverse array of languages. While only Koorete (North Omotic; kqy) and Sidamo (East Cushitic; sid) are presented here, the examples could be multiplied.

Erunxo 23

Wora Go'unxe.



Goodiyaadani lam'i e gaditini anke wonta mixe mixhiya hangicha. Mosheni Naaqoni u sungicha. Wora go'unxe erusaxi uso erussechosha wolla zeersine u hangicha. Godiyaadey be gadita "woray abasuna go'e" hidi oyco. Moshey "woray bo'ose zawa, mooshshuna ira gooche, sahay wulqana shohutte ooxesekko e go'e!" hido. Naaqoy "woray keexuse bokkule, eexuse mixe, keemose maata, gatese dooshshe ingesukko e go'e" hido.

2. Benay be pishsharo eruxi zawakko botti yoodo.

Goodiyaadey "Aadekko" hidi e itusso baane bidzi gadhesaxe u beedo. Sijaara ushichine e shife wora e dato. Ye wodey bono maaqicho gisha tamay izziyaw woraa ayto. Tamay wora miye aytine Goodiyaadani e gaditini wolla uydo. Ye ganda yecha geriti kessi tama bayssu. Goodiyaadani e gaditani ye geriti suuzzo. Ye woraa tama datta ade aytti tufe /higge/ zawa u anso.

Figure 11. An excerpt from a page in Koorete (Koorete Erunxi Pishsharo 1992: 58)

“Maammashsha afi’rinokki coyinna maxine worroonikki waasi mittoho” yinanni.

Wona qummi assinihu gede maammashshu gosa mitteyire lowore leellishanno. Lawishshu gede “a” nni-”g” geeshsha shiqinshiri sidaamu egennonso ‘Jajja roorsi’ran-no?, Cimeessa ayirrisanno?, Loosiraanchonso cee’maleessa baxanno?, halaale baxanno, Kaphphona?, wkl. Yinanni xa’ mora dawaro ikkannore afi’rino. Hakkonni daafira maammashshu lowore afisiisa dandaanno.

Mittu affi wolu ledο afi’rino fiixooma leellishate maammashshu lowo kaa’lo ikkanno. Qoleno qaallate tiro maxaafa qixeessate hawama rakkino qaallanna insano assine tuqqinanni gari maammashshu aana leellanno. Hakkonnira maxaafa togooha anga afi’ra borrote loossa duuchchu danitera irkishsha ikkitanno.

Maammashshu kulamanno gara la’nanni woyite konni afii maammashsha calla ikkikinni qaalla duuchchate maammashshi roore yanna kaimu ikkito mitemite hasaanbannita ledο afi’riño. Mitemite hinge hasaanbanni woyite kayinni, haransinenna kaima ikkire ka’a qolle agurre shiqinshanni. Konne xaphphi assini maammashsha kulinke mannino, kaima ikkire egenne hee’renni haranse calla kulanna macciishinanni. Ikkona kayinni

XIII

Figure 12. An excerpt from a page in Sidamo (Itiyophiyu... 1990: xiii)

When you accept the principle of wildcards, there is no need to stick to the Oromo solution: < x > again, but in this case for the voiced uvular fricative /χ/, has recently been proposed by SIL for Ts’amakko (East Cushitic, Ethiopia; tsb; closely related to Gawwada). Savà (this volume) has taken this proposal over in his proposed orthography of Ongota (unclassified, Ethiopia; bxe).

3.2. BREAKING AWAY FROM TRADITION

Other orthographic uses seem to point in the same direction: a progressive liberation from the bounds imposed by traditional (European) orthographic norms.

Digraphs are traditionally treated as combination of two letters and they are alphabetized under the first element only. Thus, *church* is listed in English dictionary after *campaign* but before *cut*, and the Somali-Italian dictionary (DSI 1985) follows this principle, with, e.g., *shabeeel* ‘leopard’ after *saddex* ‘three’ but before *sug* ‘to wait.’

As the number of digraphs and wildcards as well as the use of a Latin orthography increase, the weight of tradition decreases. This is when a digraph becomes a “letter:” an autonomous, single grapheme. Already in the Somali-English dictionary (Zorc 1993) all words beginning with < dh >, < kh > and < sh > (the only digraphs of the Somali orthography) are listed separately, but still after their first element: < dh > after < d >, < kh > after < k >, and < sh > after < s >.

Moreover, in a first stage the first element only is doubled in gemination; in Somali, e.g., < ddh >, rather than < dhdh >, stands for /d̥d̥/, as in *gabaddha* ‘the girl’ (more commonly actually spelled *gabaddha*). Once perceived as single letters, each of the elements of a digraph are instead doubled in gemination, yielding, e.g., Wolaytta *geeshsha* ‘clean, pure’.

The next step follows logically: if, e.g., < sh > is no longer < s > + < h > but a brand-new autonomous symbol, the order of digraphs in the alphabet can and must change. This is what happens in recent Ethiopian dictionaries of languages using the Latin orthography, with the digraphs increasingly found all together at the end of the list, as in the Wolaytta dictionary (Tophphiyaa 1991), with < ch >, < ph >, and < sh > following in this order after < z >, and in Oromo (Mekuria 1998) with < ch >, < dh >, < ny > and < sh > (no word begins with < ph >).

Still, in capitalization the first letter only is capitalized: tradition is adhered to when it implies a simpler solution.

Even a completely different tradition in listing letters may now be accommodated. The following Table is the alphabetical chart present in a Koorete primer. Following the pattern of the Ethiopic syllabary, vowels are listed as columns and consonants as rows, for a total in Koorete of ten vowels (five short and five long) and thirty consonants. Apart from this general graphical arrangement, the order of consonants is the traditional Western one (but the very last consonant is < th >). The digraphs are particularly abundant in Koorete: < ch >, < dh >, < jh >, < ny >, < ph >, < sh >, < xh >, < dz > and < th >. They are listed after (and under) their first element (< ch > after < c >, < dh > after < d >, and so on), rather than all together at the end.

The order of vowels, too, follows the traditional Western one, with each long vowel after its short counterpart.

5.4 diizo beyta tato.

	a	aa	e	ee	i	ii	o	oo	u	uu
b	ba	baa	be	bee	bi	bii	bo	boo	bu	buu
c	ca	caa	ce	cee	ci	cii	co	coo	cu	cuu
ch	cha	chaa	che	chee	chi	chii	cho	choo	chu	chuu
d	da	daa	de	dee	di	dii	do	doo	du	duu
dh	dha	dhaa	dhe	dhee	dhi	dhii	dho	dhoo	dhu	dhuu
f	fa	faa	fe	fee	fi	fii	fo	foo	fu	fuu
g	ga	gaa	ge	gee	gi	gii	go	goo	gu	guu
h	ha	haa	he	hee	hi	hii	ho	hoo	hu	huu
j	ja	jaa	je	jee	ji	jii	jo	joo	ju	juu
jh	jha	jhaa	jhe	jhee	jhi	jhii	jho	jhoo	jhu	jhuu
k	ka	kaa	ke	kee	ki	kii	ko	koo	ku	kuu
l	la	laa	le	lee	li	lii	lo	loo	lu	luu
m	ma	maa	me	mee	mi	mii	mo	moo	mu	muu
n	n	naa	ne	nee	ni	nii	no	noo	nu	nuu

ny	nya	nyaa	nye	nyee	nyi	nyii	nyo	nyoo	nyu	nyuu
p	pa	paa	pe	pee	pi	pii	po	poo	puu	puu
ph	pha	phaa	phe	phee	phi	phii	pho	phoo	phu	phuu
q	qa	qaa	qe	qee	qi	qii	qo	qoo	qu	quu
r	ra	raa	re	ree	ri	rii	ro	roo	ru	ruu
s	sa	saa	se	see	si	sii	so	soo	su	suu
sh	sha	shaa	she	shee	shi	shii	sho	shoo	shu	shuu
t	ta	taa	te	tee	ti	tii	to	too	tu	tuu
v	va	vaa	ve	vee	vi	vii	vo	voo	vu	vuu
w	wa	waa	we	wee	wi	wii	wo	woo	wu	wuu
x	xa	xaa	xe	xee	xi	xii	xo	xoo	xu	xuu
xh	xha	xhaa	xhe	xhee	xhi	xhii	xho	xhoo	xhu	xhuu
y	ya	yaa	ye	yee	yi	yii	yo	yoo	yu	yuu
z	za	zaa	ze	zee	zi	zii	zo	zoo	zu	zuu
dz	dza	dzaa	dze	dzee	dzi	dzii	dzo	dzoo	dzu	dzuu
th	tha	thaa	the	thee	thi	thii	tho	thoo	thu	thuu

Figure 13. A Koorete alphabetical chart (Koorete Bidzunxo Suma Erunxi Pishsharo 1990: 86-87)

3.3. WILDCARDS GALORE

But why to stop at “natural” wildcards? Any unused symbol may come to good use.

We go here beyond the notion that Latin symbols for consonant clusters (as in the case of < x >) or of original allophones (< c >) are “free:” any sign that happens to be useless in the language is up for grabbing and re-use.

Nara (Nilo-Saharan or isolate; nrb) of Eritrea offers maybe the most radical solution so far, and puts into relief as well a few problems.

Most solutions found in the Latin orthography of the languages of Eritrea are not peculiar at all, others partially are: e.g., a palatal nasal is < gn > in Saho, as in Italian, rather than < ny >. Again, shunning the Oromo (and therefore, in a way, the Ethiopian solution), an ejective /t'/ is not marked by < x > but by < th > in Bilin (Central Cushitic; byn) and Saho (East Cushitic; ssy), while < ch > marks an ejective palato-alveolar affricate /tʃ/ – it was seen above that in Oromo it marks the modal. It was also seen above that Saho < c > marks the voiced pharyngeal /ʕ/ (*à la* Somali); a diacritic is therefore introduced for the modal affricate /tʃ/: < č >.

The velar nasal is of course a problem for any Latin-based alphabet: its most common rendering is < ng >, but many other solutions have been or are in use, such as Swahili and Xhosa (both Bantu; swa and xho) < ng' >, < nh > in Galician (Western Romance) and Nawat (or Pipil; Aztec), or simply < g >, as in Fijian (Austronesian).

Although absent in word-initial position the velar nasal is also phonemic in Piedmontese (Western Romance), where it is also probably more common than the alveolar nasal. The orthographic solution devised for Piedmontese is to use < n > where no ambiguity may arise and have a hyphen follow it in other cases (i.e., between vowels) yielding < n- >. This of course conflicts with hyphenation, but is consonant with the liberal use of hyphens in other points of the orthography (such as in order to separate clitics, following the French model).

Nara is most illuminating in its use of < v > for the velar nasal. Certainly, the presence of prenasalized voiced stops preempted the use of < ng > – as this digraph was chosen, quite correctly, to represent a prenasalized /^hg/. Still, < v > is a brave choice, and to the best of my knowledge unique. In their proposed orthography for Ts'amakko of Ethiopia, SIL has used likewise < v >, but for a voiced pharyngeal fricative /ʕ/, and Savà (this volume) proposes to copy this in Ongota.

<i>b</i>	/b/	Voiced bilabial stop
<i>d</i>	/d/	Voiced alveolar stop
<i>t</i>	/t/	Voiceless alveolar stop
<i>g</i>	/g/	Voiced velar stop
<i>k</i>	/k/	Voiceless velar stop
<i>mb</i>	/ ^m b/	Prenasalised voiced bilabial stop
<i>nd</i>	/ ⁿ d/	Prenasalised voiced alveolar stop
<i>ng</i>	/ ^ŋ g/	Prenasalised voiced velar stop
<i>gw</i>	/g ^w /	Labialised voiced velar stop
<i>kw</i>	/k ^w /	Labialised voiceless velar stop
<i>ngw</i>	/ ^ŋ g ^w /	Prenasalised labialised voiced velar stop
<i>f</i>	/f/	Voiceless labiodental fricative
<i>s</i>	/s/	Voiceless alveolar fricative
<i>h</i>	/h/	Voiceless laryngeal fricative
<i>sh</i>	/ʃ/	Voiceless palatal fricative
<i>ch</i>	/tʃ/	Voiceless palatal affricate
<i>j</i>	/dʒ/	Voiced palatal affricate
<i>r</i>	/r/	Voiced alveolar trill
<i>l</i>	/l/	Voiced alveolar lateral
<i>w</i>	/w/	Voiced bilabial glide
<i>y</i>	/j/	Voiced palatal glide
<i>m</i>	/m/	Voiced bilabial nasal
<i>n</i>	/n/	Voiced alveolar nasal
<i>v</i>	/ŋ/	Voiced velar nasal

Table 1. List of Nara consonant phonemes

Figure 14. The consonant phonemes of Nara (Banti and Savà 2021: 240)

Furthermore: why to stop at letters? In Ethiopia, Wolaytta (North Omotic; wal) has introduced the digit < 7 > for the glottal stop, as in *lee7iyaa* ‘thin.’ It is most commonly found reduplicated, as in *ha77i* ‘now’.⁴ As elsewhere, the phonological presence of a glottal stop is not marked in word-initial position.

This solution is not totally isolated, as it is also found in Squamish (Coast Salish) of British Columbia (whether it was consciously copied from Squamish is unknown to the present writer). It is still apparently isolated in Ethiopia, where the apostrophe < ’ > is preferred. The two solutions are shown here through the initial page of the *Book of Hosea* in Wolaytta (*Hosee7a*) and Oromo (*Hose’aa*).

⁴ Examples are from the Wolaytta-Amharic dictionary (*Tophphiya...* 1991). The English translations are the most common meanings of the Amharic entries.

< 7 > is certainly more conspicuous – and therefore less likely to be forgotten in casual writing – than the apostrophe. It is obvious that language-internal considerations – such as frequency and relevance in morphological processes – will have to be taken into account in the selection.

**Hananabaa Yootiya
Hoose7a Maxaafaa**

Gidduwaa Erissuwaa

Hananabaa yootiya Hoose7i Samaaree Kiristtoosa yeletaappe kase laappun xeetanne haatamanne isiini laitann kunddanaappe kasetidi, hanotati meto gidi uttido wodiyan, hananabaa yootiis; i hananabaa yootidoogee hananabaa yootiya Amoxappe guyye gidishin, huuphessa bagga Israa7eela kawotettan yootiis. Hoose7i keehi darissidi odidoogee asai eequ goinnidoogaanne Xoossaa ammaniyo ammanuwaa pacissidoogaa. Hoose7i haggaa asai Xoossaassi ammanettennan aggiyoogaa qonccissidi odanau, issi ammanettenna maccaasiyo machcho oottidi ekkidoogaa leemiso oottidi go7ettees. A machchiyaa Goomera ayyo ammanettennan ixkidoogaadan, Xoossaa asaikka Xoossaa aggi bayiis. Hegaa gaasuwan Israa7eela asaa bolli pirdai yaana hanees. Gidikkomne wurssettan Xoossai ba asaa siiqiyo siiqoi dariyo gishshau, ba asaa baakko zaaridi, baappenne ba asaappe gidduwan de7iya siiquwaa ooraxissee. A gita siiqoi hagaappe kaaliya haasayan qonccii; “Israa7eela asatoo, taani inttena waata oloo? Taani inttena waata aattada immoo? Taani inttena siiqiyo siiqoi gita gidido gishshau, taani haggaa ubhaa oottanau ta wozanai eenenna” (11:18) yaagiis.

1 ¹Ooziyaani, Yo7aataami, Akaazinne Hizqiyaaasi kaalli kaallidi Yihudan kawotido wode, qassi Yo7aasha na7ai Iyorbbaami Israa7eelan kawotido wode, Bi7eera na7aa Hoose7akko yiida GODAA qaalai haggaa.

Hoose7a Maccaasiyoonne Naata

²GODAI koiro Hoose7a baggaara haasayido wode, Hoose7a, “Ba; baada shaaramuxa maccaasiyo neeyyo akka; ekkada shaaramuxeppe naata yela. Aissi giikko, shaaramuxiya maccaasiyaa ba azinaappe shaahettiyoogaadan, ta asai taappe shaahettiis” yaagiis.

³Yaagin biidi, Dibilaima na7iyoo Goomero ekkii; ekin shahaarada, attuma na7a au yelaasu. ⁴Yaatin, GODAI Hoose7a, “Na7aa sunntaa ‘Izira7eela’ yaagada sunnta. Aissi giikko, Izira7eelan gukkida suunntaa autiyau taani Yeehu zareta qantta wodiyan achchana; Israa7eela kawotettaakka xaissana. ⁵He gallassi taani Izira7eela Wombban Israa7eela olanchchata mentterettana” yaagiis.

⁶Goomera naa77antuuwaa shahaarada, macca na7iyoo yelaasu. Yaatin, GODAI Hoose7a, “I sunntaa ‘Loruhaamo’” yaagada sunnta. Aissi giikko, hagaappe

sinttanau taani Israa7eela asuu qarettikke; etau mulekka atto giikke. ⁷SHin taani Yihudaa asaassi qarettana; taani GODAN, eta Xoossan, eta ashshanaappe attin, wonddafiiyan woikko bisuwan woikko olan woikko paratun woikko paraasatun ashshikke” yaagiis.

⁸Goomera Loruhaamo xanttaa duuttidoogaappe guyyiyan, harantuwwaa shahaarada, attuma na7aa yelaasu.

⁹Yaatin, GODAI Hoose7a, “A sunntaa ‘Lo7aama’” yaagada sunnta. Aissi giikko, intte ta asa gidekketa; taanikka intte Xoossaa gidikke.

Israa7eela Asai Ooraxxana

¹⁰“Israa7eela asai makkanaunne qoodanau danddayettenna abbaa shafe keena gidana. Etau, ‘Intte ta asa gidekketa’ geetetti odettidosan, ‘Intte de7o Xoossaa naata’ geetettana. ¹¹Yihudaa asainne Israa7eela asai issippee shiiqana; shiiqidi banttau issi halaqaa doorana; dooridi omoodo biittaappe simmananne ishalidi de7ana. Aissi giikko, Izira7eela gallassai gita gallassa gidana.

2 ¹⁴“Intte ishannta, ‘GODAI inttena, ‘Ta asa’ yaagees’ yaagite. Qassi intte

¹ 1:6 Loruhaamo: “Loruhaamo” giyoogee “Qarettennaaro” giyoogaa. ² 1:9 Lo7aama: “Lo7aama” giyoogee “Ta asa gidenna” giyoogaa.

Figure 15. Hosea 1:10-11 in the Wolaytta Bible (Geeshsha Maxaafaa 1996: 881)

Macaafa Hose'aa Raajichaa

Ittiin Lixa

1 ¹ Kun dubbii Waaqayyoo isa bara Uziyaan, Yotaam, Ahaaz, Hisqiyaasis tarreetti biyya Yihudaa irratti mo'anii, Yerobi'aam ilmi Yeho'aash immoo biyya Israa'el irratti mo'etti, gara Hose'aa ilma Beriit dhufee dha*.
^{*1} Isa 1:1; ²Mot 14:23-29; ^{Amo} 1:1

Hose'aan Fuudhee Ijoollee Sadii Godhachuu Isaa

2 Waaqayyo dura karaa Hose'aa yeroo dubbatetti, Hose'aadhaan, "Biy-yichi Waaqayyoon dhiisuu isaatiin eja guddaa keessa lixcera; kanaaf dhaqi, dubartii ejjituu fuudhi, ijoollee ejaas godhadhu!" jedhe*.
^{*2} 2:4-10; 3:1; 4:11-19; 5:3-4; 9:1-2; 11:2; ^{Er} 2:20; ^{Kes} 23:18

3 Kana irratti inni dhaqee Gomerin intala Diiblaayim fuudhe; isheen ulfoofttee ilma in deesseef. **4** Yommus Waaqayyo, " 'Yizre'el' jedhii moggaasi! Yeroo gabaabduu booddee sababii dhiiga Yizre'elitti dhangala'eeff mana Yehuu nan adaba; mootummaan Israa'elis kanumaan akka raawwatu nan godha*." **5** Gaafas ani dachaa Yizre'el keessatti humna warra Israa'el nan cabsa" jedheen.
^{*4} 2Mot 10:10-11

6 Gomer ammas ulfoofttee, durba in deesse; Waaqayyo immoo Hose'aadhaan, "Ani si'achi mana Israa'eliif cubbuu isaaniif isaaniif dhiisuuf matumaa oo'a hin argisiisu, kanaaf 'Lo-Ruhaamaa' jedhii maqaa moggaasi! [Hiikaan isaas 'Oo'a hin arganne' jechuu dha]*." **7** Mana Yihudaatti garuu oo'a nan argisiisa, isaan nan oolchas; iddaadhaan yookiis billaadhaan yookiis lolaan yookiis fardeenii fi abboota fardeenii utuu hin ta'in, ani Waaqayyo gooftaan isaani harka kootiin isaan nan oolcha" jedhe*.

^{*6} 2:3,25

^{*7} Mik 5:9; Zak 4:6; Isa 31:1-3; Far 20:7; 2Mot 19:32-37

8 Gomer erga "Lo-Ruhaamaa" harma guuftee, ulfoofttee ilma in deesse. **9** Waaqayyo yommus Hose'aadhaan, " 'Lo-Amii' jedhii moggaasi, isinoo saba koo miti, anis Waaqayyo keessan

miti" jedhe. ["Lo-Amii" jechuun "Saba koo miti" jechuu dha*].
^{*9} 2:25; Bau 3:14; Er 7:23

Abdii Gara Fuula Duraa

2 ¹ Si'achi garuu lakkoobsi namoota Israa'el akka cirracha galaanaa isa hin safaramnee, isa hin lakkaa'annees jedhame immoo, "Ijoollee Waaqayyo isa jiraataa" jedhamuuf jiru*.

² Ijoolleen Yihudaa fi ijoolleen Israa'el tokkummaatti walitti in qabamu; walumattis nama isaan geggeessu tokko in kaafatu; guyyaan Yizre'el guddaa waan ta'uuf, isaanis biyyicha keessaa ol in dhufu*.

³ Egaa obboloota keessan, "Amii" obboleettota keessanis "Ruhaamaa" jedhaa waamaa! ["Amii" jechuun "Saba koo", "Ruhaamaa" jechuun "Oo'a argate" jechuu dha*].

^{*1} Uma 22:17; Rom 9:29

^{*2} Er 3:18; Hos 1:4; 2:23-24

^{*3} 1:6,9

Amanamuu Dhabuu Irraa Kan Ka'e Adabamuu

4 Haadha keessan hadheessaa! Isheen haadha manaa koo ta'uu waan dhiifteef, anis abbaa manaa ishee ta'uu waanan dhiiseef, ishee hadheessaa! Isheen halalummaa fuula ishee irraa mul'atu, milikkita ejjumaas harma ishee gidduudhaa haa baafu*!

5 Yoo kun ta'uu dhaabaate garuu, wayyaa ishee irraa baasee akkuma gaafa dhalatteetti qullaa nan hambisa; ishees akka lafa onaa, akka lafa isa gogaa nan godha, dheebuudhaan akka isheen duutu nan godha*.

Figure 16. Hosea 1 in the Oromo Bible (Macaafa Qulqulluu: 1118)

4. A FEW CONCLUSIONS

By definition, to use a basic Latin keyboard only excludes graphic iconization (Sebba 2015), i.e., no “icon” similar to Danish < ø > or Spanish < ñ > may arise (although the use of < 7 > in Wolaytta could come very close to be an “icon”). What is possible instead is the idiosyncratic association of a grapheme to a phoneme, as repeatedly shown in this article, and with Somali < c > and Oromo < x > being maybe the most striking cases.

Second, the case of ‘Afar (Qafar) has shown that it is difficult to outsmart wildcards: in comparison to < c >, < q > is much less of a wildcard, and it has a strong association with a uvular stop or it simply marks a back allophone of /k/ (as it was in Latin and is still generally the case in modern European languages). < c >, on the contrary, can retain its Latin value as a velar stop (as in Romance languages with a non-front vowel following) and have different values with other vowels, or still be an alveolar affricate (as in Croatian, Slovenian, Polish, Czech, Slovak, Hungarian...) or many other things. And it can of course become much else in digraphs or with diacritics. Neither a place nor a manner of articulation is strictly linked to it, synchronically.

The fact that no language, to the best of my knowledge, has followed ‘Afar in using < q > for a pharyngeal fricative or has given it yet another value than /q/ is certainly due to the imperfect ausbauization of ‘Afar: in Djibouti, French and Arabic are the only official languages, and both Somali and ‘Afar are conspicuous for their absence from the linguistic landscape. But I venture to say that there is something inherently awkward in making a wildcard out of a card that is not. In exaptation you turn into use what you have and is available; it is certainly more difficult (but not impossible) to change the use of a more or less functional tool.

Third, it is also apparent that wildcards are second bests. This strategy seems to be always secondary to the use of digraphs; e.g., /ʃ/ is always expressed by < sh > and an implosive /ɗ/ or postalveolar /ɗʲ/ is generally < dh >.

The use of digraphs is particularly shunned for vowels, except in the marking of length and notwithstanding the wide use of vocalic digraphs in many European writings.

Finally, maybe the most important – and saddest – conclusion is that many African orthographies were in a way born old: they are utterly incapable of being brought to use in the most modern technologies – pending financial investments (in developing, implementing and marketing keyboards) that the communities cannot sustain.

REFERENCES

- Abdelhay, Ashraf, Makoni, Busi & Makoni, Sinfree. 2016. "The colonial linguistics of governance in Sudan: The Rejaf Language Conference", 1928 in: *Journal of African Cultural Studies* 28 (1): 1-16.
- Banti, Giorgio & Savà, Graziano 2021. "Nara Phonology" in: *Rassegna di Studi Etiopici* 5 (3^a Serie): 235-268.
- Caguor Adḡ Manyaj. 2003. *Anyikööl ko kuḡny piööc ë dēt 1*. Khartoum: Sudan Language Programme, Dinka Rek Language Committee.
- Cahill, Michael. 2015. "Non-Linguistic Factors in Orthographies" in: Cahill, Michael & Rice, Keren (eds.), *Developing Orthographies for Unwritten Languages*. Dallas: SIL International: 9-26.
- Cahill, Michael & Karan, Elke. 2008. *Factors in designing effective orthographies for unwritten languages*. SIL International: SIL Electronic Working Papers (https://www.sil.org/system/files/reapdata/14393230040163130285698/silewp2008_001.pdf).
- Caney, John Charles. 1984. *The Modernisation of Somali Vocabulary, with particular reference to the period from 1972 to the present*. Hamburg: Helmut Buske.
- Dizionario Somalo-Italiano*. 1985. Roma: Cooperazione Italiana allo Sviluppo.
- Geeshsha Maxaafaa*. 1996 (2002). Addisaaba [Addis Ababa]: Toophphiyaa Geeshsha Maxaafaa Maabaraa [The Bible Society of Ethiopia].
- Heasty, Rev. J. A. 1937 (reprinted 1974). *English-Shilluk, Shilluk-English Dictionary*. Dolieb Hill: The American Mission.
- Heine, Bernd. 1988. "Report on the Oromo Working Group on Orthography" in: Bechhaus-Gerst, Marianne & Serzisko, Fritz (eds.). *Cushitic-Omotiic. Papers from the International Symposium on Cushitic and Omotic Languages. Cologne, January 6-9, 1986*. Hamburg: Helmut Buske: 619-621.
- Itiyophiyu Qaalla Xiinxallo Uurrinsha [Ethiopian Languages Research Center]. 1990 (1998).¹ *Sidaamu Afii Maammashsha*. Addis Ababa: Addis Ababa Yuniversiite [Addis Ababa University].
- Jaggar, Philip J. 2001. *Hausa*. Amsterdam: John Benjamins.
- Jujumbu Kendya ko Bari Buk Tomusala* [Learning to Read in Bari Book 3]. 1999. Nairobi: Institute of Regional Languages.
- Koorete Bidzunxo Suma Erunxi Pishsharo*. 1990 (1998).

¹ The publication year found in Ethiopian books is often in accordance with the local calendar (*Enkutataš*). It is followed here by the corresponding year in the Gregorian calendar.

- Koorete Erunxi Pishsharo. 1992 (2000). *Lakkunxo suma*. Awaasa.
- Kootz, Anja & Pasch, Helma (eds.). 2010. *5000 Jahre Schrift in Afrika*. Special Issue of *Afrikanistik Online* (<https://www.afrikanistik-aegyptologie-online.de/archiv/2010>).
- Kuëne ke Thok Nath Bok Ŋuaan. 1994. Maridi: Institute of Regional Languages.
- Labahn, Thomas. 1982. *Sprache und Staat: Sprachpolitik in Somalia*. Hamburg: Helmut Buske.
- Laitin, David D. 1977. *Politics, Language, and Thought. The Somali Experience*. Chicago: University of Chicago Press.
- Laitin, David D. 1992. *Language Repertoires and State Construction in Africa*. Cambridge: Cambridge University Press.
- Macaafa Qulqulluu. 1997. Addis Ababaa [Addis Ababa]: Waldaa Macaafa Qulqulluu Itoophiyaa [The Bible Society of Ethiopia]
- Mekuria, Hinsene. 1998 (2006). *Galmee Jechoota Afaan Oromoo-Amaaraa-Inglizii. Oromo-Amharic-English Dictionary*. Finfinnee [Addis Ababa]: Commercial Printing Enterprise.
- Micheli, Ilaria. forthcoming. "Limits and potential of Dyula in Burkina Faso: observations for use in cooperation" in: Serenella Baggio (ed.), *Lingua franca, lingue franche*. Alessandria: dell'Orso.
- Mumin, Meikal & Versteegh, Kees (eds.). 2014. *The Arabic Script in Africa: Studies in the Use of a Writing System*. Leiden: Brill.
- Newman, Paul. 2000. *The Hausa language: an encyclopedic reference grammar*. New Haven and London: Yale University Press.
- Owen, R. C. R. 1908. *Bari Grammar and Vocabulary*. London: J. & E. Bumpus.
- Persson, Andrew M. & Perrson, Janet R. 1991. *Mödö-English Dictionary with Grammar*. Nairobi: Summer Institute of Linguistics – Sudan.
- Savà, Graziano & Tosco, Mauro. 2008. "Ex Uno Plura: The uneasy road of the Ethiopian languages towards standardization" in "Ausbau and Abstand Languages: Traditional and New Approaches." *International Journal of the Sociology of Language* 191: 111-139.
- Sebba, Mark. 2015. "Iconisation, attribution and branding in orthography" in: *Written Language & Literacy* 18 (2): 208-227. <http://dx.doi.org/10.1075/wll.18.2.02seb>
- Sgall, Petr. 1987. "Towards a Theory of Phonemic Orthography." in: Luelsdorff, Philip A. (ed.), *Orthography and Phonology*. Amsterdam & Philadelphia: John Benjamins: 1-31.
- Tophphiyaa Doonatu Xinaatiyaanne Pilggettaa Ooso Keettaa [Center for the Study and Research of the Ethiopian Languages]. 1991 (1999). *Wolayttatto Qaalatu Amaaratto Birshshettaa* [Wolaytta-Amharic Dictionary]. Addisaaba [Addis Ababa]: Addisaaba Yuniversttiyaa [Addis Ababa University].
- Tosco, Mauro. 2001. *The Dhaasanac Language*. Köln: Rüdiger Köppe.
- Tosco, Mauro. 2008. "Introduction: Ausbau is everywhere!" in: "Ausbau and Abstand Languages: Traditional and New Approaches." *International Journal of the Sociology of Language* 191: 1-16.
- Tosco, Mauro. 2010. "Somali writings" in: Kootz & Pasch 2010 (<https://www.afrikanistik-aegyptologie-online.de/archiv/2010/2723>).
- Tosco, Mauro. 2014. "The tension between language diversity and state building in Africa" in: *Incontri Linguistici* 37: 105-116.
- Tosco, Mauro. 2015. "Short notes on Somali previous scripts" in: Cabdirashiid M. Ismaaciil, Cabdalla C. Mansuur & Saynab A. Sharci

- (eds.), *Afmaal. Proceedings of the Conference on the 40th Anniversary of Somali Orthography. Djibouti, 17th – 21st December 2012*. Djibouti: AGA (The Intergovernmental Academy of Somali Language): 189-217 [abridged version of: Tosco 2010].
- Tucker, Archibald N. 1929. "Report of the Rejaf Language Conference, 1928" in: *Africa: Journal of the International African Institute*, Vol. 2/1: 95.
- Warage O Miliny Murleye [Little Book. Pre-Primer Murle]. 2004. Entebbe: Murle Translation and Literacy Project.
- War'gat Markoká [The Gospel of Mark in Dhaasanac]. 1997. Nairobi: Bible Translation and Literacy (E.A.) and Wycliffe Bible Translators.
- Werner, A. 1929. "Report of the Rejaf Language Conference, 1928" in: *Bulletin of the School of Oriental Studies* 5/2: 426-428.
- Wori Haaggane MARKO Khore. 1993 [*The Gospel of Mark in Rendille*]. Nairobi: Bible Translation and Literacy (E.A.).
- Zorc, R. David & Madina M. Osman. 1993. *Somali-English Dictionary with English Index*. Kensington, MD: Dunwoody Press.

Part 2

Lexicography and Didactics of L2

L'italiano come seconda lingua nei CPIA (Centri Provinciali per l'istruzione degli Adulti): Questioni e Prospettive

EMILIO PORCARO
CPIA Bologna

ABSTRACT

In recent years, the adult education system has undergone a major reform that has led to the birth of the Centers for Adult Education (CPIA), autonomous educational institutions whose training offer is aimed at Italian and non-Italian adults and young adults. The present article aims to illustrate the role of CPIAs in the processes of linguistic and social inclusion and integration of foreigners. The issue requires particular attention as it poses continuous challenges, the most significant of which is the acquisition of a status of active citizenship by a large number of foreign citizens. Knowledge of the Italian language and culture, an essential tool for facilitating the integration process in the host community and participation in school courses for adults, also represents an indicator of the success of the migration process and the ability for immigrants to integrate professionally and socially in society.

KEYWORDS

CPIA; adult education; lifelong learning; Italian as a second language.

QUESTIONI

Nell'ultimo quindicennio, l'intensificarsi dei flussi migratori internazionali ha determinato un significativo aumento della popolazione straniera in Italia le cui conseguenze hanno inciso profondamente sull'intero sistema dell'istruzione e della formazione dalla scuola dell'infanzia fino alla scuola degli adulti. Parimenti, le novellazioni normative in tema di immigrazione, come ad esempio l'introduzione del requisito della conoscenza della lingua italiana per ottenere i documenti di soggiorno nonché per aspirare alla cittadinanza, hanno avuto forti ricadute sul sistema scolastico rivolto alla popolazione adulta. Alla data del 1° gennaio 2019 la popolazione straniera residente in Italia risultava pari a 5.144.440 con un incremento di 140.000 unità rispetto al 2015 e un'incidenza dell'8,2% sulla popolazione residente, mentre alla stessa data il numero degli ingressi di cittadini non comunitari ammontava a 262.770 unità¹. I paesi di provenienza risultano essere oltre 190, mentre i primi cinque paesi per numerosità sono la Romania, il Marocco, l'Albania, la Cina e l'Ucraina².

I provvedimenti emanati in diverse occasioni dal Ministero dell'Istruzione (circolari, note, Linee guida) hanno avuto l'indubbio merito di sostenere le istituzioni scolastiche nelle scelte educative, didattiche e organizzative volte a garantire agli studenti dei paesi terzi il diritto all'istruzione e alla piena inclusione culturale e sociale³.

La prospettiva interculturale, entrata gradualmente nelle pratiche didattiche delle scuole di ogni ordine e grado a partire dai primi anni novanta del

¹ Annuario statistico italiano 2019, Istat, pag. 100.

² Idos-UNAR, Dossier Statistico Immigrazione 2015, Idos Roma, 2015.

³ Tra le circolari specifiche su questo tema ricordiamo: 1) la CM n. 24 del 1 marzo 2006 che trasmette le Linee guida per l'accoglienza e l'integrazione degli alunni stranieri: «la presenza di alunni stranieri è un dato strutturale e riguarda tutto il sistema scolastico. È necessario, dunque, individuare le migliori pratiche e disseminarle nel rispetto del Piano dell'offerta formativa e dell'autonomia scolastica, d'intesa con gli Enti locali e gli altri soggetti che sul territorio interagiscono per l'integrazione». Emerge un modello italiano di integrazione scolastica basato su quattro principi chiave: l'universalismo, la scuola comune, la centralità della persona in relazione con l'altro, l'intercultura. 2) la CM n. 4233 del 19 febbraio 2014 con la quale vengono trasmesse le rivisitate Linee guida per l'accoglienza e l'integrazione degli alunni stranieri che affrontano secondo una prospettiva ancora più ampia e articolata la complessità del fenomeno degli studenti stranieri a scuola, fornendo indicazioni e strumenti di lavoro per individuare le modalità con le quali affrontare ciascuna situazione nella «consapevolezza che lo studente di origini straniere può costituire un'occasione per ripensare e rinnovare l'azione didattica a vantaggio di tutti, un'occasione di cambiamento per tutta la scuola».

secolo scorso⁴, è diventata in breve tempo il paradigma di riferimento nei processi di insegnamento e apprendimento, nella didattica, nei curricula, nelle relazioni intra ed extra scolastiche: «la via italiana all'intercultura unisce alla capacità di conoscere ed apprezzare le differenze la ricerca della coesione sociale, in una nuova visione di cittadinanza adatta al pluralismo attuale, in cui si dia particolare attenzione a costruire la convergenza verso valori comuni⁵».

Negli stessi anni altri due importanti e significativi provvedimenti aggiungono un ulteriore elemento di attenzione nei riguardi degli alunni e degli studenti giovani e adulti stranieri: la riforma, nel 2012, del sistema di Istruzione degli Adulti che ha portato alla nascita dei CPIA - Centri provinciali per l'Istruzione degli Adulti e l'istituzione, nel 2016, della classe di concorso per insegnare italiano come seconda lingua A023 – lingua italiana per discendenti di lingua straniera⁶.

Il 2012 è stato un anno significativo per l'*Adult education* in Italia. Due provvedimenti legislativi hanno contribuito al rinnovamento dell'intero sistema di *lifelong learning*: l'istituzione, grazie alla Legge 92/2012, del sistema nazionale dell'Apprendimento Permanente e la creazione dei Centri per l'Istruzione degli Adulti (CPIA). Questi ultimi hanno avuto, e tuttora continuano a svolgere, un ruolo centrale per quanto riguarda l'apprendimento della lingua italiana come L2⁷. Al momento in cui scriviamo oltre 250.000 utenti di cittadinanza non italiana (circa l'80% dell'utenza complessiva iscritta ai CPIA) frequentano i percorsi organizzati dai 130 CPIA attivi in Italia.

L'aspetto che maggiormente caratterizza l'impianto organizzativo e didattico del rinnovato sistema di Istruzione degli Adulti è il conferimento dell'autonomia scolastica: i CPIA costituiscono una istituzione scolastica autonoma, sono dotati di un proprio assetto organizzativo e didattico, sono articolati in reti di servizio e operano in stretta connessione con i soggetti pubblici e privati del territorio nel quale sono collocati. Ai CPIA si iscrivono gli adulti e

⁴ È con la CM 205 del 26 luglio 1990 che prende corpo il principio del coinvolgimento degli alunni italiani in un rapporto interattivo con gli alunni immigrati in funzione di un reciproco arricchimento.

⁵ Ministero della Pubblica Istruzione, *La via italiana per la scuola interculturale e l'integrazione degli alunni stranieri*, Roma, 2007, pag. 9.

⁶ La classe di concorso A23 è stata istituita con DPR n. 19 del 14 febbraio 2016 – Regolamento recante disposizioni per la razionalizzazione ed accorpamento delle classi di concorso a cattedre e a posti di insegnamento, a norma dell'articolo 64, comma 4, lettera a), del decreto-legge 25 giugno 2008, n. 112, convertito, con modificazioni, dalla legge 6 agosto 2008, n. 133 – pubblicato sulla Gazzetta Ufficiale n. 43 del 22 febbraio 2016 – S. O. n. 5.

⁷ I CPIA vengono istituiti con DPR 263 del 29 ottobre 2012 – Regolamento recante norme generali per la ridefinizione dell'assetto organizzativo didattico dei Centri d'istruzione per gli adulti, ivi compresi i corsi serali, a norma dell'articolo 64, comma 4, del decreto-legge 25 giugno 2008, n. 112, convertito, con modificazioni, dalla legge 6 agosto 2008, n. 133, pubblicato nella Gazzetta Ufficiale n. 47 del 25/02/2013.

i giovani adulti italiani e stranieri che abbiano compiuto il sedicesimo anno di età, che siano privi del titolo conclusivo del primo ciclo di istruzione o che non abbiano assolto all'obbligo di istruzione. In deroga al vincolo del sedicesimo anno si possono iscrivere ai percorsi di primo livello anche i Minori stranieri non accompagnati (MSNA) purché abbiano compiuto quindici anni.

La formulazione "italiano come seconda lingua" non compare mai nei dispositivi normativi di fonte legislativa. Gli ordinamenti dell'istruzione degli adulti utilizzano l'espressione percorsi di "Alfabetizzazione e Apprendimento della Lingua Italiana" (AALI) destinati agli adulti stranieri e finalizzati al conseguimento del titolo attestante il raggiungimento del livello A2 di conoscenza della lingua italiana del Quadro comune europeo di riferimento per le lingue elaborato dal Consiglio d'Europa. Le *Linee guida per il passaggio al nuovo ordinamento* del 2015 definiscono l'assetto didattico: i percorsi AALI sono articolati in due livelli A1 e A2; ciascun livello è finalizzato a sviluppare cinque ambiti di competenza: ascolto, lettura, interazione orale e scritta, produzione orale, produzione scritta⁸. L'intero percorso AALI ha una durata complessiva di 200 ore di cui 180 ore destinate alle attività didattiche e 20 ore destinate ad attività di accoglienza, orientamento e/o di messa a livello. I docenti impiegati per la realizzazione di tali percorsi sono maestri di scuola primaria posto comune e, in alcuni casi, docenti della classe di concorso A023. Tenuto conto che il livello A2 è indispensabile ai fini dell'assolvimento degli adempimenti previsti dalle novità in materia di immigrazione introdotte dalla Legge n. 94/2009, al fine di favorire quanto più possibile i processi di integrazione sociale ai percorsi di alfabetizzazione e di apprendimento della lingua italiana, possono iscriversi anche gli adulti con cittadinanza non italiana in età lavorativa in possesso di titoli di studio conseguiti nei Paesi di origine.

La tipica *classe* dei percorsi AALI, composta prevalentemente da migranti adulti, è plurilingue, multiculturale, eterogenea, intrisa di storie personali e problematiche spesso legate al processo migratorio. Soggetti diversi, ognuno con bisogni e attese differenti legati ora alla sopravvivenza, ora all'integrazione linguistica, al lavoro, alla regolarizzazione della propria posizione giuridica, all'età, al genere, alle condizioni di partenza, di scolarità e competenza alfabetica.

Un elemento che caratterizza gli apprendenti stranieri di immigrazione più recente è la loro diffusa disomogeneità quanto a livelli della scolarità esperita nei Paesi di origine. Soprattutto fra le donne immigrate per ricongiungimento familiare, fra i minori stranieri non accompagnati (MSNA) e i

⁸ Decreto Interministeriale MIUR – MEF del 12 marzo 2015 recante *Linee guida per il passaggio al nuovo ordinamento a sostegno dell'autonomia organizzativa e didattica dei Centri provinciali per l'istruzione degli adulti*, pubblicato in data 8 giugno 2015, sulla Gazzetta Ufficiale S.G. n. 130 - Suppl. Ord. n. 266.

richiedenti asilo si ritrovano situazioni di analfabetismo in lingua madre o di scarsa scolarizzazione. Ciò ha messo in evidenza la necessità di dedicare maggiore attenzione a queste nuove tipologie di bisogni linguistici. Nel 2018 il Consiglio d'Europa ha dato una prima risposta all'esigenza sempre crescente di insegnanti e formatori che si trovano spesso a dover insegnare a studenti scarsamente alfabetizzati o totalmente analfabeti mettendo a disposizione risorse per la creazione di strumenti e *framework* specifici. Nel nuovo Quadro comune europeo di riferimento, meglio noto come *Companion Volume*⁹ sono stati inseriti dei descrittori per il livello precedente all'A1. Anche i quattro Enti certificatori attualmente attivi in Italia - Università per Stranieri di Siena, Università per Stranieri di Perugia, Società Dante Alighieri e Università Roma Tre - hanno elaborato un sillabo per il livello precedente all'A1 seguendo le indicazioni del QCER e del nuovo *Companion Volume. Il Sillabo per la progettazione di percorsi sperimentali di alfabetizzazione e apprendimento della lingua italiana a livello Pre-A1*¹⁰ raggruppa i migranti adulti analfabeti funzionali e/o a debole scolarità in quattro profili di alfabetizzazione¹¹

Gruppo A. Migranti adulti che non hanno ricevuto una formazione adeguata nel loro Paese di origine, la cui lingua madre non è generalmente scritta o non è oggetto di insegnamento nel medesimo Paese. Alcuni individui di questo gruppo non hanno sviluppato l'idea di scrittura come sistema semiotico, portatrice di significato. Gli appartenenti a tale gruppo possono essere definiti pre alfabeti

Gruppo B. Migranti adulti che non hanno mai imparato a leggere e scrivere nella loro lingua madre. Nell'ambito di questo gruppo ulteriori distinzioni devono essere effettuate sulla base del sistema di scrittura della L1, nonché in relazione alla distanza tipologica tra la stessa e la L2. Gli appartenenti a tale gruppo possono essere definiti analfabeti.

Gruppo C. Migranti adulti che hanno ricevuto un'istruzione limitata nella loro lingua madre (in generale, meno di 5 anni). Gli appartenenti a tale gruppo possono essere definiti scarsamente scolarizzati o semialfabeti. Rientrano in tale gruppo coloro i quali non sono in grado di utilizzare la letto-scrittura nella maggior parte delle situazioni quotidiane, anche se riescono comunque a leggere o scrivere ad esempio parole isolate: ricordando la definizione dell'Unesco, sono considerati analfabeti. Vengono inclusi in tale gruppo anche i migranti adulti che hanno in parte perso le abilità di alfabetizzazione

⁹ *Common European Framework of Reference for Languages. Companion Volume with new descriptors*, Council of Europe, 2018.

¹⁰ *Il Sillabo per la progettazione di percorsi sperimentali di alfabetizzazione e apprendimento della lingua italiana a livello Pre-A1* è stato curato dagli Enti certificatori dell'italiano L2 con la collaborazione di Elisabetta Bonvino, Sara Di Simone, Francesca Giardini, Silvia Giugni, Giuliana Grego Bolli, Lucilla Lopriore, Eleonora Luzi, Paola Masillo, Costanza Menzinger, Sabrina Machetti, Anna Pompei e Lorenzo Rocca.

¹¹ Sillabo pre-A1, *cit.*, pagg. 5-6.

per mancanza di utilizzo della letto-scrittura, vivendo pertanto uno stato di analfabetismo di ritorno.

Gruppo D. Migranti adulti alfabetizzati: tale utenza differisce in maniera sostanziale dai tre precedenti profili in quanto con essa il percorso formativo può concentrarsi fin dall'inizio sull'apprendimento linguistico, comprendendo anche task di letto-scrittura ed individuando come obiettivi in uscita principalmente quelli stabiliti sulla base dei livelli del QCER.

Una tipologia di utenza numericamente importante nei CPIA è costituita dai richiedenti protezione internazionale, ospitati presso le strutture di accoglienza pubbliche e private sparse su tutto il territorio italiano, da nord a sud. Con riferimento agli studenti richiedenti protezione internazionale preme segnalare un fattore di criticità insito nella relazione CPIA/strutture di accoglienza e che si sostanzia in tre gradi differenti: a) è neutro laddove tale relazione rimane a un livello essenzialmente burocratico (ad es. limitata alla fase iniziale del procedimento di iscrizione); b) è ottimale laddove gli insegnanti e gli educatori delle strutture di accoglienza si pongono come figure ponte tra scuola ed extrascuola, tra CPIA e strutture di accoglienza, si manifesta un interesse reciproco rispetto ai processi educativi e di apprendimento, si definiscono momenti periodici di confronto reciproco durante l'anno scolastico; c) è pienamente efficace laddove vengono formalizzati accordi e protocolli che prevedono l'integrazione tra attività scolastiche ed extrascolastiche, la co-progettazione di percorsi individualizzati, la valorizzazione di risorse e attività complementari e di sostegno al processo formativo.

Per quanto riguarda gli strumenti di accompagnamento e di sostegno, i CPIA hanno prodotto e sviluppato numerosi strumenti e dispositivi funzionali alla didattica dell'italiano come seconda lingua nonché al miglioramento organizzativo. In ragione delle finalità possiamo definire la seguente macroclassificazione degli strumenti maggiormente utilizzati:

- a. strumenti amministrativi;
- b. strumenti per l'accoglienza;
- c. strumenti per il monitoraggio e per la valutazione.

Riportiamo nella tabella seguente, a titolo puramente esemplificativo, alcuni strumenti riconducibili alle categorie indicate:

Strumenti amministrativi	Domanda di iscrizione Informativa privacy
Strumenti per l'accoglienza	Protocollo di accoglienza Scheda di osservazione iniziale Traccia di intervista Test di ingresso di conoscenza della lingua italiana Fascicolo personale dello studente/Dossier Prove per valutare in ingresso il possesso di saperi e competenze nelle discipline non linguistiche Profilo dello studente Attestazione di riconoscimento dei crediti
Strumenti di monitoraggio e di valutazione	Patto Formativo Individuale/Piano Didattico Personalizzato Prove di verifica degli apprendimenti (iniziali, periodiche, finali) Schede di valutazione degli apprendimenti

SCHEDA DI INFORMAZIONE INIZIALE

Si tratta di una scheda che viene redatta dai docenti preposti all'accoglienza nel periodo successivo all'iscrizione amministrativa. Contiene informazioni diverse a seconda che si tratti di minori o di adulti. Nel caso dei minori la scheda iniziale è generalmente articolata nelle seguenti aree: a) dati anagrafici, b) anamnesi familiare, c) rete sociale (parenti e conoscenti significativi), d) scolarità pregressa, e) conoscenze linguistiche (lingua madre, lingue straniere conosciute), f) valutazione delle competenze in ingresso (area linguistica, area logico-matematica, area artistico-espressiva), g) interventi personalizzati da proporre (mediazione linguistica, facilitazione linguistica, sostegno allo studio), h) profilo globale dello studente.

TRACCIA DI INTERVISTA

Frequentemente, in fase di accoglienza, si utilizza una traccia-tipo di intervista allo scopo di ridurre la soggettività nella relazione docente-studente e fornire indicazioni precise per esplorare quelle aree necessarie a predisporre un piano di apprendimento personalizzato. Nel caso degli adulti l'intervista, condotta anche alla presenza di un mediatore linguistico, è tesa a identificare eventuali competenze acquisite in contesti di apprendimento informali, come ad esempio il lavoro, tali da poter essere riconosciute come credito.

TEST DI INGRESSO DI CONOSCENZA DELLA LINGUA ITALIANA

Nei CPIA viene somministrato ai nuovi iscritti prima dell'avvio del percorso con l'intento di verificare il livello di conoscenza della lingua italiana posseduto e orientare lo studente al gruppo coerente con il livello posseduto.

FASCICOLO/DOSSIER PERSONALE DELLO STUDENTE

È lo strumento che raccoglie in maniera strutturata quanto emerso durante la fase preliminare di accoglienza; documenta le evidenze identificate, contiene le prove somministrate per la valutazione delle competenze e costituisce la base per definire il successivo percorso personalizzato. Il fascicolo assolve anche la funzione di orientamento e riorientamento, di valorizzazione della storia personale, di documentazione del vissuto dello studente, nonché di supporto durante le fasi di transizione verso altri ordini di scuola, nel caso di prosecuzione degli studi.

PATTO FORMATIVO INDIVIDUALE

Elaborato al termine della fase di accoglienza definisce il percorso di studio personalizzato, gli eventuali crediti riconosciuti, l'indicazione di eventuali percorsi integrativi e/o supplementari, le proposte per la facilitazione e la mediazione linguistica.

PROSPETTIVE

Il CPIA rappresenta oggi l'istituzione scolastica pubblica alla quale gli stranieri adulti e giovani adulti possono rivolgersi per esigere il diritto, costituzionalmente garantito, di apprendere la lingua italiana, per conseguire un titolo di studio di primo o secondo livello nonché per acquisire le competenze di base necessarie per esercitare consapevolmente la cittadinanza e per un inserimento autonomo, responsabile e consapevole nella società. L'organizzazione dei percorsi sulla base dei risultati di apprendimento¹² declinati in termini di

¹² In coerenza con la Raccomandazione del Consiglio e del Parlamento europeo sulla costituzione del Quadro europeo delle Qualifiche e dei Titoli per l'apprendimento permanente approvata il 23 aprile 2008 e analogamente agli altri percorsi scolastici previsti dagli ordinamenti, gli esiti dei percorsi di alfabetizzazione e apprendimento della lingua italiana vengono descritti in termini di risultati di apprendimento. L'approccio

competenze, abilità e conoscenze favorisce negli apprendenti adulti la capacità di usare il sapere nella vita quotidiana ed è funzionale alle esigenze richieste dal contesto sociale e lavorativo. Il titolo conseguito ad esito dei percorsi di alfabetizzazione e apprendimento della lingua italiana ha anche un valore aggiunto dal momento che ha validità legale per la richiesta dei documenti di soggiorno¹³. Nell'ambito degli accordi sottoscritti tra il Ministero dell'Interno e il Ministero dell'Istruzione al CPIA è stato attribuito il compito di predisporre, somministrare e valutare i test di conoscenza della lingua italiana di livello A2 e svolgere le sessioni di formazione civica previsti dall'accordo di integrazione¹⁴.

L'approccio interculturale costituisce il paradigma di riferimento della progettazione educativa, didattica e organizzativa nei CPIA. In una società in cui la presenza di migranti è sempre più presente ad ogni livello del tessuto sociale il CPIA ha adottato approcci didattici orientati alla comprensione delle diverse culture di appartenenza. Generalmente per stabilire una relazione positiva con un'altra persona e interagire con essa in maniera efficace ognuno di noi definisce le proprie strategie comunicative facendo ricorso a norme e valori che si porta dentro fin dalla nascita. Per stabilire altrettanto positi-

basato sui risultati di apprendimento – di derivazione europea - sposta il focus dagli input dell'apprendimento (durata dei percorsi, discipline insegnate ecc.) agli output ovvero agli esiti formativi declinati in competenze, abilità e conoscenze. Per la descrizione dei risultati di apprendimento, le Linee guida assumono a riferimento il Common European Framework of Reference for Languages con riguardo ai livelli A1 e A2, è articolata in cinque ambiti (ascolto, lettura, interazione orale e scritta, produzione orale e produzione scritta) a loro volta declinati in termini di competenze, abilità e conoscenze.

¹³ L'Italia si è uniformata alla linea europea riconoscendo la lingua come chiave di accesso al sistema: «il rilascio del permesso di soggiorno CE per soggiornanti di lungo periodo è subordinato al superamento, da parte del richiedente, di un test di conoscenza della lingua italiana» (l. 94/2009, art. 1, c.22, lett i).

¹⁴ Il 10 marzo 2012 è entrato in vigore il *Regolamento concernente la disciplina dell'accordo di integrazione tra lo straniero e lo Stato*, emanato con D.P.R. 14 settembre 2011, n.179. Da tale data, gli stranieri di età superiore ai 16 anni che faranno ingresso nel territorio nazionale per la prima volta e richiedano un permesso di soggiorno di durata non inferiore ad un anno, dovranno sottoscrivere tale accordo presso le Prefetture o le Questure. Con la sua sottoscrizione invece lo straniero si impegna ad acquisire un livello adeguato di conoscenza della lingua italiana parlata (equivalente almeno al livello A2 di cui al quadro comune europeo di riferimento), una sufficiente conoscenza dei principi fondamentali della Costituzione della Repubblica, della cultura civica e della vita civile in Italia (con particolare riferimento ai settori della sanità, della scuola, dei servizi sociali, del lavoro e degli obblighi fiscali) e, laddove presenti, a garantire l'adempimento dell'obbligo di istruzione da parte dei figli minori. All'atto della sottoscrizione vengono attribuiti allo straniero 16 crediti che corrispondono al livello A1 di conoscenza della lingua italiana parlata ed a conoscenze di base di formazione civica e le informazioni sulla vita civile in Italia e, al fine di favorire questo percorso di formazione, allo straniero viene fornita la possibilità di frequentare, entro 90 giorni dalla data della sottoscrizione, un corso gratuito di formazione civica della durata complessiva di 10 ore presso i CPIA.

vamente una relazione non stereotipata con persone appartenenti a culture distanti dalla nostra dobbiamo consapevolmente compiere uno sforzo supplementare volto a ridurre quanto più possibile il filtro comunicativo creato dai valori e dalle norme della cultura di appartenenza. L'approccio interculturale alla comunicazione ci spinge, quindi, a riformulare le nostre consuete categorie culturali e di pensiero: concetti storicamente definiti come identità e senso di appartenenza vanno intesi in senso dinamico e fluido; l'interazione spesso conflittuale con l'altro dovrebbe diventare un'opportunità di arricchimento e di crescita personale e collettiva, una possibilità di confronto rispetto a valori, regole e comportamenti.

Nell'ambito delle azioni messe in campo per promuovere l'offerta formativa e per intercettare l'utenza, assumono un ruolo molto importante le relazioni extrascolastiche e le connessioni formali e informali con il territorio. A livello di microsistema il CPIA fa riferimento alla comunità straniera, ai CAS, allo SPRAR, ai luoghi di aggregazione sportiva, culturale e religiosa, ai luoghi di lavoro; a livello di mesosistema il CPIA fa riferimento alla rete dei servizi sociali, sanitari, ai centri per l'impiego e alle reti per l'apprendimento permanente. La prospettiva interculturale si arricchisce a ben vedere di significati molto più ampi e complessi in relazione alle numerose contaminazioni tra il CPIA e i diversi soggetti pubblici e privati operanti nei contesti sociali, economici e organizzativi del territorio di riferimento.

Il sistema dei CPIA appare quindi strategico per l'inclusione e la coesione sociale dei cittadini stranieri e dei migranti in quanto, grazie alla cura posta nelle pratiche didattiche di insegnamento dell'italiano come seconda lingua, favorisce l'acquisizione e lo sviluppo di strumenti culturali e cognitivi fondamentali per una piena integrazione nella comunità di accoglienza, per l'inserimento nel mercato del lavoro e per un esercizio libero e consapevole della democrazia. Occorre tuttavia intervenire a livello di macrosistema incoraggiando una radicale revisione e un potenziamento delle vigenti disposizioni normative in materia di istruzione degli adulti secondo una prospettiva onnicomprensiva di apprendimento permanente che tenga conto delle continue trasformazioni sociali, demografiche, culturali ed economiche del mondo globalizzato.

REPERTORIO NORMATIVO

- Decreto Interministeriale del 12 marzo 2015 pubblicato in Gazzetta Ufficiale n. 130 del 6 giugno 2015 - Linee guida per il passaggio al nuovo ordinamento.
- DPR n. 286/1998 - Testo unico delle disposizioni concernenti la disciplina dell'immigrazione e norme sulla condizione dello straniero.
- DPR n. 394/1999 - Regolamento recante norme di attuazione del testo unico delle disposizioni concernenti la disciplina dell'immigrazione e norme sulla condizione dello straniero, a norma dell'articolo 1, comma 6, del decreto legislativo 25 luglio 1998, n. 286.
- Legge n. 40 del 6 marzo 1998- Disciplina dell'immigrazione e norme sulla condizione degli stranieri.
- Legge n. 132 del 1 dicembre 2018 - Disposizioni urgenti in materia di protezione internazionale e immigrazione, sicurezza pubblica, nonché misure per la funzionalità del Ministero dell'interno e l'organizzazione e il funzionamento dell'Agenzia nazionale per l'amministrazione e la destinazione dei beni sequestrati e confiscati alla criminalità organizzata.
- MIUR, CM n. 301 del 8 settembre 1989 - Inserimento degli alunni stranieri nella scuola dell'obbligo: promozione e coordinamento delle iniziative per l'esercizio del diritto allo studio.
- MIUR, CM n. 205 del 26 luglio 1990 - La scuola dell'obbligo e gli alunni stranieri. L'educazione interculturale.
- MIUR, CM n. 73 del 2 marzo 1994 - Dialogo interculturale e convivenza democratica: l'impegno progettuale della scuola.
- MIUR, CM n.24 del 1 marzo 2006 - Linee guida per l'accoglienza e l'integrazione degli alunni stranieri.
- MIUR, CM n. 4233 del 19 febbraio 2014 - Linee guida per l'accoglienza e l'integrazione degli alunni stranieri.
- MIUR, prot. n. 3298 del 23 marzo 2016 - Linee guida per la progettazione dei Piani regionali per la formazione civico linguistica dei cittadini di Paesi terzi finanziati a valere sul FAMI - OS 2 - ON 2 - Azioni formative specifiche - percorsi sperimentali.

BIBLIOGRAFIA

- Balboni, Paolo. 2015. *Le sfide di Babele. Insegnare le lingue nelle società complesse*, Torino: UTET.
- Baldacci, Massimo. 2006. *Ripensare il curriculum*, Roma: Carocci.

- Cerrocchi, Laura. 2018. "La famiglia come osservatorio e progetto migratorio in prospettiva pedagogica". *Sguardi Simmetrici*, Bologna: Ufficio Scolastico Regionale per l'Emilia Romagna.
- De Mauro, Tullio & Mancini, Federico & Vedovelli, Massimo. 1993. *Lessico di frequenza dell'italiano parlato*. Milano: Etas-Fondazione IBM.
- Gass, Susan & Selinker, Larry. 2008. *Second Language Acquisition*. New York London: Routledge.
- Giacalone Ramat, Anna. 2003. *Verso l'italiano. Percorsi e strategie di acquisizione*. Roma: Carocci.
- Lo Duca, Maria G. 2006. *Sillabo di italiano L2*. Roma: Carocci.
- Maher, Vanessa. 2012. *Genitori migranti*. Torino: Rosenberg & Sellier.
- Memmo, Daniela. 2012. *Cittadini e stranieri nel diritto privato della modernità*. Torino: Giappichelli.
- Ongini, Vinicio. 2011. *Noi domani, Un viaggio nella scuola multiculturale*. Roma: Bari: Laterza.
- Porcaro, Emilio. 2018. "Il ruolo dei CPIA nel processo di integrazione linguistica dei migranti con focus sui MSNA", in: *Sguardi Simmetrici*, Ufficio Scolastico Regionale per l'Emilia Romagna.
- Porcaro, Emilio. 2019. "L'alfabetizzazione: formazione scolastica e CPIA", in: *Narrare la migrazione come esperienza formativa. Strumenti e strategie di comunità e corresponsabilità educativa*. Milano: Franco Angeli.
- Porcaro, Emilio. 2019. "Adulti stranieri analfabeti o debolmente alfabetizzati nei CPIA: aspetti normativi", in: *Insegnare italiano ad analfabeti*. Torino: Bonacci – Loescher.
- Rocca, Lorenzo. (a cura di). 2008. *Percorsi per la certificazione linguistica in contesti di immigrazione*, Perugia: Guerra Edizioni.
- Turco, Livia & Tavella, Paola. 2005. *I nuovi italiani. L'immigrazione, i pregiudizi, la convivenza*. Milano: Arnoldo Mondadori editore.
- Vedovelli, Massimo. 2002. *Guida all'italiano per stranieri. La prospettiva del Quadro comune europeo per le lingue*. Roma: Carocci.

A bottom-up experience: The DiM project*

MORMONE A.**

USR Campania

BATTISTA M. S.***

CPIA Avellino

PENSABENE A.****

CPIA Avellino

ABSTRACT

This contribution aims to outline the experience gained while working on the DiM project, the context within which it arose, its development as an Erasmus+ project and its resulting outcomes. The first part of this paper introduces the adult education system in Italy, which is considered a significant tool for the country to enhance its growth; it also explains the purpose of the Centri Provinciali per l'Istruzione degli Adulti (CPIAs). The second part of the paper outlines the context in which the idea of the DiM Project took shape, having been designed from the outset as a teaching aid for students and teachers. The final section focuses on the development of the project from its inception, and highlights the work of the students who, thanks to their

* The authors discussed the content of this article in strict cooperation and agreement; however, for academic purposes, Angela Mormone is responsible for the section "Adult education: an opportunity for the country"; Maria Stella Battista for the section "The DiM Project in the context of CPIA Avellino"; Lia Pensabene for the sections "The idea of DiM", "First steps in the building of the DiM", "The DiM Project", "The key contribution of the experts" and "The results".

** Angela Mormone, Head Teacher of USR Campania.

*** Maria Stella Battista, Head Teacher of CPIA Avellino.

**** Amelia Pensabene, Teacher at CPIA Avellino, creator and coordinator of the DiM Project.

languages and with the assistance of teachers and experts, were the real key players in the creation of the multilingual dictionary.

KEYWORDS

Adult education; DiM Project; CPIA Avellino; foreign students inclusion; L2 teaching.

1. ADULT EDUCATION: AN OPPORTUNITY FOR THE COUNTRY

Over the last few years, lifelong learning, as an intentional individual process aimed at the acquisition of roles and skills involving a relatively stable change in the course of time, has become a key word in the welfare policies of the European Union. Among the many definitions of the term lifelong learning that have emerged over time, what makes the difference is certainly the principle of intentionality at the basis of which there is an individual who, proactively, seeks the training opportunities best suited to his/her goals and needs. On the basis of this assumption, it can be said that lifelong learning appeared as a structured concept in the second half of the 20th century, when UNESCO introduced it as a guiding principle for the renewal of education.

The need to talk about lifelong learning emerged primarily in the 1970s, in an international climate characterized by wars and poverty, when it became necessary to rethink a new educational paradigm. Later on, in the 1980s, the neo-liberal wave profoundly changed the context, and even the concept of lifelong learning shifted from a more humanistic vision to a prevalently economic one. Education was no longer understood as a tool to improve the conditions of society as a whole, but as a means to foster individual enterprise from a professional point of view. Moreover, since the 1990s there has been a new perspective on lifelong education in our country in the wake of the social inequalities created and in light of the challenges launched by the emergence of new technologies. The synthesis of this new paradigm is well framed in the working document of the European Commission, “A Memorandum on Lifelong Learning” dated 2000¹, which underlines the interconnection between social and economic changes. Starting from this assumption, lifelong learning must promote collective development starting from the personal realisation of the individual.

¹ See: Commissione delle Comunità Europee, *Memorandum sull'istruzione e la formazione permanente* Bruxelles, 30.10.2000 (https://archivio.pubblica.istruzione.it/dg_postsecondaria/memorandum.pdf).

While we are strongly convinced of the intrinsic potential of the adult education sector in affirming the importance of lifelong learning as a driving force for personal and collective development, the adoption of suitable tools to promote such training has not taken place with the same speed and awareness. In recent years, the new adult education system has become an indispensable strategic infrastructure of our country's development policies, capable of responding to new training needs and anticipating and prefiguring the necessary innovations in the system. In fact, it represents a real driving force for the relaunch of learning for those who, for various reasons, have left the school system and thanks to bespoke courses devised for them, are afforded the possibility to make up for lost time, but above all have the opportunity to complete or enhance their skills.

Nevertheless, it must be said that the adult education system, although fully part of the Italian school system, still suffers from problems related to the real recognition of its identity. Starting from the principle that satisfied adults are and will be adequate citizens in society, we should note that having moved beyond a first historical phase in which it was limited to a single, albeit indispensable, compensatory function (i.e. primary literacy), adult education is now configured as a true 'existential regulator' and a tool for social inclusion. People placed in a position to exercise this right are able to guarantee full participation in social life as citizens. Talking about adult education today, therefore, means referring to the twofold instrumental and existential function it performs: the first aimed at compensating for training gaps or returning illiteracy; the second more linked to the cultural, ethical, value-based and self-fulfilling dimension of the individual's development process. In this sense, it is worth remembering that it is now recognised at an interdisciplinary level that fostering a culture of lifelong learning means promoting strategies of individual and collective empowerment, of intra- and intercultural mediation, of improvement of equal opportunities and even of promotion of well-being and democracy. In order for all this to become concrete in practice, it is first necessary to resolve the many problems/critical issues that, obviously, in Covid 19's time have become exceedingly complicated and thus have pushed for the need to renew motivation and relationships to be implemented in adults, whether they are students or not. Today, there are increasing numbers of young people with complicated paths behind them, who opt for the aforementioned evening courses. The typical student is in fact no longer just a worker who wants to resume his or her studies to improve future working opportunities, but a young man or woman who is unemployed or has never worked before and is trying to get back into the game (the so-called 'NEET'² group).

Yet data from around the world give us a scenario that is certainly not encouraging and which, among other things, sees the presence of 750 mil-

² NEET is the acronym for "Not in Education, Employment, or Training".

lion illiterate adults. A picture that, in the specific case of the Italian situation, is even starker because of the persistence of strong inequalities between regions, due to the gap between Southern Italy and the national average³. A new glimmer of hope comes from the opportunity offered by Next Generation EU to create in Italy, finally, a truly structured system of lifelong learning in order to reach by 2025 the European objective of 50% of adults participating in training activities at least once a year.

It is in this scenario that the work of the Provincial Centres for Adult Education (CPIAs)⁴ scattered across the country should be enhanced. If, on the one hand, the CPIA represents a clear institutional point of reference for the “reception, orientation and accompaniment aimed at the adult population with particular reference to disadvantaged groups” (Guidelines, 2015), on the other, this sector of our education system is called upon to implement the enhancement of the cultural and professional heritage of the person, starting from the reconstruction of his or her individual history, conveyed by the Individual Training Pact, stipulated by a Commission appointed by the CPIA itself, and certified by means of the Individual Booklet.

The CPIAs are organized as proper educational institutions and are autonomous; however, they have a rather complex internal structure. A territorial service network articulated in three levels, namely:

a) administrative unit: the CPIAs consists of a central office and delivery points (associated offices) of first level paths and literacy and Italian language learning;

b) didactic unit: CPIAs stipulate network agreements with educational institutions that provide second level courses for adults, also in order to connect first and second level courses;

c) training unit: in order to widen the training offer, the CPIAs stipulate network agreements with local authorities and both public and private organizations in order to integrate/enrich/favour the connection between different types of education and training courses.

The CPIAs also carry out research on adult education through the Regional Centre for Research, Experimentation and Development (CRRS&S). In the Guidelines of 2015⁵, in consideration of the curricula of the CPIAs, it is still stated that the training offer made by the CPIAs is aimed at obtaining the

³ Suffice it to say that the proportion of graduates aged 30-34 is 21.6% in the South, compared to the national average of 26.9% (Source: OECD, 2018).

⁴ In the 2014/2015 school year, the new Provincial Centres for Adult Education (CPIA) were established. The new centres carry out the functions previously carried out by the Permanent Territorial Centres (CTP) and by the schools hosting evening classes. (see <https://www.miur.gov.it/istruzione-per-gli-adulti-centri-provinciali-per-l-istruzione-degli-adulti>).

⁵ DECREE 12 March 2015 - Guidelines for the transition to the new system to support the organisational and teaching autonomy of the Provincial Centres for Adult Education. (15A04226) (OJ General Series n.130 of 08-06-2015 - Ordinary Supplement n. 26).

certification attesting the level of education corresponding to that provided by the system in force at the end of elementary school; first and second cycle qualifications; the certification attesting to the acquisition of basic skills related to compulsory education; the qualification attesting to the achievement of a level of knowledge of the Italian language not lower than level A2 of the Common European Framework of Reference for Languages.

In this context, attention to policies for the inclusion of foreign students at the linguistic level is particularly important, considering that languages can be a source of discrimination and an obstacle to inclusion, but they can also be great resources and factors of resilience. The promotion of native languages and of the linguistic, human and cultural heritage that each immigrant brings with him/her cannot but be the point from which to start and on which to build new knowledge and new practices, linguistic but not only.

The result is the need to rethink education systems at all levels, with a special focus on the most vulnerable learners, who must be given specific instruction in order to guarantee social integration, even before educational/training integration. This is the direction in which the work of the eight CPIAs operating within Regione Campania is heading.

The educational and training dimensions that we find in this context are substantiated by competences: learning in prison, with migrants, at school, with adults, in the CPIAs, cannot but be dense with pathways that have “transversal and non transversal” competences as main axes to achieve adequate learning outcomes.

2. THE DiM PROJECT IN THE CONTEXT OF CPIA AVELLINO

On September 1, 2015 the Provincial Centre of Adult Education Avellino-Benevento came into being. It was to maintain this inter-provincial dimension until August 31, 2017 when two distinct entities, the CPIA Avellino and CPIA Benevento, were created. The centre included twenty-six service delivery points, of which six were prisons, with a student body made up primarily of foreigners mostly from Sub-Saharan Africa and Eastern European countries with a level of education ranging from illiteracy to a university degree. Currently the CPIA Avellino consists of the central headquarters situated in the city of Avellino, three associated branches located in the province, four branches of prison schools located in four different prisons. All venues offer literacy courses for Italian adults as well as literacy and Italian language learning courses for adult foreigners aimed at obtaining the qualification attesting to the achievement of level A2 of knowledge of the Italian language of the Common European Framework of Reference for Languages drawn up by the Council of Europe. Additional courses offered include first level courses leading to a lower secondary level qualification (equivalent to the Italian

terza media qualification); first level course - second educational period for the fulfilment of compulsory schooling comprising senior school years 1 and 2, for the fulfilment of compulsory schooling.

The DiM Project with its specific focus on learners' languages and cultures was born in this context. From the outset, it was clear that the interest shown in their own mother tongue made the students proud and happy, as they felt they were directly involved in the realisation of a project that concerned them closely. In addition, participation in the project yielded important results in the learning of Italian as a second language. There was also an impact on students not directly involved in the project, who enthusiastically took part in related workshop activities such as, for example, the production of drawings, and put forward suggestions that their own languages be included in the dictionary at a future stage.

2. 1 THE IDEA OF DiM

The idea of DiM arose within the CPIA Avellino in an Italian language course for foreigners (Italian L 2) during the 2016/2017 school year. The course was held at Poggio dei Signori, a facility located in Forino, which housed some sixty adult asylum seekers. The CPIA Avellino had signed a memorandum of understanding with this facility for the running of the Italian L2 course within it.

During the welcome and orientation phase which took place in the same structure, the class teacher conducted sociolinguistic interviews in the presence of the cultural mediator on duty at Poggio. The questions concerning the assessment of formal and informal competences, but also the understanding of the linguistic, cultural and experiential background, were asked by the teacher in Italian so that she could be aware of the level of L2 knowledge, and if necessary were translated into French or English. Often it was necessary, if not indispensable, to have an 'internal mediator' who was close to the language and/or ethnicity of the interviewee and could translate the answers into French, English or Italian.

The emerging class group, the vast majority of which came from sub-Saharan Africa, was composed of a group of French-speaking students from Mali, Togo, the Ivory Coast, Burkina Faso and Senegal; a group of English-speaking students mostly from Nigeria but also from Gambia, Ghana, Liberia and finally two students speaking only Arabic or its variants from Oman and Chad. Within the groups there were students who were not literate in their native language and had never attended school in their own country. Among them, three types of non-literate people could be distinguished: 1. those who spoke only their mother tongue and other local dialects and were unable to write in any language; 2. those who could speak but not write in the official

language (English or French) of their countries of origin; 3. those who had never attended any kind of school but spoke and wrote in French or English. Among those who had attended school, only a few had finished the course of study and obtained a diploma, all the others had attended school for a few years and only occasionally; yet another group, Muslim by faith, had only attended Koranic school and were not familiar with the Latin alphabet. A further group, from Nigeria, claimed to be able to read and write English, when in fact they were only able to hand copy flawlessly any text, despite being unable to read.

It was evident that the linguistic competence of all the students, including the non-literate ones, was remarkable since they spoke and understood at least three or four languages. Their ability to communicate within the various groups was surprising. There was something that I would call “linguistic solidarity” that allowed everyone to communicate with each other and with everyone else. This solidarity manifested itself very clearly within the same language group and was reinforced during classroom activities by reaching out to the other language groups, as in the case of the two Arabic-speaking students whom everyone tried to help even though they did not know Arabic. The idea of DiM was conceived in the multilingual and multicultural context of this classroom where a vast and varied world of different languages, sounds and cultures was enclosed within a few square metres. It was an opportunity to see the students from a different vantage point, as knowledge holders, giving voice to their identity through their native languages.

2. 2 FIRST STEPS IN THE BUILDING OF THE DiM

The idea behind the project was to take advantage of the linguistic skills of the students in the class to create a glossary of their languages, to be used as a teaching tool useful to both teachers and learners. The starting point for the experimental work was the lexical list A1 of the *Profilo della lingua italiana*⁶ made up of about 500 headwords of level A1 of the CEFR (Common European Framework of Reference for Languages).

The lemmas were listed in alphabetical order on a sheet containing three columns: the first column contained the lemma in Italian; the second column

⁶ *Profilo della lingua italiana. Livelli di riferimento del QCER A1, A2, B1 e B2* is the result of the project entitled “Descrizioni dei livelli di riferimento per le lingue nazionali e regionali” (DLR), promoted by the Language Policy Division of the Council of Europe in collaboration with the CVCL (Centre for Language Evaluation and Certification) of the University for Foreigners of Perugia and other international institutions (Cervantes Institute, University of Cambridge, Université Paris III - Sorbonne nouvelle, among others). (see: https://www.hubscuola.it/profilo_lingua_italiana/origini.html).

contained the translation into English or French (for French or English speaking users); the third column was empty for the insertion of the translation into one's own language. The students were divided into groups speaking the same language.

From the teacher's observations it could be inferred that the task was carried out with a great deal of participation and interest, often outside class time, and was quite challenging. There were lengthy discussions about each term before finding the one that everyone thought best corresponded in translation from one language to another.

The writing phase was the most complicated and was entrusted by the group to the best student, usually to the pupil with the greatest number of years of schooling. The difficulty in transliterating the sounds of one's own language, mostly used orally, into a written form was evident. After the arduous phase of writing, the teacher recorded the lemmas with a mobile phone.

During the school year, the following languages were collected and recorded for A1 level: Kotokoli spoken in Togo; Bambara spoken in Senegal; Igbo, spoken in Nigeria. During the following school year (2017/2018), Twi language lemmas for A1 and A2 level were collected and recorded in another class consisting of 23 learners all from Ghana and belonging to the same ethnic group. Urdu language for A1 level was also collected with the help of a Pakistani student with a degree in foreign languages from Pakistan, attending a first period course (corresponding to the eighth grade).

3. THE DiM PROJECT

Already in its experimental phase, in the two years preceding the funding of the project by the European Community, the students involved had shown great interest and participation by working also after school hours. The project in its embryonic stage was presented at an interregional conference of CPIAs (Campania, Piemonte and Emilia Romagna) in December 2017. Colleagues appreciated the results of the project and suggested submitting an application for an Erasmus project so that it could be adopted as good working practice in teaching activities for the teaching of L2, and as a useful tool for the inclusion of foreign students.

A project application was submitted by CPIA Avellino as lead school for the Call 2018 within the framework of the Erasmus plus KA204 Projects - Strategic Partnerships for Adult Education for Good Practice and Social Inclusion, and was approved and funded by the European Community (Grant Agreement Number 2018-1-IT02-KA204-048332).

The project officially began on 15 October 2018 with the First International Meeting of all partners in Avellino to define the expected objectives, the languages to be included and the modalities of implementation. Seven European

partners participated in the realisation of the DiM Project: CPIA Avellino as Coordinator, CPIA Benevento, CPIA Ravenna, Active Citizens Partnership (Greece), 36.6 Competence Centre (Scotland), Future Focus Ltd (Malta), IMS (Cyprus). All the schools and institutions involved are active in the field of adult education for foreigners, and in particular in the teaching of L2. It should be emphasised at this point that the real protagonists of the project were the students of the various organisations and their native languages.

The choice of Bengali by the CPIA Avellino as a language to be included in the DiM was made following the monitoring that is carried out every year to gain a general overview of the students' catchment area. During the 2017/2018 school year, 1098 students were enrolled at the CPIA Avellino (91% of whom were foreigners, subdivided as follows: 69% from Africa, 16% from Asia and 15% from Eastern Europe. The Bengali language, or Bangla, spoken in Bangladesh but also in the Indian state of West Bengal, was one of the most widely spoken languages given the high percentage of enrolled students from Bangladesh, and was collected by the CPIA Avellino. The other student languages chosen were Bambara (spoken in Mali), collected by the CPIA Benevento; Wolof (spoken in Senegal) collected by the CPIA Ravenna; Fârsi (spoken in Iran), collected by the Greek partners; Russian (spoken in Russia), collected by the Cypriot partners; Tagalog (spoken in the Philippines), collected by the Maltese partners; Tigrinya (spoken in Eritrea), collected by the Scottish partners.

The CPIA Avellino coordinated the various phases and activities of the project by creating the format for the collection, and the database used to process the data for the construction of the online dictionary. It also designed the DiM project website. The linguistic data was acquired by each partner through laboratory activities in which the students translated the words of the lexical list A1 of the Profile of the Italian Language, first into English (vehicular language of the project) and then into their mother tongue. They subsequently recorded the audio of the translated word; transcribed the word onto an Excel file shared on Google drive; produced a drawing (these were created by students who lacked or possessed lower levels of literacy) and wrote a definition consulting one or more dictionaries. The students' work was supported by a teaching team, usually a teacher acting as a technician for the audio recordings and data entry into Google drive and the Language 2 teacher. For languages with an alphabet other than Latin, Google's input tools were installed as a Chrome extension in order to be able to write with characters not present on the keyboard.⁷

⁷ <http://www.dimproject.net/>

A1	A	B	C	D	E	F	G	H
1	id_term	ve_lang	ve_lang_gr_ref	semantic_field	ben	grammar reference	audio_file_name	transliteration
2	1	at	(prep.)	Describing things	তে	পদার্থী অব্যয়	ben_1_at.mp3	te
3	2	reside	(verb)	Home and building	বাস করা	ক্রিয়া পদ	ben_2_reside.mp3	bās karā
4	3	vinegar	(n. mass.)	Food and drink	সির্কা	অগণনীয় বিশেষ্য পদ	ben_3_vinegar.mp3	sirkā
5	4	water	(n. uncount.)	Food and drink/Natural world	পানি	অগণনীয় বিশেষ্য পদ	ben_4_water.mp3	pāni
6	5	now	(adv.)	Time	এখন	ক্রিয়া বিশেষণ পদ	ben_5_now.mp3	ekhan
7	6	airplane	(n. count.)	Travel	বিমান	গণনীয় বিশেষ্য পদ	ben_6_airplane.mp3	bimān
8	7	airport	(n. count.)	Travel	বিমানবন্দর	গণনীয় বিশেষ্য পদ	ben_7_airport.mp3	bimānabandar
9	8	affection	(n. var.)	Relationship	স্নেহ	পরিবর্তনশীল বিশেষ্য পদ	ben_8_affection.mp3	sneha
10	9	agency	(n. count.)	Home and building	সংস্থা	গণনীয় বিশেষ্য পদ	ben_9_agency.mp3	sangasthā
11	10	August	(n. var.)	Time	আগস্ট	পরিবর্তনশীল বিশেষ্য পদ	ben_10_august.mp3	āgast̪
12	11	help	(convention)	Communication	সাহায্য	নেতিবাচক	ben_11_help.mp3	sāhāy
13	12	tree	(n. count.)	Natural world	গাছ	গণনীয় বিশেষ্য পদ	ben_12_tree.mp3	gāch
14	13	alphabet	(n. count.)	Communication	বর্ণমালা	গণনীয় বিশেষ্য পদ	ben_13_alphabet.mp3	barnamālā
15	14	tall	(adj.)	Describing things/People.appearance	লম্বা	বিশেষণ পদ	ben_14_tall.mp3	lambā
16	15	love	(verb)	People/personality/Relationship	অলবাসা	ক্রিয়া পদ	ben_15_love.mp3	bhālabāsā
17	16	embassy	(n. count.)	Home and building	দূতাবাস	গণনীয় বিশেষ্য পদ	ben_16_embassy.mp3	dūtābas
18	17	friend	(n. count.)	Relationship	বান্ধবী	গণনীয় বিশেষ্য পদ (স্ত্রী)	ben_17_friend.mp3	bāndhabī
19	18	friend	(n. count.)	Relationship	বন্ধু	গণনীয় বিশেষ্য পদ (পুরুষ)	ben_18_friend.mp3	bandhu
20	19	love	(n. uncount.)	Relationship	প্রেম	অগণনীয় বিশেষ্য পদ	ben_19_love.mp3	prem
21	20	also	(adv.)	Communication	এছাড়াও	ক্রিয়া বিশেষণ পদ	ben_20_also.mp3	echhārō
22	21	go	(verb.)	People actions/Travel	যাওয়া	ক্রিয়া পদ	ben_21_go.mp3	jāoyā
23	22	animal	(n. count.)	Animals	পশু	গণনীয় বিশেষ্য পদ	ben_22_animal.mp3	Pashu

Google sheet for Bengali language data collection

All the shared online work was evaluated by the CPIA Avellino team, who were responsible for checking the technical quality and constantly monitored all the data collected. Once the processing phase had been completed, data was uploaded onto the online DiM platform through the project website www.dimproject.net.



Bengali alphabet drawn by students

3. 1 THE KEY CONTRIBUTION OF THE EXPERTS

The project presented a number of challenges that had to be dealt with at different stages. From the outset, participating partners assumed that creating a dictionary would be a very simple task, the same assumption applying to the collection of linguistic data entrusted to them. Only during the implementation phase did they realise the difficulties and complexity of the project. The greatest problems were encountered both in the data acquisition phase and above all during the revision phase of the linguistic material at our disposal. It is widely believed, even among teachers, that whoever speaks a language must also know it. There was a lack of awareness among participants that speakers do not necessarily possess knowledge of their own language, just as, in many cases, they do not possess the metalinguistic ability to reflect on how their language works and how it is used. This led to several difficulties in the collection of language data as in the case of Tigrinya where the group of students was heterogeneous in terms of geographical origin (some from Eritrea and others from Ethiopia) and speaking very different and mutually unintelligible dialect varieties. In addition, the students were not very literate, having studied for only a few years in their country. For this reason, during the revision phase entrusted to the experts, the linguistic data of Tigrinya turned out to be mostly inaccurate (on the peculiarities of Tigrinya variants and their script, see Lusini in this volume).

Many difficulties were overcome thanks to the participation in the DiM Project of various experts, highly qualified people in the various languages, who generously gave their time and experience to advise, guide and review the work done by our students. Immediately after the approval of the DiM Project funding, the coordinator got in touch with Professor Adriano V. Rossi⁸, an expert in lexicography and Iranian languages, to request an opinion on the project and to seek advice on how to proceed with the design of the multilingual dictionary. Following an initial meeting with Professor A. Rossi, subsequent meetings were also attended by Professor Maddalena Toscano⁹, who has coordinated a number of European projects on multilingual education in multicultural contexts in the past, and Professor Flavia Aiello¹⁰. All of them found the project valuable and interesting and agreed

⁸ Professor Emeritus Adriano V. Rossi was full Professor of Iranian Linguistics and Iranian Philology at the University of Naples "L'Orientale". Since 2016 he is President of ISMEO – The International Association for Mediterranean and Oriental Studies, Rome.

⁹ Maddalena Toscano is a former researcher and teacher in Swahili language at UNIOR with experience as main coordinator in various SOCRATES Comenius projects (see for e.g.: T.I.M.E for Teachers <http://opar.unior.it/179/1/Intercultural.pdf>).

¹⁰ Flavia Aiello is Associate Professor of Swahili Language and Literature at the University of Naples L'Orientale.

to collaborate to ensure its success, especially in the linguistic revision phase. Thanks to them, it was possible to get in touch with Professor Ilaria Micheli¹¹, who revised the Bambara language resources and provided useful information on the linguistic structures of African languages. Professor Gianfrancesco Lusini¹², was the consultant for Tigrinya; and Emiliano Minerba, a PhD student at the University L'Orientale in Naples, took care of the language revision for Wolof and all the problems related to data gathering for this language.

These academics, to quote Professor Lusini's words, decided to "take an active militancy" by going out into the field, working side by side with our native speakers and supporting the teachers with their specialist linguistic expertise. As mentioned earlier, the biggest problems for African languages were related to the prevalence of oral language and the lack of familiarity with written variants. This phenomenon occurs even among the most educated speakers, since the language of education often corresponds to the language of the former colonial administration.

3. 2 THE RESULTS

In its final format, the project consists of an online multilingual dictionary (<http://dictionary.dimproject.net/index.php>) called DiM, mounted on a digital platform and made freely available to all. The DiM currently contains eleven languages: Bambara; Bengali; Farsi; Russian; Tagalog; Tigrinya; Wolof; Italian; Greek; Maltese and English. A list of about five hundreds entries is available for each language. Each word is accompanied by audio; a schematic drawing made with vector graphics (for concrete names); a transliteration for words written in non-Latin scripts; a grammatical reference; the definition of the word (for highly coded languages with a great written literary tradition); a drawing made by the students for some languages (Bambara, Bengali and Wolof); language profiles containing the alphabet and descriptions of the main features of each language; semantic labels referred to words belonging to the same semantic field; examples of activities that can be carried out in L2 classes using words, audio, drawings or semantic lists.

¹¹ Ilaria Micheli is Associate Professor of African Linguistics at the University of Trieste - Department of Legal, Language, Interpreting and Translation Studies (IUSLIT).

¹² Gianfrancesco Lusini is Full Professor for Ge'ez and Amharic languages and literatures at the Dipartimento Asia, Africa e Mediterraneo of the University of Naples "L'Orientale".

clothes

	Bambara	Bengali	Greek	English	Farsi	Italian	Maltese	Russian	Tagalog	Tigrinya	Wolof
	susseti	সসেতি	κόλτσα	sock	حوراب	calza	kalzetta	носок	medyas	ከሳሊ	kawas b-
	tirikó	জামা	πουκάμισο	shirt	پیراهن	camicia	qmis	рубашка	damit pang itaas	ከገግቶ	simis b-
	kalisó	জামা	παλτό	coat	کت	cappotto	kowt	пальто	amerikana	ከገገት	west b-
	nèné	জাকেট	σακάκι	jacket	ژاكت	giacca	gakketta	куртка	dyaket	ጃገት	west b-
	jipu	ফাঁট	φούστα	skirt	دامن	gonna	dublett	юбка	palda	ከገ	sipp b-
	dùtoki dé	টি-শার্ট	μπλουζά	t-shirt	تی شرت	maglietta	t-shirt	футболка	t-shirt	ሜልቶ	siletmaa b-
	dùtoki	সোয়েটার	πουλόβερ	sweater	ژاكت	maglione	flokk	свитер	damit pang ginaw	ጎልፎ	mbubb b-
	núneti	চশমা	γυαλιά	glasses	عینک	occhiali	nucčali	очки	salamín sa mata	ሜንገር	lonet y-
	sánninèlàn	ছাতা	ομπρέλα	umbrella	چتر	ombrello	umbrella	зонт	payong	ዩ-ሌል	seelukaay b-
	kùlusi já	ট্রাউজার	παντελόνι	trousers	شلوار	pantalone	qalziet	брюки	pantalon	ከረ	tubay j-
	sàbara	জুতা	παπούτσι	shoe	كفش	scarpa	žarbut	туфля	sapatos	ዲላ	däll w-
	dùtoki	পেপশাক	φόρεμα	dress	لباس	vestito	libsa	платье	damit	ከገገ	yére b-
	boré	কুল ব্যাগ	σακίδιο	backpack	كوله پشني	zaino	barzakka	рюкзак	backpack	ሜገደር	saaku b-

Semantic list of clothes

The user interface created for the DiM is very simple and is accessible from any device. Moreover, it is also designed to be used by people with a low literacy level and is suitable for embedding (the inclusion of the dictionary in any other website). Accessing the dictionaries is possible directly through the link: dictionary.dimproject.net (<http://dictionary.dimproject.net/>) and also through the website link: <http://www.dimproject.net/>. The DiM Project website chronicles the development of the project in its various stages and also contains articles related to the project written by participants or external persons. In addition, it contains the profiles of the countries involved in the project; a recording of the Multiplier Event (a one day webinar to showcase the project, with the participation of all partners, students, public figures and the academics who were instrumental to its success); and texts in Italian and English of the contributions of speakers who participated in the webinar.

CONCLUSIONS

The DiM does not claim to be a dictionary, but rather a multilingual glossary. Furthermore, it was not produced by expert lexicographers. The work produced in collaboration with the students was nonetheless a significant achievement for all those involved. The project offered a great opportunity for human and professional enrichment resulting from the exchange of experiences between the different partners, as well as between students and teachers from different countries, in order to gain a new awareness of the students’ languages and cultures. It also provided the opportunity to reflect on the importance of languages as a vehicle of identity and culture; on being open to different cultures, languages and literatures and on other ways of receiving, teaching and learning with the belief that all languages are part of the common human heritage and are to be safeguarded and protected.

The DiM experience highlighted the importance of language and the need for language education in schools. Not all teachers who deal with language on a daily basis are aware of linguistic diversity and the importance of language knowledge in the learning process. The sociolinguistic interviews conducted with the students highlighted the strong emotional bond they have with their mother tongue. This powerful tie is encapsulated in the words of Rahman, a Bengali student at the CPIA Avellino and one of the protagonists of the work on Bengali. During the sociolinguistic interview, when asked if he was interested in participating in the DiM Project, he replied: “Sure, so finally I can do something for my language¹³ too”. In fact, his commitment to the project has been truly incomparable, as he has devoted much of his free time in collecting data and writing definitions. Rahman concluded the interview with these words: “If I could express everything in my language, I would say it better and you would understand more”. This final statement captures well the meaning of the DiM Project.

¹³ The Bengali language movement arose as a political effort in East Pakistan, advocating the recognition of Bengali as one of the languages of Pakistan. The language issue played a very prominent part in the creation of Bangladesh.

REFERENCES

Commissione delle Comunità Europee, 30.10.2000 - *Memorandum sull'istruzione e la formazione permanente:- Bruxelles*. https://archivio.pubblica.istruzione.it/dg_postsecondaria/memorandum.pdf (Site consulted before 15/06/2021).

Istruzione per gli adulti - Centri Provinciali per l'Istruzione degli Adulti

<https://www.miur.gov.it/web/guest/istruzione-per-gli-adulti-centri-provinciali-per-l-istruzione-degli-adulti> (Site consulted before 15/06/2021).

Spinelli B. e Parizzi F. (2010) *Profilo della lingua italiana. Livelli di riferimento del QCER A1, A2, B1, B2*, Firenze, La Nuova Italia / RCS Libri. https://www.unistrapg.it/profilo_lingua_italiana/site/index.html (Site consulted before 15/06/2021).

OECD (2018), *Education at a Glance 2018: OECD Indicators*, OECD Publishing, Paris, <https://doi.org/10.1787/eag-2018-en>. (Site consulted before 15/06/2021).

Lexicography and language learning of Swahili L2 at UNIOR: the Swahili-Italian online dictionary project

F. AIELLO, M. TOSCANO, R. TRAMUTOLI*
Università degli Studi di Napoli “L’Orientale”

ABSTRACT

The present essay outlines the progress of a lexicography project, namely a Swahili-Italian online dictionary, which was developed at the University of Naples “L’Orientale” (henceforth UNIOR) and conceived as a useful digital tool for Italian L1 learners of Swahili in a context of renewal of the teaching/learning of Swahili at the UNIOR, encouraged by web-based distance tools and platforms.

As will be explained in detail, the software of the online lexical database, first developed in the period between 2003 and 2009, is in the process of being updated with some technical and lexicographical improvements, aiming for a resource that will be increasingly user-friendly to Swahili language students, and in general to Italian-speaking learners.

* The authors discussed the content of this article in strict cooperation and agreement. However, for academic purposes, Flavia Aiello is responsible for the sections “Introduction” and “Swahili L2, lexicography and web-based learning at Unior”; Maddalena Toscano for the sections “The online dictionary UWAZO” and “Swahili word structure and dictionary entry in UWAZO”; Rosanna Tramutoli for the sections “From UWAZO to KIU: software access and management tools” and “New learner-oriented features in the Swahili-Italian online dictionary KIU”.

KEYWORDS

Swahili L2 language learning; Swahili-Italian online dictionary; lexicography.

INTRODUCTION

This contribution aims to chart the developments of the Swahili-Italian online dictionary project at UNIOR, which was led by the three authors, and discusses some of the inherent challenges in Bantu languages/Swahili lexicography which targets foreign learners, along with the solutions recently adopted in the course of the dictionary updating.

The paper is organised in five sections, the first of which provides a background to the development of the dictionary's project by giving an overview of the teaching of Swahili language and literature at UNIOR. It also touches upon the related need for lexical resources for Italian speakers and the progressive use of distance learning and web-based didactic materials and activities by the teaching staff. The subsequent two parts are devoted to the origins of the project, namely the online lexical database UWAZO: the history of the project, its participants and the initial design of the database, which originates from lexicographical research applied to Swahili language and word structure. Owing to the old software being out-dated, UWAZO is currently in the course of being reprogrammed from an IT point of view by an IT specialist. This new phase has provided an opportunity to redesign some features of the dictionary, both in terms of software access and management tools, and of learner-oriented characteristics of the entries, as will be discussed in the last sections of this paper.

SWAHILI L2, LEXICOGRAPHY AND WEB-BASED LEARNING AT UNIOR

The teaching of Swahili language and literature at UNIOR was established in 1969 by Elena Bertoncini Zúbková, an internationally renowned scholar of Swahili language and literature, whose educational activity was closely connected with research on the Swahili lexicon¹, resulting in the production of lists of words and vocabularies². These were (and still are) much needed tools for the Italian students of Swahili, given the ongoing rapid expansion of this

¹ See, for instance, Bertoncini Zúbková 1973.

² The work of Bertoncini Zúbková in this field remains largely unpublished. The only publication, out of print, is a small vocabulary published in 1977, *Vocabolario swahili-italiano e italiano-swahili* (1977).

Bantu language, called *Kiswahili* by its speakers, in Eastern and Central Africa. From the Indian Ocean coast and its neighbouring islands, where the Swahili people (*Waswahili*) live, the Swahili language has spread in the continent due to various historical factors, such as the caravan routes of the Arabo-Swahili merchants, the activity of missionaries, the language policies of colonial and postcolonial governments and the flows of migrants and refugees. It is currently used by approximately 99,000,000 people, mainly in Kenya, Tanzania and in the eastern regions of the Democratic Republic of Congo (DRC)³. It is also spoken in Rwanda, Burundi, Uganda and, to a lesser degree, in Somalia, the Comoro Islands, Mozambique, Malawi, Southern Sudan and Zambia⁴. As a consequence of the post-independence development of Swahili in a wide geo-cultural area and in multiple contexts of language use, especially in Tanzania, where it has become the dominant medium of communication in the public sector (basic education, administration, courts, national assembly, media etc.), the Swahili lexicon has been constantly growing, also due to a huge lexicographical initiative undertaken by experts from the National Swahili Council (Baraza la Kiswahili la Taifa, BAKITA) and the Institute of Swahili at the University of Dar es Salaam (Taasisi ya Uchunguzi wa Kiswahili⁵, TUKI)⁶. These latter institutions, besides producing general dictionaries, have been creating and disseminating terminologies for many domains of language use, such as law, science, and IT glossaries⁷.

Regarding bilingual lexicographical works aimed at Italian learners, apart from E. Bertoncini Zúbková's work which, as remarked above, remains substantially unpublished or out of print, three dictionaries have been published so far. These are Vittorio Merlo Pick's *Vocabolario kiswahili-italiano e italiano-kiswahili* (EMI, Turin 1961, re-edited in 1978, currently out of print), Maddalena Toscano's pocket-size *Dizionario swahili. Swahili-italiano, italiano-swahili* (Vallardi, Milano 2004) and Gianluigi Martini's *Dizionario swahili. Swahili-italiano, italiano-swahili* (Hoepli, Milano 2016). Furthermore, a terminological work has appeared, namely a Swahili-Italian linguistic glossary by Rosanna Tramutoli (*Kamusi ya isimu Kiswahili – Kiitaliano*, TUKI, Dar es Salaam 2018), published by the university press of the Institute of Swahili at the University of Dar es Salaam within

³ See <https://www.ethnologue.com/>.

⁴ Swahili has been declared national/official language in Kenya, Tanzania, DR Congo, Uganda, and more recently in Rwanda. As underlined by M. Mulokozi (2003), the Swahili language has also an international status, being amongst the official languages of the African Union since 2005. Furthermore, since 2019 Swahili is an official working language of the Southern Africa Development Community, an organisation composed of 16 states of Central, Eastern and Southern Africa.

⁵ In 2009 the Institute was renamed TATAKI (*Taasisi ya Taaluma za Kiswahili – Institute of Swahili Studies*).

⁶ See Sewangi 2007.

⁷ See Aiello & Toscano 2017.

the sphere of a long standing cooperation agreement with UNIOR. All these useful works for Italian speakers, are however in print format, and partly only available at the library of UNIOR. Thus, in the early 2000s, the development of an online lexicographical resource for Italian learners was initiated by M. Toscano, as will be explained in detail in the next paragraph, in a context of renewal of the courses of Swahili at UNIOR, encouraged by Internet and digital technologies, which have revolutionised the teaching of foreign languages by introducing web-based distance tools and platforms.

Over the years, different digital materials and activities have been developed at UNIOR by the teachers of Swahili (E. Bertoncini Zúbková, M. Toscano, F. Aiello), in cooperation with mother-tongue collaborators, researchers and technicians⁸, to be used for “blended” or “hybrid” teaching/learning, i.e., to be combined with traditional face-to-face language courses⁹. These include a list of online resources for autonomous learning, available on the website of UNIOR’s Language Centre, CLAOR (Centro Linguistico di Ateneo Università L’Orientale)¹⁰, as well as e-learning modules which offer teaching materials and exercises integrating the contents of the Swahili language courses. The e-learning resources and activities are offered through the Moodle digital platform managed by CLAOR, and consist of two courses, one for beginners (“Swahili livello iniziale”) and one for intermediate-advanced level (“Swahili livello avanzato”)¹¹, which can be accessed by registered students who are given the course login by the teacher. The basic-level Swahili e-learning course targets students, enrolled in BA or MA programs at UNIOR¹², who are starting the study of one or two African/Oriental languages; the intermediate-advanced course is offered to 3rd year BA students and MA students (4th/5th year). The learning objectives of the first course are to consolidate Swahili basic grammar knowledge and develop listening and comprehension skills. The intermediate-advanced course has been designed for students learning Swahili language

⁸ All the individual credits are available on the web pages quoted in footnote n. 11.

⁹ Smyrnova-Trybulska, Eugenia. 2009.

¹⁰ Swahili is the sub-Saharan language most widely taught at university level, inside and outside Africa, therefore a great deal of academic and non-academic teaching materials is available online, alongside Swahili-language media (information portals, musical videos, films etc.). The resources are classified in different sections, such as “integrated abilities”, “information portals”, “culture and society” etc., see <https://www.unior.it/ateneo/7644/1/swahili-sul-web.html>.

¹¹ See respectively <https://elearning.unior.it/course/view.php?id=84>; <https://elearning.unior.it/course/view.php?id=5>.

¹² Undergraduate degree courses: “Oriental and African Languages and Cultures”, “Political Science and International Relations”, “Comparative Languages and Cultures”; Master’s Degree Courses: “Languages and Cultures of Asia and Africa”, “International Relations and Institutions of Asia and Africa”, “Comparative Literatures and Cultures”. See https://www.unior.it/index2.php?content_id=17159&content_id_start=2&titolo=lauree-e-lauree-magistrali&parLingua=ITA.

and literature, mostly to further develop the written comprehension of literary texts and the awareness of translation strategies. The switch to “blended” teaching/learning has been facilitated by the user-friendly features of Moodle, for which basic computer literacy is sufficient, and has involved the need to redesign course materials, as one simply cannot shift content from one medium to another, i.e. from textbook to digital, asynchronous technology¹³. Rather, it is important to acknowledge the characteristics and potentials of the new medium. Therefore, fair amount of new content has been introduced, based on audio/picture/video resources (such as vocabulary quizzes, cloze tests for listening comprehension etc.), which expose the students to “realia” and/or to language use in full contexts. Moreover, an advantage of Moodle-based activities is that they promote student-centred learning since the students can self-evaluate their achievements step by step.

Furthermore, from a partnership of UNIOR with Federica Web Learning, a platform created by the University of Naples Federico II to offer high-quality MOOC (Massive Open Online Courses), a MOOC for Swahili language and culture has been created by F. Aiello and published online in 2020, entitled “Karibuni! Introduzione alla lingua e cultura Swahili” (Welcome! An introduction to Swahili language and culture)¹⁴. The course is structured in 10 teaching units, containing slides (with text, pictures and links), short video-lessons and auto-evaluation quizzes meant both for students and a wider public.

Finally, a digital corpus of Swahili texts was collected by M. Toscano in collaboration with a number of students and researchers. It is a small (one million words) Swahili untagged raw corpus consisting close to fifty full texts, taken mainly from contemporary written literature, with the addition of some oral narratives and non-literary works (socio-political essays, handbooks about agriculture, media studies, information technology). It is available to researchers and MA students for their dissertation research on Swahili language, literature and linguistics.

AINI, a Swahili lemmatizer based on MSDOS¹⁵, was used for extracting lemma from the corpus.

The corpus was available online for a short while and was also searchable through the use of regex (regular expressions)¹⁶ incorporated in Shika Neno¹⁷. For reasons of copyright restrictions only two lines of the retrieved contexts text were accessible to the user.

¹³ See Azaryad Shechter, Deborah 2015.

¹⁴ See https://www.federica.eu/mooc/c/karibuni_introduzione_alla_lingua_e_cultura_swahili.

¹⁵ By Professor T. Schadeberg and P. S. E. Elias, Leiden (no more available).

¹⁶ For info on regular expressions see <https://www.sketchengine.eu/guide/regular-expressions/>.

¹⁷ SHIKA NENO was a Linux environment, language dependent software for context retrieval. It allowed search for Swahili bases (by M. Sorrentino and M. Toscano, no more available).

The corpus is now only accessible offline. The digital corpus is used for searching examples and quotes to be included in the Swahili-Italian online dictionary.

The main tool used for information retrieval is CONCORDANCE¹⁸, a software for context retrieval which operates in a Windows environment. It is a language independent software which allows search for forms. CONCORDANCE also allows use of regular expressions and provides access to full text when the need for disambiguation occurs. It is for offline use only.

Here below are some screen shots from CONCORDANCE, the main software used by UWAZO operators to provide reliable examples of lemma usage.

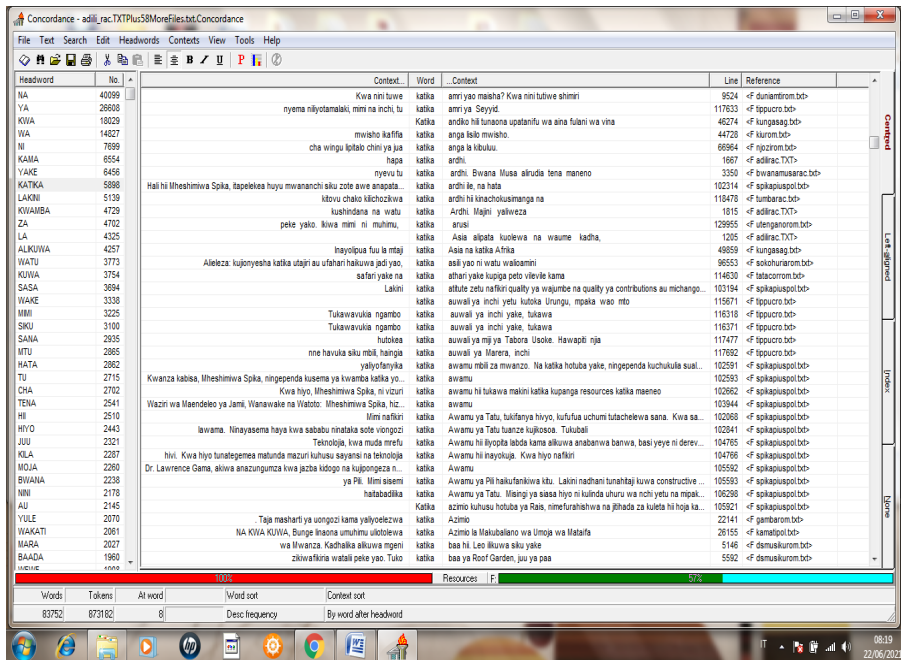


Figure 1. Concordance - Context sort of KATIKA, by word after headword

¹⁸ Concordance: <http://www.rjcw.freereserve.co.uk/>.

2 right		3 right		4 right	
No.	Word	No.	Word	No.	Word
183	Ya	663	Ya	203	Ya
161	Nchi	312	Na	175	Na
74	Maisha	192	Wa	80	Wa
53	Nyumba	179	Kwa	70	Katika
50	Dunia	164	Za	67	Kwa
50	Mj	161	Yake	55	Za
50	Sehemu	130	Chi	43	Ni
48	Hutaba	106	Lakini	38	Yake
47	Bunge	100	Yetu	37	Watu
46	Kia	93	Chi	30	Kama
46	Mambo	79	Huu	25	La
45	Jani	73	Wala	24	Hi
45	Kpindi	66	Sasa	24	Kwamba
45	Ushari	60	Sana	23	Hata
44	Maka	41	Vya	19	Kuwa
42	Kiji	38	Zote	19	Nchi
38	Chumba	37	Haya	17	Cha
38	Shule	37	Hiki	16	Hi
37	Maeneo	36	Huo	15	Wake
35	Kundi	34	Mj	15	Ameyo
35	Safari	30	Ile	15	I
34	Mj	29	Kwa	15	Maisha
34	Umwingu	29	Zote	15	Alikuwa

Figure 2. Concordance - Collocations of KATIKA

Concordance - adili_rac.TXT\Plus58MoreFiles.txt

Headword	No.	Context	Word	Context	Line	Reference
WA	40599	nipokuwa nimekaa, Wambie naseba	katika	amani ya vurumi. Staki	58074	<F mzinglerom bdt>
YA	26508	pokuwa kweli, wewe u	katika	amani, maana babu yako kana	117558	<F sippocro bdt>
KWA	18029	Nawe una	katika	amara yake, wambie na wasenzaji wote.	117197	<F sippocro bdt>
WA	14827	nyema nilyotamaki, nimi na nchi, tu	katika	amri ya Seyyid.	117833	<F sippocro bdt>
NI	7699	Kwa nini tuwe	katika	amri ya masha? Kwa nini tutuwe shirini	9524	<F dunantrom bdt>

View - FA...chivetta\SWAHILI-Archivo_texti_swahili\CORPUS\adili_rac.TXT\Plus58MoreFiles.txt

Clipboard Search Options Window Help

Manayema.

Alipoambwa khabari ile Mirambo akamwita Sef: Waambie wale Waarabu, kesho waende zao, maana zimekuja khabari amri kadha wa kadha. Nami nasema khabari hizo uwongo. Ziya-pokuwa kweli, wewe u **katika** amani, maana babu yako kana ndugu yangu. Lakini nataka, waondoke hawa Waarabu, waende zao, wasipate khabari wapagari wao, watu wa Tabora, watacupa misigo wakimbie. Itakuwa khabara kuu kwao.

144. Akenda Sef bin Hamed akawambia wale Waarabu na watu wangu: Hesho safari, mwende zemu. Nami nataka kurudi Tabora. Wakasafiri wale Waarabu pamoja na watu wa Mirambo. Na siku ya pili akapewa Sef bin Hamed pembe tano kwa frasila tisa wa robo akapewa na nguo, madeule na bushi na vitambi bure vinne. Aktoa na watu wake sita akampa Sef bin Hamed, Menda nao hawa watu. Watatu awachukue Hamed bin Muhammed pwan, akawambie Seyyid Barghash: Sasa baada ya kuja Hamed bin Muhammed sitaki tena vita naye. Hamed

Dr. Lawrence Gama, akiva amanzunguzwa

hivi Kwa hyo tunategemea matunda lewama. Ninayasema hay

NA KWIA KUWA

Words: 83752 Tokens: 873182 At word: 8 Desc frequency: By word after headword

Figure 3. Concordance - Access to full text, for needs of disambiguation

THE ONLINE DICTIONARY UWAZO

UWAZO¹⁹ is a Swahili-Italian online bilingual lexical database for Italian speakers. It contains about 5000 headwords selected from a Swahili frequency list²⁰ and from various other sources, including, in addition to the above quoted works of Merlo-Pick and Toscano, the lexicon used in the teaching materials of Swahili courses (by E. Bertoncini Zúbková) and some specific lexicons (e.g. immigration, body parts) resulting from various unpublished works.

Based on T.E.I. guidelines²¹ and developed over a period of time between 2003 and 2009²², UWAZO came into being mainly thanks to IsIAO²³ funding and the collaboration of various experts²⁴.

Alongside the T.E.I. guidelines, reference was made also to the guidelines contained in the materials of the lexicography and terminology courses²⁵, acquired by the staff team (M. Toscano, F. Aiello, R. Tramutoli) through participation in specific training courses, as well as reference works for lexicog-

¹⁹ Originally available at the link http://old.iuo.it/diprapa/swahili_leksikoloja/uwazo/default.asp, temporarily hosted at the link <http://www.siamoinsieme.org/mawazo/>.

²⁰ See Bertoncini Zúbková 1973.

²¹ T.E.I. (Text Encoding Initiative): <http://www.tei-c.org/index.xml>.

²² UWAZO was presented at the DWS 2004. Maddalena Toscano, Giuseppe Marzatico, Salvatore La Gala and Massimiliano Sorrentino: *Building a corpus based Kiswahili-Italian online lexical data base*, Third International Workshop on DICTIONARY WRITING SYSTEMS (DWS 2004), Brno, Czech Republic, 6-7 September 2004 (<https://nlp.fi.muni.cz/dws2004/pres/#17>).

²³ Funding came from the bilateral agreement IsIAO-UNIOR, 2002-2003 and from the bilateral agreement UNIOR-Univ. of Dar es Salaam, 2002-2005.

²⁴ M. Toscano for project planning, Swahili language consultancy from E. Bertoncini Zúbková and S. Sewangi (Univ. of Dar es Salaam). Inserting data by F. Aiello, M. Toscano, with occasional cooperation from C. Marzio and G. Aquaviva. As for the software, the first version was by Giuseppe Marzatico and Salvatore La Gala (MARS, Napoli). The second version, now in use, was built by Tommaso Borrelli and Luciano Piedimonte.

²⁵ June 1996. *Stage Inter-regional de formation à l'utilisation du logiciel MARIAMA*, a 10 days training course organized by the GDRE 1172 of CNRS, at the Centre de Formation de Luminy, Marseille, FR.

September 7-18 1998. *Intensive course on lexicography*, by AFRILEX-SALEX, Univ. of Pretoria, R.S.A.

August 5-8 1999. *16th Terminology Summer School, Module 1 – Terminology Teaching and Training*.

August 9-11 1999. *1st Terminology Summer Campus*, by TermNet and Infoterm, Univ. of Donau, Krems, AU.

July 16-20 2001. *Training workshop in lexicography and lexical computing*, by Information Technology Institute, University of Brighton and Lexicography Masterclass, Univ. of Brighton, (UK).

June 8-12 2015. *The 2015 Lexicom Workshop in Lexicography and Lexical Computing*, Masaryk University's Centre in Telč, Czech Republic.

raphy and publications on lexicographic standards. In the project phase other available DWSs were considered²⁶.

UWAZO was specifically tailored for a list of needs which include flexibility in setting masks and labels, controlled on-line access, search facilities, output in .doc and .xml, and a user friendly and corpus based approach. Unlike a traditional printed dictionary, the online dictionary supports the transfer of knowledge from research to teaching/learning of Swahili (teaching module of computer applications to African languages), while also enhancing tools for distant learning and cooperation.

The UWAZO general framework is based on the aims and objectives of target users, and on language aspects which are relevant to the project. These elements were taken into account while defining the data base structure. Resources and tools available at the time were also considered.

UWAZO target users are Italian speakers/learners of Swahili, namely language students also majoring in (literature, political science, etc.) who are required to learn grammar (morphology and syntax) up to advanced level, to read and translate literature and political texts; to produce proper Swahili texts. Potential interested users are also non-language students (other individuals from NGOs, tourists, etc.) who need to know grammar at beginner/intermediate level.

UWAZO's basic aim is to support Italian mother tongue speakers in the autonomous learning of Swahili by providing detailed information on grammar/morphology and clues to syntactic structures. It also aims to provide updated lexical information, i.e., loanwords from English (e.g. *skrini* 'screen'); new meanings assigned to already existing words (e.g. *kibarua* 'temporary work') or neologisms (e.g. *simu ya mkononi*, 'mobile phone'; *tovuti* 'website'; *mtandao* 'internet').

All entries contain an indication of the grammatical category and at least one simple gloss. Most of the items are divided into sub-items and completed with examples of use. The inflected grammatical forms, in particular, contain various examples of use referring to the different meanings and functions. All examples were taken from the above-mentioned Swahili corpus.

Following T.E.I. guidelines for printed dictionaries which include free and structured entries, the UWAZO database structure is based on various groups which include a fixed list of elements with free position. Sub-class elements are also available. Main T.E.I. groups list used in UWAZO include: Gruppo grammaticale, Traduzione, Esempio, Etimologia, Confronta, DictScrap. Open lists of labels can also be created by the operator.

²⁶ MARIAMA, by R. Nicolai, Univ. of Nice - (historical linguistics, restricted access) LinguaLinks; Shoe Box, by S.I.L. (field work, commercialised); DBT - LEXXIKO, by E. Picchi, ILC-CNR, Pisa (old version in MSDOS environment, new WINDOWS version not commercialised).

UWAZO management tools allow parallel login by various operators; it is possible to identify different log in identities but it has only two levels of management (control on giving new access only).

UWAZO does not allow off-line work (e.g. lexicographers working at home with their notebooks); it allows online work only. The corpus, presently available only offline, has to be made available to each operator.

Apart from the simple search for lemma, various advanced search options are available to the user: by PoS (part of speech), by type of word (form, morpheme, stem). It is also possible to search for words, Swahili and Italian, used in the examples, as shown in the following screen shots from UWAZO:

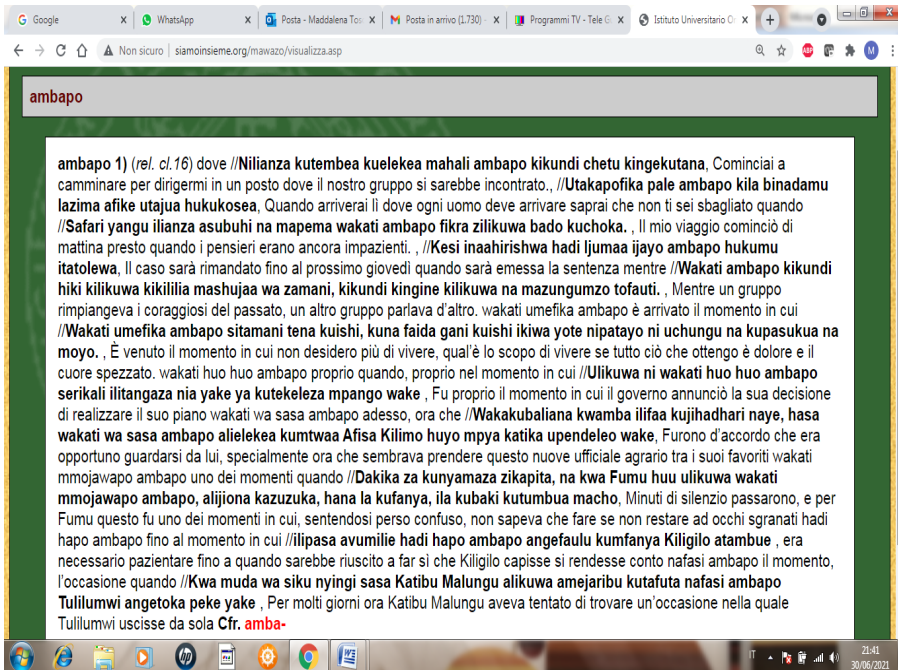


Figure 4. UWAZO - Basic search for lemma 'ambapo'

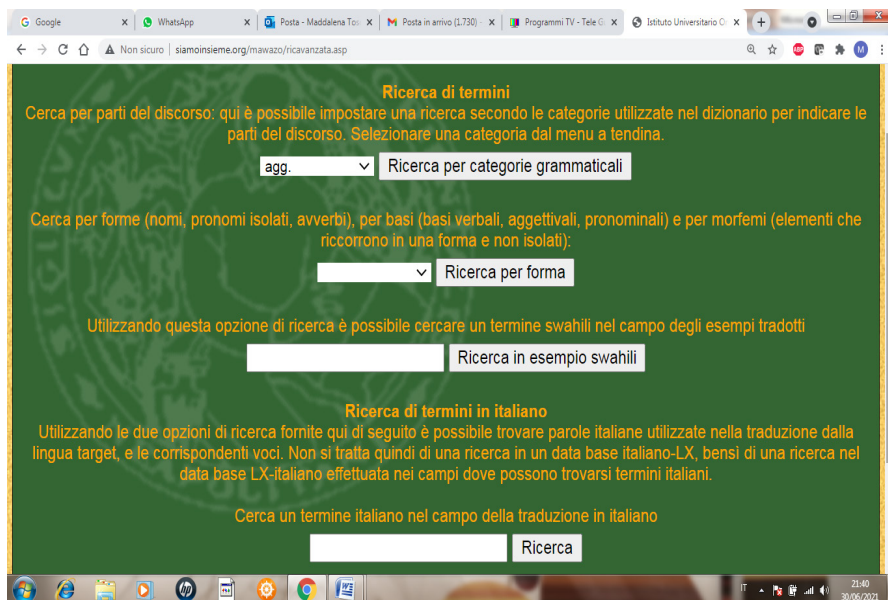


Figure 5. UWAZO - Advanced search options

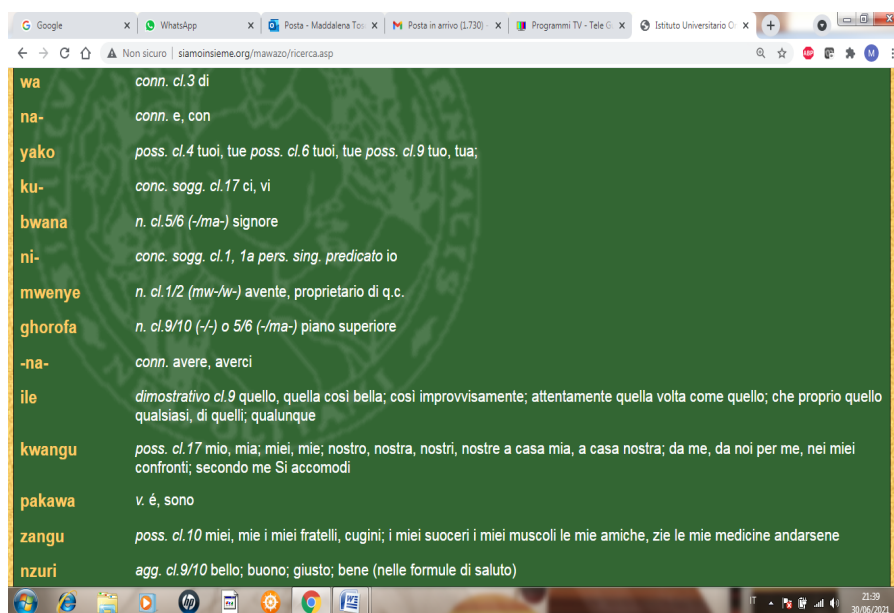


Figure 6. UWAZO - Search for 'nyumba' (casa) in Swahili examples. By clicking on each lemma shown in the list, the user will see different Swahili examples containing the word 'nyumba'

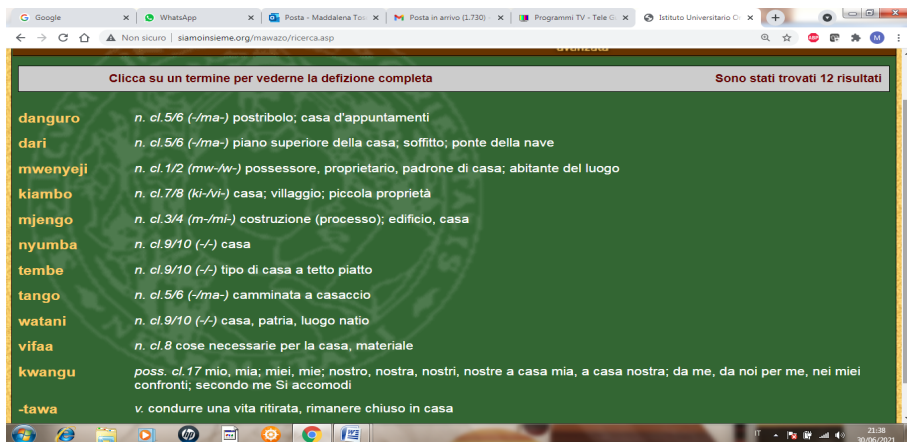


Figure 7. UWAZO - Search for word ‘casa’ in the Italian translation of examples in Swahili

SWAHILI WORD STRUCTURE AND DICTIONARY ENTRIES IN UWAZO

Bantu languages are characterized by a complex system of noun classes, represented by noun prefixes. With the exception of classes 12 and 13 (which are no longer productive in the language), and class 11-14 which have merged into a single class, the Swahili noun class system includes 15 noun classes. Swahili words follow a main basic order: *Morpheme(s) + Stem*:

- nouns: class concord + noun stem (noun classes: cl. 1/2, 3/4, 5/6, 7/8, 9/10, 11-14/10; verbal nouns: class 15; locative classes: 16, 17, 18);
- adjectives: adjectival class concords (cl. 1...18) + adj. stem;
- pronominals: pronominal concords (cl. 1...18) + stem; stem + pr. conc. (1...18);
- verbs: subject class concords (cl. 1...18) + tense marker + (relative conc. cl.1...18) + (object concords cl. 1...18) + verbal stem + final vowels/extensions.

Swahili is an agglutinative language, which means that morphemes are juxtaposed to form words. Within the Swahili lexicographical tradition, the accepted lemmatization strategy is to list nouns in their full forms with class prefixes, whereas the prefixes of verbs, numerals, pronouns and inflected adjectives are ignored, and the stems alone are listed²⁷. Thus, in standard printed dictionaries, we find the following types of Swahili entries:

²⁷ See Wójtowicz 2016: 410.

- nouns: singular form (e.g.: *moyo* ‘heart’ → *mioyo*; *moyoni*, *mioyoni*²⁸);
- adjectives: stem (e.g.: *-baya* ‘bad’ → *mbaya*, *wabaya*, *mibaya*, *baya*, *ma-baya*, *kibaya*, *vibaya*, *kubaya*, *pabaya*);
- pronouns: stem (e.g.: *h-* ‘this’ → *huyu*, *hawa*, *huu*, *hii*, *hili*, *haya*, *hiki*, *hivi*, *hizi*, *huku*, *hapa*, *humu*; *huyo*, *hao*, etc.);
- verbs: stem (e.g.: *-soma* ‘to read’ → *-somea*, *-someka*, *-somesha*, *-somwa*, *-somana*, etc.);
- invariable forms (some prepositions, conjunctions, adverbial forms, etc.).

Considering that Swahili students are used to this lexicographical tradition, for the online Swahili-Italian dictionary we have decided to choose solutions adopted by almost all Swahili dictionaries in regard to the process of lemmatization. Consequently, we have listed only the stems alone for verbs, numerals, and inflected adjectives, ignoring subject concord and agreement prefixes.

On the other hand, in some other cases we have decided to reject traditional lexicographic solutions usually adopted in printed dictionaries, in response to beginner learners’ needs and we have listed closed sets of grammatical words, such as pronouns, in their full forms (with the stems also included as separate entries). We have also included all grammatical morphemes as separate entries, adding an explanation about their function.

Thus, the basic structure of an entry in the online Swahili-Italian dictionary includes the following elements:

- headword;
- variant/variants of the headword;
- poS (part of speech), that is, the grammatical category;
- noun class;
- translation (gloss and/or description);
- examples (quotations, collocations, idioms, proverbs, etc.);
- example translations;

Moreover, each entry is categorized according to the type:

- form (nouns or invariable entries);
- grammatical stem (verbs, pronouns, adjectives, variable entries in general);
- morpheme: e.g. noun prefixes, subject prefixes, object markers, derivational suffixes, etc.

²⁸ *Moyo/mioyo* (cl. 3/4): ‘heart’; *moyoni*: ‘in the heart’; *-ni* is a locative suffix that expresses the locative relations indicated in Italian through prepositions, such as ‘in/at/from’ (Bertoncini Zúbková 2009: 7).

In considering the Swahili word structure, it is evident that, while working with a Bantu language, we have to address problems not experienced by lexicographers working with European languages. These problems are connected primarily to two issues: the form of headwords and the presentation of the numerous derivatives of a single root²⁹. In the following sections, we will explore challenges and difficulties regarding the design of a new Swahili online dictionary as a learning/teaching language tool.

FROM UWAZO TO KIU: SOFTWARE ACCESS AND MANAGEMENT TOOLS

The UWAZO software is presently being redesigned from an IT perspective owing to the old software now being obsolete. The upgrade is carried out by an IT specialist on the basis of the lexicographical indications provided by the three authors of this contribution. It includes some improvements aimed at the creation of a resource that will be increasingly useful to Swahili language students, and in general to Italian-speaking learners, as outlined in the next paragraph.

In the first phase of assembly of the new dictionary, the main effort was devoted to redesigning an updated version of the software, importing and editing previous data by tidying inconsistent entries, adding relevant missing information, and completing sets of grammar inflected forms. Closed sets belonging to semantic and grammatical categories, like days of the week, months and adjectives were verified, and additional vocabulary was also collected by students and researchers who worked on chosen sets they found useful in their studies, such as body parts or the Swahili COVID lexicon.

The new version of the dictionary, called KIU (Kiswahili-Italian-UNIOR), will be published online on an ad hoc page of the UNIOR website, and will be freely accessible to university students and the general public. The number of entries in the database (currently about 5,000) will also be increased.

The KIU dictionary, accessible by users from computers or smartphones, supports two different levels of access depending on the role: administrators, having the highest level of access to the database, can implement and edit data, and set up and manage the accounts of students and collaborators. Students and learners can have full access to database and software tools except for data publishing, which needs administrator approval.

UWAZO management tools also allowed parallel login by various operators, where it was possible to identify different login identities. However, unlike UWAZO, which only provided control on creating new login access, the

²⁹ Wójtowicz 2016: 410.

updated software enables the administrator to control both access and data editing as well as publishing of all other operators.

Another implementation which characterizes the new software consists in the export/import of full or selected data to facilitate editing and updating by remote access.

Moreover, in addition to the Swahili-Italian dictionary, the new software was designed to be extended to other Bantu languages; in particular a Zulu-Italian lexical database (including a collection of body vocabulary and a small literary corpus) is being developed by R. Tramutoli.

NEW LEARNER-ORIENTED FEATURES IN THE SWAHILI-ITALIAN ONLINE KIU DICTIONARY

New features for the entries have been introduced in the updated version of the Swahili-Italian online dictionary, in order to address specific Swahili L2 learner needs.

The new dictionary is going to be a learner's dictionary, that is, a bilingual dictionary with features that until recently have been primarily associated with monolingual learner's dictionaries: "extended grammatical information (meant to make the creation of real sentences easier, by providing hints for constructing the proper agreement patterns) and with visualisation of derivative forms that will provide extra lexical information and make navigation across the dictionary easier"³⁰.

As regards to Swahili nouns, the entire plural form is entered immediately after the class prefixes and the searches can be carried out on headwords and plural forms of headwords; thus, if a user looks up for *maembe* 'mangos', he/she will be directed to the entry for *embe* 'mango'.

Differently from paper dictionaries and other online lexical databases, the online Swahili-Italian dictionary offers the possibility to search both for regular and irregular forms of the plurals, such as *jiko* (sg. cl. 5 'kitchen') - *meko* (= *ma* + *iko*; pl. cl. 6 'kitchens'); *jino* (sg. cl. 5 'tooth') - *meno* (= *ma* + *ino*, pl. cl. 6 'teeth'), where the noun class prefix of cl. 6 *ma-* coalesces with the vowel *-i* changing to *me-*. Indeed, students at a beginner level might not be familiar with similar plural forms of class 6 (e.g. *meko*, *meno*), which change consistently compared to the corresponding singular forms, following the Bantu rules of vowel coalescence. The possibility to search for full plurals is thus very convenient for learners at a beginner level who do not have sufficient knowledge of grammar to enable them to identify easily singular and

³⁰ Bański & Wójtowicz 2008: 269-275. Bański & Wójtowicz 2012: 60-72.

plural forms which carry different noun prefixes and are hidden in the entries of the singular form³¹.

Another important issue with respect to macrostructure is related to the handling of derivation, which in Swahili, like in other Bantu languages, is very robust and typically creates dozens of complex lexemes from a single root, especially with regard to verbal roots³².

Verbs are described by the type of derivation and references to their base/root in the case of derivatives, or to derivatives in the case of roots. Through a mechanism of cross-entry references, derivatives are inserted as searchable sub-entries showing both sides of the derivational process (derivative→ root/root→ derivative)³³. In standard paper dictionaries, only frequently used derivative verbs (e.g. *-pendeza* ‘be attractive’; *-endelea* ‘continue’) or those extended forms of the verb which have become lexicalized or fossilized, and which meaning is independent from the derivational process, (such as *-sima-ma* ‘stand’; *-kumbatia* ‘hug’; *-patikana* ‘be available’; *-wezekana* ‘be possible’) are usually searchable as separate entries and do not require learners to be aware of their morphological structure.

The Swahili dictionary for learners provides the possibility to look up all verbal derivative forms, i.e. all extended verbs, which have been listed as searchable sub-entries linked to the corresponding verbal bases. Thus, learners can either directly search for extended verb stems (e.g. *-fundishana* ‘teach each other’; *-jibizana* ‘answer each other’) even without being familiar with the derivational process. Alternatively, they can look for the basic form of the verb and consult the corresponding possible derivations. The extended mechanism of cross-entry references, which also characterizes other types of learner-oriented dictionaries designed for Bantu languages (see T-Lex; isiZulu.net)³⁴, helps “to maintain a system whereby derivatives have the status of headwords, while simultaneously the derivational and semantic relationships between forms are preserved”³⁵.

Swahili diminutive and augmentative forms are generally created by adding noun class prefixes of class 5/6 and 7/8 respectively to the nominal stem, for instance:

mbuzi cl. 3 (goat) → *kibuzi* cl. 7 (little goat) → *buzi* cl. 5 (big goat).

Thus, augmentatives and diminutives have also been entered as searchable derivatives forms linked with the corresponding nominal base, except for those forms which have acquired independent meaning in Swahili and thus

³¹ Kiango 2005: 264.

³² Wójtowicz 2016: 410.

³³ Wójtowicz 2016: 411.

³⁴ TshwaneLex: <https://tshwanedje.com/dictionary/swahili/>; <https://isizulu.net/>.

³⁵ Wójtowicz 2016: 410.

constitute separate headwords (e.g. *m-ji* cl. 3 ‘town’ → *ki-ji* cl. 7 ‘village’; → *ji-ji* cl. 5 ‘metropolis’)³⁶.

Moreover, the updated Swahili-Italian dictionary (KIU) includes different entry variants (orthographic, dialectal/continental, allomorphs), searchable as cross-reference entries and visualized as part of the headword. Thus, entries are searchable not only in the plural form, but also according to the different orthographic and phonetic variants. For instance, both alternative forms *asante*, *ahsante* ‘thank you’, have been inserted in the entry.

Similarly, the search of noun class prefixes includes all allomorphs, that is the possible variants occurring in different phonological contexts, such as:

cl. 1 **m-**: *mw-* and *mu-* are two different variants which occur in different phonological contexts; *mw-* before vowels except before vowel “u”, where the prefix *mu-* does not undergo vowel assimilation -, e.g. *muungwana* ‘gentleman’)

cl. 2 **wa-**: *w-* before vowel (e.g. *mwalimu*, pl. *walimu* ‘teacher/s’);

cl. 3 **m-**: *mu-* and *mw-*; similar situation as class 1 prefix;

cl. 7 **ki-**: *ch-* before vowel;

cl. 8 **vi-**: *vy-* before vowel.

The Swahili noun class system is quite standardised and homogeneous; each noun class is represented by noun prefixes which mark all elements of a Swahili sentence and thus encodes the grammatical information necessary for the grammatical agreement. Nevertheless, in some cases, apart from noun class prefixes, other semantic and grammatical skills, which can hardly be included in the entry of standard dictionaries, are required in order to construct a correct Swahili sentence.

In order to enrich the learning tools and support beginner learners in the acquisition of Swahili grammatical knowledge, the design of the updated Swahili-Italian dictionary provides a tool for adding grammatical comments or usage notes where appropriate. This option allows us to enhance the quality of information contained in the entry with the aim of supporting the acquisition of grammar skills and expanding the learner’s vocabulary through the addition of:

- indications on the correct grammatical agreement for more complex cases: e.g. the agreement of animate nouns from non-human classes (e.g. *kijana* ‘young man’ cl. 7; *mama*, ‘mother’, cl. 9; *waziri* ‘minister’ cl. 5); possessive agreement with animate nouns in class 9/10 referring to close relationship (e.g. *bibi yangu* ‘my grandmother’: possessive agreement is in class. 9 and not in class 1 **bibi wangu*), etc.;
- notes on the semantic features of a term in order to disambiguate meanings and facilitate the appropriate choice/use of a term in translations

³⁶ Class 12, including the diminutives in *ka-*, has disappeared in standard Swahili and has been reintroduced by some Tanzanian authors from the hinterland under the influence of other Bantu languages; in a few cases, also the plural *tu-* of class 13 was imported.

and oral/written production³⁷. For this purpose, semantic explanations can also be accompanied by a number of labels/tags indicating status (formal, informal, slang, derogative, euphemism, vulgar, colloquial etc.); register (literary, familiar, popular, etc.); semantic field (biology; zoology; military; music; legal; medicine; religion, etc.), frequency of use (common, rare); figurative or extended meaning.

Swahili learners will rely on phonological tools as well, such as audio recordings (and possibly the transcriptions) of difficult words to pronounce. Since Italian L1 Swahili learners do not usually face difficulties in Swahili pronunciation due to the high degree of similarity with the Italian phonological system, pronunciation has not been inserted by default for each entry; nevertheless, a few terms containing difficult sounds (not present in the Italian phonological system) are supported by audio recordings with pronunciation (e.g. words containing the sound *ng'*, such as *ng'ombe* or the sound *j*, such as *jambo*). This learner-oriented feature is also useful when adding hints for Swahili words with different accents and meaning, e.g. *barabàra* 'highway', *baràbara* 'perfectly'. Since in Swahili accent regularly falls on the penultimate syllable, the dictionary offers the possibility of inserting an audio support for those exceptional cases where an accent falling on the third last syllable of a word is distinctive in that it produces difference in meaning.

Furthermore, the dictionary provides another useful learner-oriented tool through a function which allows cultural terminology to be supported by images/pictures, together with a description/definition which substitutes the gloss/translation.

The possibility of adding images and descriptions to an entry is particularly helpful in order to clarify specific terms, which cannot be easily translated into Italian or for which the Italian gloss is not exhaustive enough to explain a cultural concept (e.g. *ugali*: "typical Swahili food similar to 'polenta'"; *kanga*³⁸: "coloured women's textile"). This additional information is generally avoided in printed dictionaries and cannot be included due to printing size restrictions.

³⁷ A cross-reference system is also exploited to link synonymous entries, such as *kinywa - mdomo* 'mouth'.

³⁸ "The kanga is a widely spread printed cloth, mainly used by women as a dress. The cloth measures ca. 110 cm in height and 150 in length. It is defined by a border (*pindo*), a central field (*mji*) and usually contains on the lower third a printed proverbial inscription (*jina*). We know that it was "invented" around the 1880s in Zanzibar and imported from Europe in this form, i.e. already imprinted with patterns and inscription. It played an important role in the emancipation of slaves and their integration into the Muslim Swahili community of the East African coast. [...] the kanga allows for the communication of the unspeakable, whereby the interactants cannot be held responsible for their interaction." Beck, 2001: 157, 166.

REFERENCES

- Aiello, Flavia & Toscano, Maddalena. 2017. "On some Swahili ICT terms". In Batic, G. C. & Leger, R. (eds.), *Studia Africana, Papers in Honour of Sergio Baldi*, 19-34. Köln: Rüdiger Köppe Verlag.
- Azaryad Shechter, Deborah. 2015. "Blended Learning Course Format on Moodle: A Model for Beginner Level Foreign Language Courses in Higher Education" n: Gómez Chova, L., López Martínez, A., Candel Torres, I. (eds.), *INTED2015 Proceedings* (9th International Technology, Education and Development Conference Madrid, Spain. 2-4 March, 2015), 183-209 Madrid: IATED.
- Bański, Piotr & Wójtowicz, Beata. 2008. "Multi-level reference hierarchies in a dictionary of Swahili". In Bernal, E. & De Cesaris J. (eds.), *Proceedings of the XIII EURALEX International Congress* (Barcelona, 15-19 July 2008), 269-275. Barcelona: Institut Universitari de Lingüística Aplicada – Universitat Pompeu Fabra.
- Bański, Piotr. & Wójtowicz, Beata. 2011. "New XML-encoded Swahili–Polish Dictionary: Micro- and Macrostructure". In Goźdz-Roszkowski, S. (ed.). *Explorations across Languages and Corpora. PALC 2009*, 497-514. Frankfurt a. Main: Peter Lang.
- Bański, Piotr. & Wójtowicz, Beata. 2012. "Swahili lexicography in Poland: its history and immediate future". *Studies of the Department of African Languages and Cultures* 46, pp. 60-72.
- Beck, Rose M. 2001. "Ambiguous signs: the role of the kanga as a medium of communication." *Swahili Forum* 8, pp. 157-169.
- Bertoncini Zúbková, Elena. 1973. "A tentative frequency list of Swahili words", *Annali dell'Istituto universitario Orientale di Napoli* 33, pp. 297-363.
- Bertoncini Zúbková, Elena. 1977. *Vocabolario swahili-italiano e italiano-swahili*. Napoli: Opera Universitaria, IUO.
- Bertoncini Zúbková, Elena. 1985. "Quantitative analysis of Swahili vocabulary", *Studies in African Linguistics*, supplement 9 (présis from the 15th Conference on African Linguistics, UCLA, March 29-31, 1984), pp. 22-26.
- Bertoncini Zúbková, Elena. 2009. *Kiswahili Kwa Furaha*. Corso di lingua Swahili (Tomo I). Roma: Aracne.
- Kiango, John G. 2005. "Problems of Citation Forms in Dictionaries of Bantu Languages". *Nordic Journal of African Studies* 14 (3), pp. 255-273.
- Mulokozi, Mogyabuso. 2003. "Kiswahili as a national and

- international language”,
Kiswahili 66, pp. 66-80.
- Sewangi, Selemani. 2007.
“Uundaji wa istilahi za
Kiswahili kileksikografia na
athari zake”, *Nordic Journal
of African Studies* 16 (3), pp.
333-344.
- Smyrnova-Trybulska, Eugenia.
2009. “About Some
Basic Aspects of Distance
Learning”. In Eugenia
Smyrnova-Trybulska (ed.),
*Theoretical and practical
aspects of distance learning:
collection of scholarly papers*,
13-35. Cieszyn: University of
Silesia in Katowice.
- Wójtowicz, Beata. 2016.
“Learner Features in a
New Corpus-based Swahili
Dictionary”. *Lexikos* 26, pp.
402-415.

Unveiling oral and writing skills of low-literate learners of L2 Italian: from research to teaching practice

MARTA MAFFIA

Università degli studi di Napoli “L’Orientale”

ABSTRACT

This contribution is meant to give an insight into the topic of L2 acquisition and teaching in the case of low-literate adult learners. To the aim, two applied researches will be presented, both focusing on L2 Italian: one investigates the oral skills of Senegalese learners with different educational backgrounds in the country of origin; the other concerns the assessment of L2 writing skills in a multilingual group of refugees and asylum seekers. The results of both studies contribute to look at this peculiar target of learners from an unusual perspective, unveiling skills that can often be “invisible” to L2 literacy teachers.

KEYWORDS

L2 Italian; low-literate learners; oral skills; writing skills; L2 teaching.

1. INTRODUCTION

According to UNESCO most recent data (2017), worldwide there are still about 750 million people aged over 15 (women in 63% of cases) who didn't acquire reading and writing abilities in their first language or any additional language. Despite the great effort made by the UNESCO Institute of Statistics to obtain global data, the computation of this number is complicated by the absence of international standard tools to elicit data and to assess literacy levels and by the ambiguity of the definition of (il)literacy itself (Nitti, 2020)¹.

The lowest adult literacy rates (below 50% of the population) are observed in several countries of Southern Asia (49%) and sub-Saharan Africa (27%): Afghanistan, Benin, Burkina Faso, Central African Republic, Chad, Côte d'Ivoire, Ethiopia, Gambia, Guinea, Guinea-Bissau, Mali, Senegal, Sierra Leone etc. The highest rates (around 100%) are reported in the other regions of Asia, Europe and Northern America (UNESCO, 2017).

Data from the last census in Italy (Istat, 2020) confirm that illiterates represent 0.6% of the population (339.585 individuals), while 4% (about 2 million) is composed of literates without history of formal education and 16% (almost 9 million) is represented by people who only attended primary school. In order to obtain an overall picture of the illiteracy phenomenon in Italy, these data should be accompanied by those regarding non-native, migrant population, coming also from the most "illiterate countries", especially in recent years. Unfortunately, official data on levels of literacy, education and languages of migrants in Italy are still lacking (D'Agostino, 2017; Mocciaro, 2019) and some information can be only derived from reports of the SPRAR (System of Protection for Asylum Seekers and Refugees). In 2018, 12% of 41.113 guests of reception centres in Italy are described as illiterates, 63% as having a brief educational history (corresponding to Italian primary and lower secondary schools), while 19% attended high school (or equivalent) and only 6% university (Cittalia, Fondazione Anci, Ministero dell'Interno, 2019). In 2019, 15% of 21.108 migrants learning Italian language in the SPRAR centres were taking pre-literacy classes while 38.5% were attending a basic course (Cittalia, Fondazione Anci, Ministero dell'Interno, 2020).

Despite the presence of low-literate and illiterate learners is not a new phenomenon for teachers in L2 Italian classes (Minuz, 2005), it has gained importance with the recent migrations towards Italy and other European countries. People from rural areas of the world, often affected by violent con-

¹ In this contribution the UNESCO definition of (functional) literacy will be adopted: «A person is functionally literate who can engage in all those activities in which literacy is required for effective functioning of his (or her) group and community and also for enabling him (or her) to continue to use reading, writing and calculation for his (or her) own and the community's development» (UNESCO, 2005, p. 22).

flicts, are entering societies in which every kind of communication is based on the written medium and they are therefore experiencing a semiotic shock (Adami, 2009). For those, learning to read and write in the second language is an arduous task but it's obviously essential to be able to interact with the literate community and to conduct an autonomous life.

Literacy acquisition and literacy teaching in a second language have recently become subjects of interest in the academic European context, as demonstrated by the introduction of the Pre-A1 Level descriptors in the *Companion Volume of the Common European Framework of Reference for Languages* (Council of Europe, 2020), as well as in Italy (Borri et al., 2014)².

2. ILLITERATE OR LOW-LITERATE ADULT LEARNERS OF A SECOND LANGUAGE

L2 literacy classes for adults are usually very heterogeneous in terms of learners' profiles, with a great amount of variables of different nature. Minuz (2005) and Borri et al. (2014) proposed the following distinction in:

- pre-literate: learners whose first language doesn't have a written codification or it is not used as language of education in their country of origin;
- illiterate: learners who didn't develop reading and writing skills in any language;
- low-literate (or semi-literate): learners who have a brief history of formal education (usually less than five years). They are technically able to read and write in at least one language but they can't use literacy skills in daily communicative situations (they are not functionally literate).

Other linguistic variables that must be taken into account in an educational context are the kind of writing system of the learners' mother tongue (Latin, non-Latin, alphabetic, logographic) and the typological distance between L1 and L2, which can both have an impact on the perception of familiarity with the new language and particularly with the new written code to be learned. As we will observe in the following paragraph, also the didactic approach to which learners have been exposed in the few years of schooling could give important information to L2 literacy teachers.

Moreover there could be a high variability also in terms of oral abilities already acquired in the second language (null, initial, basic or intermediate).

Finally, personal variables (such as age, motivation, presence of disabilities or experience of trauma) and contextual factors in the country of arrival

² Previous researches mainly focused on L2 English and Dutch have been carried out especially by members of the international association LESLLA (Literacy Education and Second Language Learning for Adults) founded in 2005.

(nature and frequency of contacts with the L2, domains of use of the second language) can obviously influence the L2 acquisition process.

Pre-literate, illiterate and low-literate learners of a second language are usually described in literature as complex and difficult learners. Some of their features are:

- the slowness in the process of second language acquisition, with frequent failures in achieving high levels of L2 competence and obtaining official language certifications, so important for them in order to regularize their situation in the host country (Kurvers & Stockmann, 2009);
- the lack of metalinguistic skills both in L1 and L2 and the difficulty in focusing on linguistic forms instead of on their meanings (see, among the others, the studies on grammatical judgment by Van de Craats, Kurvers & Young-Scholten, 2006);
- the scarcity of abstraction ability, due to a concrete, pragmatic way of thinking, closely connected to objects and experiences that can be directly observed (on this topic, see the study on syllogism by Kurvers, 2002);
- the frequent lack of logical and numeracy skills;
- peculiar educational needs, also regarding the developing of motor and spatial orientation skills and, of course, of competencies related to formal classroom expectations and to studying techniques and learning strategies (Ardila, Roselli & Rosas, 1989; Minuz, 2005);
- an uncertain motivation to second language learning, both instrumental and integrative, that in some cases can be easily undermined by a sense of frustration and insecurity due to a previous unsuccessful history of formal education.

According to the European guidelines (Council of Europe, 2017) L2 teaching approach in the case of this specific target must necessarily be learner-centred³. Taking into account the characteristics and the communicative needs of learners, L2 literacy teachers must propose educational paths strongly linked to the concreteness of learners' lives outside the classroom context (Peyton, Moore, Young 2010; Brichese, 2018). Adults must be made aware of the objectives of the language acquisition process, in which learning to read and write in the L2 means also to develop linguistic autonomy, self-confidence and self-esteem. Valorising learners' past experiences and (also poor) linguistic abilities previously developed in formal or informal contexts of education and in any language become therefore necessary and essential.

In this respect, two studies conducted at the University of Naples "L'Orientale" and focusing on L2 Italian low-literate learners will be presented and discussed in the following paragraphs.

³ For a recent in-depth look at the topic of L2 Italian literacy teaching, see Caon & Brichese, 2019.

3. UNVEILING ORAL SKILLS: THE CASE OF SENEGALESE LEARNERS OF L2 ITALIAN

The first is a study conducted by Maffia & De Meo (2015) aiming at investigating the oral skills of low-literate Senegalese learners of L2 Italian.

Senegalese immigrants in Italy represent a small community: in 2020 they were 106.198 (2.1% of all foreign population – Istat, 2020), mostly resident in Northern regions of the country and represented by adult and not married men. Nevertheless, they have always been and still are a very visible minority group, especially for their common occupation as local street vendors and for their strong ability to create community based support structures and link with other ethnic groups (de Filippo, 2003).

Senegalese usually present a rich sociolinguistic repertoire, as it is often observed in people from countries with a history of colonization. In Senegal, French, in the two varieties of Standard and Non-standard, is the linguistic legacy of the colonial period, with the former usually associated with high education and prestige, the latter usually spoken in market places and other informal situations (Ngom, 1999). French is until today the official language of the country although it is spoken only by 10% of the population. The most widely spoken language is Wolof instead, which is the first of several national languages (Pulaar, Mandinka, Noon, Serer, Soninke, Arabic etc.). All these languages present a high vitality in oral communication but they have received a standardised Latin orthography only after the independence, in the early Seventies. Moreover, they are nowadays in the process to be introduced in the Senegalese education system, but exclusively in the first years of primary school (Universalia, 2019). As a consequence, in Senegal literacy skills are developed in a second language, French or Arabic, depending on the kind of school attended. While French schools propose a “European” didactic approach, that gives priority to the acquisition of literacy abilities and to the development of metalinguistic awareness, the educational approach adopted in Qur’anic schools (*daaras*) is mainly based on oral learning and memorization, that conversely gives priority to oral skills (speaking and listening), trained through the reciting of Qur’an (Gandolfi, 2003).

3.1 PARTICIPANTS AND METHODS

In order to evaluate if and how these different educational contexts in the country of origin could influence the acquisition of L2 Italian oral skills, 20 subjects were involved in the study: 10 learners who attended French school in Senegal and 10 learners who attended Qur’anic school (average of 7 years school attendance for all). They were all male, aged between 20 and 40, liv-

ing in Italy for a period ranging from 1 to 7 years and they all indicated Wolof as their mother tongue. At the moment of the research they were all attending L2 Italian classes offered by a voluntary association in Naples. Their literacy levels in French and Italian languages were assessed through standardised tests and resulted very poor. A test for Arabic was not administered because they all declared they were unable to read and write Arabic or to use it in real communicative situations.

An elicited imitation task was constructed and administered to all the participants: they were asked to listen once to 18 Italian utterances and to imitate them immediately after, as accurately as possible, regardless of the effective understanding of their meaning. The model utterances presented different degrees of complexity in terms of morpho-syntactic structures, length, lexical frequency and prosodic contours (assertion, questions and orders).

The entire corpus of imitations was object of two different kinds of analysis:

- an error analysis, conducted by a group of 10 experienced teachers of L2 Italian without any competence in phonetics, aimed at evaluating L2 utterances in terms of accuracy and kinds of errors;
- a spectroacoustic analysis, conducted by a phonetician through specific software, aimed at observing rhythmical and prosodic characteristics of Senegalese learners' imitations, in comparison to the utterances produced by the Italian models.

3.2 RESULTS

The results of the error analysis showed a higher level of accuracy in the imitations produced by French school learners, when compared to the Qur'anic school learners' productions. This difference in the performance of the two groups of subjects, somehow expected, was probably due to superior (even if very basic) analytic skills developed by learners in the context of French school education. Both in the case of simple and complex model utterances, French school learners were able to produce a higher number of complete and correct imitations, with lower percentages of errors, particularly in the case of segmental pronunciation and lexicon.

However, the spectroacoustic analysis revealed a "hidden" skill in the group of learners who attended in Senegal the Qur'anic school. Compared to the other learners, they were found to better imitate the intonational contour of simple and short model utterances, especially in the case of assertions and questions. Moreover, even when they were not able to correctly or entirely reproduce long and complex Italian utterances, they did not interrupt the imitations but instead they used a peculiar strategy in order to preserve the rhythmical structure of the original utterance: they correctly imitated the first

and the last syllables of the model utterances, while in the central portion they produced a meaningless sequence of hypo-articulated syllables, called “mumbling”, as in the two examples reported below (Maffia, Pettorino & De Meo, 2015).

Model utterance: *Fossi in te, non avrei la presunzione di essere impeccabile.*

Imitation (speaker 15): *Fossi in te xxx xxxxxxxxxxxxxxxx cabile*

Model utterance: *Perché usi ancora il cucchiaino di plastica?*

Imitation (speaker 18): *Perché usare xxx di plastica?*

This study unveiled oral skills in a group of Senegalese learners of L2 Italian, specifically linked to the prosodic competence developed in the context of Qur’anic school education that can be normally “invisible” to teachers. A pilot study on L2 Italian teaching to this specific target of learners demonstrated that taking into account their peculiarities, grounding on them the entire educational path, can have a positive effect on language learning motivation and improve L2 literacy acquisition process (Maffia & De Meo, 2017).

4. UNVEILING WRITING SKILLS: FUNCTIONAL ADEQUACY IN L2 ITALIAN TEXTS OF VULNERABLE LEARNERS

In this paragraph the results of another study focusing this time on the writing skills development of low-literate refugees and asylum seekers learners of L2 Italian are reported (De Meo, Maffia & Vitale, 2019). The low level of literacy in the first language and, often, a brief and uncertain history of formal education in the country of origin are only two of the several aspects that contribute to the definition of vulnerability of this group of L2 learners. Personal features such as experience of trauma, anxiety, depression, lack of concentration, fragility, perception of invisibility and isolation can characterize refugees and asylum seekers, and have, of course, a negative impact on the motivation to L2 learning (Bigelow & Schwarz, 2010; SPRAR 2010; Gordon 2011; Galos et al., 2017; Nitti, 2018).

The research originated from the observation of frequent failures in passing the written task of L2 Italian Certification Exams in classes for refugees and asylum seekers at the CLAOR, the Linguistic Centre of the University of Naples “L’Orientale”. The study had, therefore, a twofold objective: firstly it intended to monitor the development of writing skills in low-literate vulnerable learners of L2 Italian in the context of formal education; moreover and secondly it meant to test the effectiveness of different assessment methods of L2 Italian writing skills for this specific target.

4.1 THE CORPUS AND THE TWO SCALES

In order to reach these goals, 50 refugees and asylum seekers were involved in the research (only 5 women), coming from 16 different countries of Northern and sub-Saharan Africa and Southern Asia and with 20 different mother tongues (Arabic, Bambara, Bangla, Bissa, Mandinka, Urdu, Wolof, Yoruba etc.). A group of low-literate learners (0-8 years of schooling in the home country) was distinguished from a literate group (9-18 years), in order to understand if the variable level of literacy could have an impact on the development of writing skills and on the effectiveness of different assessment methods.

At the moment of the research, all the involved learners were hosted in an Extraordinary Reception Centre and in SPRAR centres in Naples and they were all attending L2 Italian basic classes. Their writing skills were monitored in a period of six months through the regular administration of written tasks. A corpus of 450 written productions, composed by narrations, descriptions and written interactions, was collected and all the texts were evaluated by experts using two different rating scales:

- a traditional scale, focused mainly on formal accuracy and in which the analytical criteria used in L2 Italian Certifications were considered and rated (language use, morpho-syntactic correctness, vocabulary, spelling and punctuation);
- a functional adequacy scale, assessing the ability of the writer to successfully transfer information and the socio-pragmatic appropriateness of his/her production. This scale is composed of four global dimensions - content, task requirements, comprehensibility, coherence and cohesion (Kuiken & Vedder, 2017; Vedder, 2016).

4.2 RESULTS

The first result of the study was a confirmation that writing skills development of low-literate refugees and asylum seekers learners of L2 Italian is a long and slow process. Nevertheless, a specific focus on L2 writing in the classroom context was found to be very helpful to improve quality and appropriateness of written productions in the second language.

Furthermore, through the application of the functional adequacy scale, higher and more stable scores were obtained, even at the very beginning of the observation and especially in the assessment of low-literates' productions and in the case of narrations and written interactions. While with the traditional scale the low scores given by raters in particular to the parameter of

morpho-syntactic correctness resulted in very negative evaluations, the functional adequacy scale appeared to be a reliable and efficient tool for valorising also poor writing skills, instead of underlying the limits of low-literates' productions.

Such results led to suppose that an assessment method focused on socio-pragmatic appropriateness of a written (but also oral) production, independently from the formal accuracy of grammatical structures, could be effectively used in classroom context but above all in L2 Italian Certification Exams for low-literate vulnerable learners, at least for A1 and A2 levels. This could avoid further educational failures that, in the case of this specific target, risk to have a disastrous effect on second language learning motivation.

5. CONCLUSIONS

In this contribution two different studies have been presented: the former focusing on the oral skills in the L2 Italian of Senegalese learners; the latter on the assessment method used to evaluate written texts produced by a heterogeneous group of refugees and asylum seekers attending Italian language classes offered by reception centres in Naples.

Despite being very different in terms of objectives, methodologies and linguistic materials analysed, the two studies have at least three aspects in common.

The first one regards the participants involved: adult low-literate learners of a second language, a target that has been neglected for a long time, drawing only recently the attention of academic community in Europe and also in Italy, due to the general growth of the migrant population and to the influx of refugees and asylum seekers also from countries with low literacy rates.

The second aspect is that both studies represent attempts to change perspective when describing low-literate learners' abilities: instead of pointing out to what they lack, these researches try to unveil and valorise what they can actually do with language, their "hidden" and maybe unexpected oral and writing skills.

Furthermore, what the two studies are also sharing is that they are deeply rooted in the L2 classroom experience and have important implications for teaching, constituting metaphorical bridges between the university and the different contexts in which L2 Italian courses are often provided, such as reception centres or voluntary associations.

REFERENCES

- Adami, Hervé. 2009. "The Role of literacy in the Acculturation Process of Migrants", in *Case Studies: prepared for the Linguistic Integration of Adult Migrants Seminar*, 19-27. Strasbourg: Council of Europe Publishing.
- Ardila, Alfredo; Rosselli, Monica & Rosas, Patricia. 1989. "Neuropsychological assessment of illiterates. Visuospatial and memory abilities". *Brain and Cognition* 11, pp. 147-166.
- Bigelow, Martha & Schwarz, Robin Lovrien. 2010. *Adult English Language Learners with Limited Literacy*. Washington DC: National Institute for Literacy.
- Borri, Alessandro; Minuz, Fernanda; Rocca, Lorenzo & Sola, Chiara. 2014. *Italiano L2 in contesti migratori. Sillabo e descrittori dall'alfabetizzazione all'A1*, Quaderni della ricerca, XVII. Turin: Loescher.
- Brichese, Annalisa. 2018. "Lo studente adulto analfabeta e semi-analfabeta. Alfabetizzare in lingua seconda". *Educazione Linguistica. Linguistic Education EL. LE* 7/1, pp. 7-24.
- Caon, Fabio & Brichese, Annalisa (eds.). 2019. *Insegnare italiano ad analfabeti*. Turin: Bonacci editore.
- Cittalia; Fondazione Anci; Ministero dell'Interno. 2019. *Atlante SPRAR/SIPROIMI 2018. Rapporto annuale SPRAR/SIPROIMI 2018. Sistema di protezione per titolari di protezione internazionale e per minori stranieri non accompagnati*, < <https://www.retesai.it/wp-content/uploads/2019/11/Atlante-Sprar-Siproimi-2018-leggero.pdf> >; Site consulted on 17/08/2021.
- Cittalia; Fondazione Anci; Ministero dell'Interno. 2020. *Atlante SPRAR/SIPROIMI 2019. Rapporto annuale SPRAR/SIPROIMI 2019. Sistema di protezione per titolari di protezione internazionale e per minori stranieri non accompagnati*, < <https://www.retesai.it/wp-content/uploads/2020/12/ATLANTE-SIPROIMI-2019.pdf> >; Site consulted on 17/08/2021.
- Council of Europe. 2017. *Language support for adult refugees: a Council of Europe toolkit*, < www.coe.int/en/web/language-support-for-adult-refugees/list-of-all-tools >; Site consulted on 17/08/2021.
- Council of Europe. 2020. *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*, Strasbourg, Council of Europe Publishing, < www.coe.int/lang-cefr >; Site consulted on 17/08/2021.

- D'Agostino, Mari. 2017. "L'italiano e l'alfabeto per i nuovi arrivati". *Testi e Linguaggi* 11, pp.141-156.
- De Filippo, Elena. 2003. *Gli immigrati nella città di Napoli, Centro di cittadinanza sociale per immigrati*, Naples: Cooperativa Dedalus..
- De Meo, Anna; Maffia, Marta & Vitale, Giuseppina. 2019. "La competenza scritta in italiano L2 di apprendenti vulnerabili. Due scale di valutazione a confronto". *Educazione Linguistica. Linguistic Education EL. LE* 8/3, pp. 637-653.
- Galos, Eliza; Bartolini, Laura; Cook Harry & Grant, Naomi. 2017. *Migrant Vulnerability to Human Trafficking and Exploitation: Evidence from the Central and Eastern Mediterranean Migration Routes*. Geneva: International Organization for Migration.
- Gandolfi, Stefania. 2003. "L'enseignement islamique en Afrique noire". *Cahiers d' études africaines* 43, pp. 261-277.
- Gordon, Daryl. 2011. "Trauma and Second Language Learning Among Laotian Refugees". *Journal of Southeast Asian American Education & Advancement* 6/1, < https://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=1029&context=j_saaea >; Site consulted on 17/08/2021.
- Istat. 2020. *Il Censimento permanente della popolazione e delle abitazioni. Prima diffusione dei dati definitivi 2018 e 2019*, < <https://www.istat.it/it/archivio/251687> >; Site consulted on 17/08/2021.
- Kuiken, Folkert & Vedder, Ineke. 2017. "Functional Adequacy in L2 Writing: Towards a New Rating Scale". *Language Testing* 34/3, pp. 321-336.
- Kurvers, Jeanne. 2002. *Met ongeletterdeogen*. Amsterdam: Aksant.
- Kurvers, Jeanne & Stockmann, Willemijn. 2009. *Watwerkt in Alfabetisering NT2. Leerlast en succesfactoren*. Tilburg: Universiteit van Tilburg.
- Maffia, Marta & De Meo, Anna. 2015. "Literacy and prosody. The case of lowliterate Senegalese learners of L2 Italian", in: van de Craats, Ineke; Kurvers, Jeanne & van Hout, Roeland (eds), *Adult literacy, second language, and cognition. LESLLA Proceedings 2014*, 129-147. Nijmegen: Centre for Language Studies..
- Maffia, Marta & De Meo, Anna. 2017. "Tecnologie per l'analisi del parlato e alfabetizzazione in italiano L2. Il caso di immigrati senegalesi adulti". *Italian Journal of Educational Technology (IJET)* 25/1, pp. 86-93.
- Maffia, Marta; Pettorino, Massimo & De Meo, Anna. 2015. "To mumble or not to mumble. Articulatory accuracy and syllable duration in L2 Italian of Senegalese learners", in: Russo, Domenico (ed.), *The Notion of Syllable across History, Theories and Analysis*, 354-371. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Minuz, Fernanda. 2005. *Italiano L2 e alfabetizzazione in età adulta*, Rome: Carocci.
- Mocciaro, Egle. 2019. "Emerging Constructions in the L2 Italian Spoken by Low Literate Migrants". *Languages* 4/86, pp. 1-24.
- Ngom, Fallou 1999. "A sociolinguistic profile of the Senegalese speech community". *Studies in the Linguistic Sciences* 29/1, pp. 131-146.
- Nitti, Paolo. 2018. "I bisogni linguistici nei corsi di italiano L2 rivolti ad utenti vulnerabili. Un'indagine sui corsi di lingua seconda erogati dai centri di accoglienza in Piemonte". *Educazione Linguistica. Linguistic Education EL.LE* 7/3, pp. 413-428.
- Nitti, Paolo. 2020. *L'alfabetizzazione in italiano L2 per apprendenti adulti non nativi*, Quaderni di Expressio, II, Milan-Udine: Mimesis.
- Peyton, Joy Kreeft; Moore, Sarah Catherine K. & Young, 144 Sarah. 2010. *Evidence-*

- Based, *Student-Centered Instructional Practices*. CAELA Network Brief, Washington, DC: Center for Applied Linguistics, < www.cal.org/caelanetwork/pdfs/EvidenceBasedFinalWeb.pdf >; Site consulted on 17/08/2021.
- SPRAR, Sistema di protezione per richiedenti asilo e rifugiati. 2010. *Le dimensioni del disagio mentale nei richiedenti asilo e rifugiati. Problemi aperti e strategie di intervento*, < www.sprar.it/wp-content/uploads/2017/02/Ricerca-Le-dimensioni-del-disagio-mentale-Cittalia-2010.pdf >; Site consulted on 17/08/2021.
- UNESCO. 2005. *Education for all: literacy for life; EFA global monitoring report, 2006*, < <https://unesdoc.unesco.org/ark:/48223/pf0000141639> >; Site consulted on 17/08/2021.
- UNESCO. 2017. *Literacy Rates Continue to Rise from One Generation to the Next*. Fact Sheet n. 45, September < <https://bit.ly/2xOPli2> >; Site consulted on 17/08/2021.
- Universalialia. 2019. *Summative Evaluation of GPE's Country-level Support to Education, Batch 4, Country 11: Senegal FINAL REPORT*, < www.globalpartnership.org/content/summative-evaluation-gpes-country-level-support-education-senegal-august-2019 >; Site consulted on 17/08/2021.
- Van de Craats, Ineke; Kurvers, Jeanne & Young-Scholten, Martha. (eds.). 2006. *Low-educated adult second language and literacy acquisition: Proceedings of the inaugural symposium Tilburg 2005*. Utrecht: LOT.
- Vedder, Ineke. 2016. "Il ruolo dell'adeguatezza funzionale nelle produzioni scritte in lingua seconda: proposta per una scala di valutazione" in: Santoro, Elisabetta & Vedder, Ineke (eds.), *Pragmatica e interculturalità in italiano lingua seconda*, 79-92. Florence: Franco Cesati Editore.

Part 3

Relationships between orality and writing in multilingual contexts, linguistic policies and case studies

Farsi Language and Iranian immigration in Europe

ADRIANO V. ROSSI

Professor Emeritus, Università degli studi di Napoli “L’Orientale”

ABSTRACT

Following a short outline of the history of the Iranian languages and Farsi, the Author analyses the different waves of Iranian emigrations before and after the Islamic Revolution. Data on the more recent Iranian emigration in Europe and Italy are reviewed, and a short discussion on the cultural and translational problems connected to emigration closes the essay. A bibliography of some 20 titles on the problem is also appended.

KEYWORDS

Farsi; Iranian emigration; Iranian languages.

Fârsi, a Western variety of Persian (also known as Modern Persian, Iranian Persian; while *Fârsi* is the official denomination in the Iranian Constitution) is an Indo-European language spoken in Iran since time immemorial, and by minorities of different consistence in Iraq and the Persian Gulf states,

Armenia and the Caucasian area, the Indian Subcontinent and (former Soviet) Central Asia¹.

A brief premise is necessary to explain why we call 'Persian' what the Iranians call *Fārsi*. Until 1934, the toponym for Iran used in Europe was that which had stabilized in the culture of European countries since the time when reliable news began to flow about the organization of that kingdom, that is, roughly from the Renaissance (or just before).

This toponym Persia was a creation of Greek geographers, prior to Herodotus but made popular by the *Histories* of this author, and later adopted by geographers and more generally by men of culture, and therefore by their respective languages, throughout Europe. However, no one in the West was aware that the name 'Persia' was used only in foreign languages other than *Fārsi* to indicate the country or its government: in *Fārsi* and in Iran, instead, starting around the time of the Ilkhanids (1256-1335), the Sassanid concept of 'Irān' (derived from the Middle Persian toponym *Ērānšahr* 'Iranian homeland') had been revitalized.² Of course, at the time of the first contacts between Greeks and Iranians, the ruling dynasty, that of the Achaemenids, had royal seats in Pārsa (today's Fārs region) and the king proclaimed himself 'King Pārsa' (adjective to the King) or, at times, also 'king in Pārsā' (locative of the region). The reports of the early Greek travellers were the cause of the extension of this dinastic (and perhaps originally tribal) name to the country, its inhabitants and its main language³. In essence, for millennia Iran has been known by its own heteronym, in the same way that Italy has been called *Włachy* in Polish for centuries, or the entire Indian subcontinent is called in Europe and in the world by the name of its westernmost region, Sind/Hind,⁴ the region that was first reached by Western travelers.

Fārsi is the predominant and the only official language of Iran (even if protection of the local cultures of Iran is mentioned in the Constitution). Eastern Persian varieties (known as *Dari* Persian, Afghan Persian, or *Dari* [its constitutional denomination], where is one of the two official languages, along with Pashto, recognised by the Afghan Constitution) are spoken and understood in Afghanistan. *Dari* is spoken by about 50% of the Afghan population as a

¹ For an overview on the centuries-old process of settlement of Persian population on the Plateau and subsequent expansion on the surrounding areas (the Persianate World), see Spooner 2012.

² On this see Fragner 1989, p. 89 ff.

³ In addition to *Fārsi*, numerous other languages of both Iranian and other language families (mainly forms of Turkish and Arabic) are spoken in Iran.

⁴ To be more exact, the name 'India' dates back through Ionic Greek Ἰνδός and Old Persian *Hinduš*, Old Indian *Sindhu*.

native language, but used in everyday interaction by much more than this⁵. While *Fârsi* and *Dari* are in theory the ‘same language,’ *Dari* is what we can call the ‘old way’ of speaking Persian, while *Fârsi* is the more modern one. If messages are only shared in one dialect or the other, common people might miss some keywords. However, for interpretation purposes – if you have a *Dari* / *Fârsi* speaker, he should be able to interpret for both⁶.

In general, Persian immigration came in Europe in two different waves, before and after the Islamic Revolution of 1979. Iranians who emigrated before the 1979 Revolution were mostly students while those who left the country after the Revolution were mainly exiles or political refugees. The industrialization drives in the 1960s created a need for educated and skilled labour in Iran. Despite a rapid growth in the enrollment of college students, Persian universities could not absorb the large number of high school graduates. Thus, Iran became one of world’s major exporters of college and university students, most of whom pursued higher education in the advanced industrial countries of Europe and North America. In comparison to students, there were fewer economically motivated immigrants since employment opportunities were in those years abundant in Iran. Although students initially were temporary migrants, they formed the original links in the subsequent migration chains of Iranians. After the Revolution, many students decided to settle abroad.

In the 1977-1978 academic year, about 100,000 Iranians were studying abroad, of whom ca. 36,000 were enrolled in U.S. institutes of higher education; the rest were mainly in the United Kingdom, West Germany, France, Austria, and Italy. In the 1978-1979 academic year, the number of Iranian students enrolled in the United States totaled 45,000, peaking at ca. 51,000 in 1979-1980.

The second wave consisted of a more sizable number of Iranians than the first. While the earlier exiles fled the Revolution, the later ones included whole families escaping the Iran-Iraq war (1980-1988). The exile outflow has continued, even in a much reduced form, to this day.

⁵ According to Kieffer 1983, p. 504, “a written language common to all educated Afghans, *Dari* must not be confused with *Kābolī*, the dialect of Kabul and surrounding areas that is more or less understood by eighty percent of the non-Persian speaking population and is fast becoming the nation’s koine.”

⁶ “Although *Paštō* has enjoyed official favor, it is little propagated among Persian speakers. First, it is difficult to proceed from a less complex (more analytical) language such as Persian to a more complex (more synthetic) language such as *Paštō*; and second, *Paštō* has been poorly taught, despite the efforts of the Ministry of Public Instruction. The tendency is, rather, for *Paštūns* to learn Persian in the course of their movements or during their military service, while the number of Persian-speakers who can express themselves in *Paštō* remains stationary” (Kieffer 1983, p. 504).

At the end of 2005, the United Nations High Commissioner for Refugees (UNHCR) estimated there were 111,684 refugees, asylum seekers, internally displaced persons (IDPs) and other persons of concern from Iran. The countries hosting the largest populations of Iranian refugees were Germany (39,904), the United States (20,541), Iraq (9,500), the United Kingdom (8,044), the Netherlands (6,597), and Canada (6,508).

The distinctive characteristic of this wave is the rise of asylum applications lodged in Europe. In 2004, Iran ranked tenth among the top countries of origin for asylum seekers across Europe. Fifty-five percent of the total Iranian asylum applications in 2000 were submitted in Western European countries, including the United Kingdom, Germany, Belgium, Austria, and the Netherlands.

In 2001 alone, there was a 300 percent increase in the number of Iranians seeking asylum in Britain. Furthermore, in 2004, Iran was the top nationality of asylum seekers to the UK, accounting for 10 percent of all applications. Preliminary observations indicate that these migrants often come from smaller provinces outside of Tehran and that they often face greater obstacles to integration in their host societies. Given that European Union countries have made it difficult to obtain asylum, Iranians who are not recognised as refugees often go to another country, remain illegally in the country where they applied, or return to Iran.⁷

Diasporas are rarely homogeneous groups, and the Iranian diaspora is no exception. Although the exact size of the diaspora remains unknown, a common yet disputed estimate of its size is two to four million people.⁸

Regardless of size, the Iranian diaspora is extremely heterogeneous with respect to ethnicity, religion, social status, language, gender, political affiliation, education, legal status, and timing and motivation for departure (ranging from political to sociocultural to economic).

In terms of ethnic origin, while the majority of the Iranian diaspora are authentically Persian in origin, there are also large communities of Azeris, Kurds, Assyrians, Turkmens, Baluchi and Armenians, who, speaking Fârsi as their medium language, tend to be perceived as generally 'Iranian' and Persophones. This ethnic diversity parallels linguistic heterogeneity, with large populations of Turkish-speaking Iranians especially in the peripheral areas of the country. A religious divide also exists between the majority, who are Shi'ia Muslims, and the minority groups, such as the Baha'is, Jews, Christians, Zoroastrians, and Sunni Kurds and Baluchi.

⁷ For demographic information and evaluation of the Iranian diaspora in Britain around the early 2000s see Spellman 2004, p. 37 ff.

⁸ Haghghat 2020; "Over 3.1 million Iranian-born people have emigrated from Iran, out of whom over 2.6 million (83%) have left the country since 1979" (Azadi et al. 2020, p. 22).

In the aftermath of the revolution, political divisions among supporters of the former Shah, groups such as the Mujaheddin, and apolitical groups decided for one reason or another to join the emigration flow. However, the heterogeneity of the diaspora is not a recent expression; on the contrary it mirrors the internal diversity long rooted in the homeland.

After the 1979 Revolution, some Iranians immigrated to Western Europe and especially to Scandinavia with the aim of acting there as political opponents. However, in the last two or three decades, Iranians from different social classes have migrated all over the world, with higher education as their main goal. At present, Iranians are dispersed in all parts of the world and have created what has been called “the two torn parts of the Iranian Nation,” which despite some similarities like language, history, and cultural heritage, are far more different in life style, while their mutual cultural and social influences are and remain important. The study of attractive areas for Iranians shows that the first immigration in recent times was that of the Armenian immigrants to Caucasus, and Iranian workers to Baku, Kuwait, UAE and other Arabic countries.

Even in Italy, the relatively recent phenomenon of Iranian emigration can be divided into two large waves: one before and one immediately following the Revolution of 1979. The first wave of migration, consisting mainly of university students, was essentially aimed at acquiring the skills necessary for the technical and scientific progress of the country. Although this type of migration was born in the 1960s as temporary, following the events related to the Revolution, many students decided to settle permanently in Italy or, moving from this country, to emigrate to other European countries. The second wave of migration, far more consistent than the previous one, consisted mainly of people fleeing for political reasons – including members of religious minorities, supporters of the monarchical regime and intellectuals of different backgrounds – with a much greater female and family component. Some scholars speak of a third wave of migration that began in the second half of the 1990s, consisting of both students, researchers and intellectuals, and lower-middle-class people looking for better job opportunities abroad (but the phenomenon has been comparatively little investigated).

Since there are no reliable data on the total Persian diaspora population in the world, the estimates of this population are exaggerated. For example, according to one source, “More than a million, and perhaps, as many as two million [Persians], remained in exile.” This source further quotes Ayatollah Ḥosayn-‘Alī Montazerī, a high-ranking official, as saying, “we have several million refugees abroad” (World Refugee Survey, 1988 in *Review*, p. 73). In the absence of comprehensive worldwide data, the data presented here are culled from national population censuses of some of the receiving countries. (...) Around

1990 there was a total of half a million Persians in Germany, France, Sweden, Norway, USA, Canada, Australia, and Israel, but if we include Turkey, they could well exceed one million” (Bozorgmehr 1996, p. 381)⁹.

It is now believed that there are about three million Iranians living outside the national borders. In Europe, the countries with the largest Iranian presence remain Sweden, Germany, the United Kingdom and France. Although the favorite destinations of the Iranians have been, since the pre-revolutionary period, the countries of Northern Europe, many (much more than those who have remained) have made a first stop in Italy, mainly for study reasons. Some of these have chosen to settle permanently on the peninsula, creating a family and even reaching important jobs.

According to recent estimates, it is assumed that more than 25 thousand Iranians reside in Italy. The prudent estimates available only take into account Iranian citizens who apply for a valid residence permit, while the total estimate also includes temporary visa-owner residents, illegal immigrants, nomads and detainees without passports, members of the diplomatic corps in a broad sense with their families, etc.

According to the Italian Ministry of the Interior, between 2006 and 2013 requests for residence permits from Iranian citizens increased by 2000, while the Italian municipalities of residence of Iranian citizens increased, in the same years, from 823 to 881. According to an age group breakdown in 2011, the most represented Iranian groups were those aged 25-29 and those aged 30-34. The smaller group corresponded to the 60-64 age group. The Iranian population was mainly concentrated in four regions: Lombardy (with 2,409 Iranians), Lazio (1,948), Emilia-Romagna (1,098) and Piedmont (1,009). Compared to the gender split, it appears that in 2014 53% of Iranians in Italy were men, in a substantial decrease, considering that only 15 years earlier, in 1999, the male population represented 70%.

Given the high percentage of the youth population in Iran and the difficulty of accessing university courses in Iran, many students decide to continue their studies abroad. Since Italian universities offer the possibility to enroll in numerous open-number courses and grant scholarships to foreign students on low incomes, every year about a thousand young Iranians arrive in Italy with the intention of enrolling in Italian universities. Of these, however, there are many who opt to give up their studies and choose to return to Iran during the first academic year.

In 2012/2013, 2,667 Iranian students (1,418 women and 1,249 men) were enrolled in the peninsula’s public universities alone, 9.79 percent of the total

⁹ Mehdi Bozorgmehr’s article, although generically entitled ‘Diaspora in the Postrevolutionary period,’ is centered, like many others in *the Encyclopædia Iranica*, on US data. For a recent evaluation of the Iranian diaspora in the USA see Mostafavi Mobasher 2018.

of the 27,246 foreign students. An interesting fact concerns the latest information available on the number of foreign graduates in 2014: out of 10,719 foreigners (3.5% of the total number of graduates), Iranians were in fifth place after Albanians, Chinese, Romanians and Cameroonians with 378 degrees obtained, amounting to 3.5% of foreign graduates.

With regard to the presence of asylum seekers from Iran, it should be stressed that the data provided by the Government on the number of refugees concern only the five nationalities from which the highest requests come. Since 1990, Iran has been in the top five countries of origin of asylum seekers only in 1994 (with 102 applications), 1995 (134), 1996 (46), 1997 (75) and 1999 (239).

There are also a very small percentage of nomads on Italian territory, most of them without identity documents, who declare themselves Iranians but who are most likely of Afghan origin. The largest community consists of about two hundred subjects residing in the municipality of Castel Volturno in the province of Caserta. I doubt very much that they can speak any Persian at all.

In order to fully understand the linguistic component of emigration from Iran to Europe, it must also be borne in mind that we often register as Iranians emigrants who arrived in Iran from two Middle Eastern realities from which large waves of migration date back about 40 years, namely Iraq and Afghanistan.

Though Afghans have a long history of visiting Iran as migrant workers, pilgrims, or merchants, the Soviet invasion in 1979 was a turning point in Afghan migration to Iran, resulting in a population of three million refugees at its peak in 1991.

The other major source of emigration to Europe is that of the Iraqi refugees. Although Iraqi refugees come from various backgrounds, they can be divided into three main groups: Iraqi Shiite, i.e. Muslim Arabs who were persecuted under Saddam Hussein's regime, Sunni Muslim Kurds, and Feili Kurds who Saddam's Iraq stripped of citizenship because their ancestors were from Iran. All fled Iraq to escape persecution under Saddam's regime.

The first Iraqi refugees arrived in the 1970s, mainly when Saddam's regime was crushing a Kurdish rebellion in northern Iraq. The Feili Kurds (who are Shiite, unlike most Kurds, who are Sunni) were declared Iranian by Saddam, even though Iran considered them Iraqis. The deportation of Feili Kurds continued in the 1980s during the Iran-Iraq War. The greatest number (700,000) arrived following the Halabja slaughter in 1988, when the Iraqi government used chemical weapons in the Iraqi Kurdish town of Halabja. Many members of this group returned home in 1992.

In both of these cases, hundreds of thousands of Afghan and Iraqi citizens who have learned only a few rudiments of Persian in the years of passage through Iranian refugee camps or other similar realities, use Persian in Europe only because their native language would have an even less chance of being understood by the European citizens, and therefore any program for

the recovery of their culture through their language would encounter even greater difficulties than those that can characterize an ethnically and culturally homogeneous group as for example Romanians or Sri Lankans in Italy.

With regard to this last aspect, that of cultural recovery, I would just like to mention an issue that has been studied at length by Joshua Fishman and more recently by Brian Spooner. On the basis of interviews with speakers of a variety of languages, Joshua Fishman found in his researches that the 'weak' (symbolic and reflective) and the 'strong' (essential and causal) connections between the native language and ethno-national subjectivity are expressed in parallel, but the latter connection tended always to predominate. "[W]ith [its own] particular chosen and beloved language," each ethno-cultural or ethno-national group "has felt, claimed and created an association and identity that are simultaneously experienced as eternally, creatively and inspirationally inseparable," to quote just a few words from one of Joshua Fishman's seminal essays on this matter.

Though Fishman does not refer directly to Persian, it is easy to understand why this notion of a 'beloved language' was picked up by Babak Elahi in an interesting essay on language and identity in Iranian-American writers (particularly in women's memoirs). In Iranian national culture (not only for Iranians proper, but in a more general sense for the whole Persianate World), Persian, which is the common language, is at the same time the repository for a long tradition of poetry, including many pre-Islamic themes of Iran's great national epic, Ferdowsi's *Shahname*, a poems of some 50,000 double-lines composed in Transoxiana in the 10th century CE. Today, the Persian language - as an ethno-linguistic resource as well as a cultural performance - continues to be important for Iranian exiles in the United States, in Europe and elsewhere. This privileging of the 'beloved language' must be complicated by Iran's multi-cultural and multilingual history, by the fact that many Iranians also speak forms of Turkish, variant dialects of Persian and Iranian non-Persid languages. Judging from the context of the United States, e.g., where the problem has been studied since many decades, one would venture to say that many Iranian Americans will eventually have the same relationship to Persian that third and fourth generation Italian Americans or other white, non-Latino ethnic groups have to their old-world languages.

Babak Elahi reminds us of the productivity in these contexts of the notion of 'writing with an accent' outlined by Taghi Modarressi in a short autobiographical essay:

The notion of an accented identity, and specifically of accented writing, can help us understand how Iranian-Americans might transform the trauma of a language lost into the celebration of a self regained or reconstructed, a translation of identity into a new language or through a dialectical relationship between two languages.

Although this is perhaps not the place to discuss the problems of the valorization of the sociolinguistic heritage of our students, no one can fail to see the great perspective implicit in projects like the one we have been involved in, in which every institution has tried to involve a congruous number of students in the realisation of its own dictionary and also create internal workshops on different languages, giving centrality to languages less known on the territory.

Thanks, therefore, to the organizers of this Project, who have allowed us to participate in such an unusual experience in the practice of university teaching, but so forward-looking in the slow progress towards the new multilingual Europe!

REFERENCES

- Azadi, Pooya, Matin Mirramezani, Mohsen B. Mesgaran. April 2020. *Stanford Iran 2040 Project: An Academic Platform for Research on Iran's Long-term Sustainable Development. Migrations and Brain Drain from Iran*, Working Paper No. 9, Stanford,.
- Bozorgmehr, Mehdi. 1996. "Diaspora. viii: Diaspora in the post revolutionary period," in: *Encyclopædia Iranica*, edited by Ehsan Yarshater, Vol. XIV, Fasc. 3, Costa Mesa, California: Mazda Publishers, pp. 380-383.
- Casari, Mario. 2012. "Iranians in Italy," in: *Encyclopædia Iranica*, edited by Ehsan Yarshater, Vol. XIV, Fasc. 3, New York: Encyclopædia Iranica Foundation Inc., pp. 290-292.
- Elahi, Babak. 2006. "Translating the Self: Language and Identity in Iranian-American Women's Memoirs," *Iranian Studies*, Vol. 39, No. 4, pp. 461-480.
- Fishman, Joshua. 1997. *In Praise of the Beloved Language: A Comparative View of Positive Ethnolinguistic Consciousness*, Berlin: De Gruyter Mouton.
- Fragner, Bert. 1989. "Historische Wurzeln neuzeitlicher iranischer Identität: zur Geschichte des politischen Begriffs 'Iran' im späten Mittelalter und in der Neuzeit," in: Maria Macuch, Christa Miller-Kessler & Bert G. Fragner (eds.), *Studia semitica necnon iranica Rudolpho Macuch septuagenario ab amicis et discipulis dedicata*, Wiesbaden: Harrassowitz, p. 79-100.
- Haghighat, Elhum. 2011, published 2020. "The Iranian Diaspora: Its Formation and Transformation," *Diaspora: A Journal of Transnational Studies*, 20, pp. 372-382.
- International Organization for Migration. 2003. *World Migration Report 2003: Managing Migration - Challenges and Responses for People on the Move*.
- Kieffer, Charles M. 1985. "AFGHANISTAN v. Languages," *Encyclopaedia Iranica*, Online Edition, 1982, available at <http://www.iranicaonline.org/articles/afghanistan-v-languages> (updated July 22, 2011, also available in print: *Encyclopaedia Iranica*, edited by Ehsan Yarshater, Vol. I, Fasc. 5, London-Boston-Henley: Routledge & Kegan Paul, pp. 501-516).
- Miggiano, Alice. 7 February 2017. "Chi sono gli iraniani in Italia?" *Il Caffè geopolitico*.
- Modarressi, Taghi. 1992. "Writing with an Accent," *Chanteb*, Vol. I, pp. 7-9.
- Mostafavi Mobasher, Mohsen (ed.). 2018.

The Iranian Diaspora: Challenges, Negotiations, and Transformations, Austin: University of Texas Press.

Roma nella storia e negli ultimi 50 anni, Internet Report retrieved on 23 May 2021.

Refugee Documentation Centre (Ireland)/Legal Aid Board, *Iran* – Researched and compiled by the Refugee Documentation Centre of Ireland on 14 August 2012.

Saint-Blancat, Chantal. 1988. *Nazione e religione fra gli Iraniani in Italia*, Padova: C.S.S.R.

Saint-Blancat, Chantal. 1989. “L’immigrazione iraniana in Italia: vera o falsa parentesi?”, in: Giovanni Cocchi (ed.), *Stranieri in Italia: caratteri e tendenze dell’immigrazione dai paesi extracomunitari*, Bologna: Istituto di studi e ricerche Carlo Cattaneo, pp. 109-125.

Saint-Blancat, Chantal. 1990. “La presenza iraniana in Italia,” *Inchiesta*, no. 90, pp. 59-67.

Spellman, Kathryn. 2004. *Religion and Nation: Iranian Local and Transnational Networks in Britain*, New York-Oxford: Berghahn

Spooner, Brian. 2012. “Persian, Dari, Farsi, Tadjiki: Language names and language policies,” in: Harold H. Schiffman (ed.), *Language Policy and Language Conflict in Afghanistan and its neighbors*, Leiden: Brill, pp. 89-120.

Vartanian, Vahed Massihi. 2013. *La comunità iraniana di*

Lingua e storia in Africa: considerazioni sul caso del tigrino (Eritrea ed Etiopia)

GIANFRANCESCO LUSINI

Università degli Studi di Napoli “L’Orientale”

ABSTRACT

Tigrinya (Təgrəñña) is a Semitic language spoken in Eritrea and in the Ethiopian regional state of Tigray (Təgray) by about seven million people. In Eritrea it is also working and school language and the medium of a rich literature. Here, one can find a brief discussion of the main issues concerning the history of language and literature, as well as the development prospects of an idiom used in a highly troubled African territory.

KEYWORDS

Tigrinya; Təgrəñña; Eritrea; Ethiopia; Tigray; Gə‘əz.

CLASSIFICAZIONE

Quando parliamo di Eritrea ed Etiopia, in Africa nord-orientale, ci riferiamo a due Paesi che complessivamente contano una popolazione di più di centodieci

milioni di abitanti, e che presentano una delle situazioni linguistiche più ricche e articolate di tutto il continente africano, con più di ottanta lingue, appartenenti a due distinte macro-famiglie, il nilo-sahariano e l'afro-asiatico (Fig. 1).¹

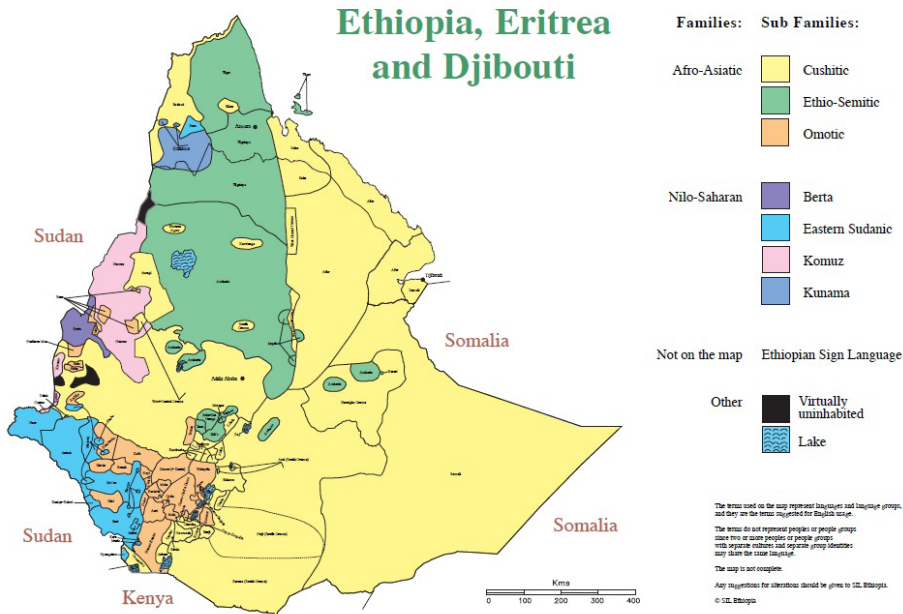


Figure 1. Distribuzione delle lingue attualmente parlate in Eritrea, Etiopia e Gibuti [riprodotto da <https://www.silethiopia.org/sil-ethiopia-publication-matters/>].

Del primo fanno parte una ventina di idiomi, parlati a cavallo della lunga frontiera che unisce l'Eritrea con il Sudan, e l'Etiopia con il Sudan e il Sud-Sudan. Della macro-famiglia afro-asiatica, che come indica il nome si estende su due continenti, tre rami (o all'inglese *phyla*, plurale di *phylum*) sono parzialmente presenti nei due paesi, ovvero il cuscitico, l'omotico e il semitico.² All'interno del ramo semitico della macro-famiglia afro-asiatica, fin dal secolo scorso (Renan 1863: 323-340) si è riconosciuta l'esistenza nella regione eritreo-etiopica di un gruppo di lingue che presentano nella fonetica, nella struttura grammaticale e nel vocabolario una grande affinità con tutte le lin-

¹ In passato indicato anche come camito-semitico, con una denominazione oggi totalmente abbandonata per le sue implicazioni extra-linguistiche, culturali e razziali (Sanders 1969).

² Gli altri tre rami, non presenti in Eritrea ed Etiopia, sono costituiti dall'egiziano (con la sua fase più recente rappresentata dal copto), oggi estinto, il libico-berbero e il ciadico.

gue semitiche note, sia quelle attestate fin dal III millennio a.C. e in massima parte oggi estinte, sia quelle attualmente parlate.³

Dunque, il semitico d'Eritrea e d'Etiopia è, insieme all'arabo e ai suoi dialetti, l'unica diramazione africana del semitico, e questo solleva immediatamente un rilevante interrogativo storico. Se, infatti, l'arabo è arrivato certamente in Nord-Africa in seguito all'espansione del Califfato (a partire dal 640), le origini del semitico d'Eritrea e Etiopia sono molto più antiche e tuttora non chiarite. Secondo la ricostruzione più accreditata, l'intera macrofamiglia afro-asiatica avrebbe avuto origine in Africa, nel Sahara centrale, forse tra il massiccio del Tibesti (Ciad settentrionale) e il Darfur (Sudan occidentale). Circa diecimila anni fa, a seguito del processo di desertificazione della regione, sarebbe avvenuta la sua frammentazione e la conseguente diaspora dei sei rami che oggi conosciamo (Diakonoff 1965: 102-107). Il semitico, dunque, deriverebbe da un'antica migrazione di gente dal Nord-Africa al Vicino Oriente, mentre il ramo etiopico, a sua volta, sarebbe conseguenza di un movimento a ritroso verificatosi a cavallo fra il II e il I millennio A.C., e diretto dalla Penisola Arabica verso l'Africa di nord-est. Il semitico d'Eritrea e d'Etiopia conta diverse lingue, ovvero (per limitarci a quelle pienamente documentate) il gə'əz (oggi estinto), l'amarico, l'argobba, il gafat (oggi estinto), il gurage, il harari, il tigré e il tigrino. Circa la sua storia molto resta da chiarire, ma da un punto di vista descrittivo, quindi senza una dimostrazione stringente (Bulakh, Kogan 2010), si ammettono comunemente due capisaldi: in primo luogo, la sua origine unitaria da un'unica proto-lingua formata nella Penisola arabica più di tremila anni fa; in secondo luogo, la distinzione tra due varietà fondamentali, denominate etiopico meridionale⁴ ed etiopico settentrionale. Di questo secondo ramo fan parte solo tre lingue, ovvero il gə'əz, il tigré e il tigrino. Poiché da un migliaio d'anni il gə'əz è uscito dall'uso vivo,⁵ l'etiopico settentrionale risulta costituito oggi da due sole lingue: il tigré, che conta circa un milione di parlanti, quasi tutti abitanti dell'Eritrea settentrionale e occidentale, e solo in piccola parte cittadini sudanesi, e il tigrino, usato quotidianamente da circa due milioni e mezzo di parlanti nativi in Eritrea centrale e da quattro milioni e mezzo nella regione etiopica del Tigray (Fig. 2).

³ Ovvero, oltre al semitico d'Eritrea e d'Etiopia, l'arabo, l'ebraico moderno, l'aramaico moderno (in un'area a cavallo fra Turchia, Siria, Iran e Iraq) e il sud-arabico moderno (Yemen e Oman).

⁴ Ne fanno parte l'argobba, il gafat (oggi estinto), il gurage, il harari e l'amarico, *working and school language* della Repubblica Federale Democratica d'Etiopia e lingua letteraria di notevole tradizione.

⁵ Anche se è sopravvissuta in quanto lingua liturgica e letteraria della Chiesa Etiopica Ortodossa *Täwähädo*, e fino ad oggi è oggetto di studio nelle scuole religiose d'Eritrea e d'Etiopia.



Figure 2. Areale del tigrino [riprodotto da John Huehnergard & Na'ama Pat-El (eds), *The Semitic Languages*, Second Edition. London and New York: Routledge, 2019, p. 174].

STORIA LINGUISTICA

Il glottonimo oggi universalmente impiegato dagli stessi parlanti è *təgrəñña*, da cui il termine tigrino deriva per adattamento (cf. l'inglese *tigrinya* o il francese *tigrina*). Esso è formato da un suffisso *-əñña* impiegato per indicare i nomi delle lingue (cf. *ənglizəñña* per l'inglese, *arəbəñña* per l'arabo, ecc.), preso in prestito dall'amarico e poi inglobato nel sistema morfologico del tigrino, soppiantando una formazione più antica con suffisso *-ay*, da cui deriva la forma *təgray*, oggi in disuso. La sua storia sembra risalire molto indietro nel tempo, ed è certamente connessa al toponimo Tigray (Təgray), indicante in passato una regione storica dell'Etiopia, e attualmente uno dei nove stati regionali della Repubblica Federale Democratica d'Etiopia. Dunque, in maniera certamente non casuale, i nomi delle lingue tigré (*təgre*) e tigrino (*təgrəñña*) e il toponimo Tigray (Təgray) appaiono etimologicamente collegati. E infatti, sulla base di testimonianze scritte in greco e in gə'əz, sappiamo che la base onomastica *təgr-* indicava una regione e i suoi abitanti fin da età altomedievale, se non già in epoca tardoantica.⁶ Non solo ancor oggi la grande maggioranza degli abitanti

⁶ Lo confermano inequivocabilmente alcuni testi antichi. In due manoscritti greci databili al X secolo, e che trasmettono un'opera geografica composta intorno al 547-549 (la

dell'Eritrea e del Tigray parla tigrino, ma proprio in queste regioni, tra I e VII secolo, sulle montagne intorno alla città etiopica di Aksum e nei bassopiani che scendono verso l'antico porto eritreo di Adulis (odierna Zula), affacciato sul Mar Rosso, è fiorita una civiltà monumentale, un regno africano alleato di Roma e di Bisanzio, la cui classe dirigente seppe trarre vantaggio allo stesso tempo dallo sfruttamento di risorse agricole e dai proventi di traffici commerciali. Successivamente, la regione del Tigray ha svolto un ruolo cruciale nella storia dell'Etiopia medievale e moderna, e la stessa Aksum, una volta perso il suo primato politico, ha mantenuto una posizione di rispetto fra i cristiani d'Etiopia in quanto città santa, luogo di pellegrinaggio, e sede della gerarchia, prima della fondazione di Addis Abeba. Nell'antica Aksum, la lingua dell'*élite* politica era il gə'əz, un idioma ben documentato ed estintosi quando il regno collassò, anche se coltivato fino ad oggi come lingua liturgica e letteraria dal clero cristiano. In effetti, dal punto di vista storico-linguistico vi è un legame non superficiale tra gə'əz e tigrino. Le convergenze fra le due lingue semitiche sono consistenti, senza contare che i territori dove attualmente si parla tigrino sono più o meno gli stessi dove in passato si parlava il gə'əz. Infatti, se le regioni dell'altopiano un tempo soggette all'autorità politica del Regno di Aksum oggi si trovano all'interno di due distinti stati, ovvero l'Eritrea e l'Etiopia, il *continuum* linguistico tigrino è indipendente dalle frontiere politiche attuali, e sostanzialmente si sovrappone all'areale del gə'əz.⁷

Al momento in cui si cominciarono a redigere testi scritti in tigrino, come già era avvenuto per l'amarico (e come avvenne contemporaneamente per il tigré), si ricorse al sistema di scrittura in uso da più di un millennio e mezzo per le opere in gə'əz. Alla base di questo fenomeno di conservazione e trasmissione di un elemento culturale tanto caratterizzante vi è certamente il fattore religioso, perché parliamo di territori che, a partire dal IV secolo, hanno condiviso l'adesione alla fede cristiana, e per tutta l'età medievale e moderna la scrittura del gə'əz è stata prerogativa di preti e monaci, che l'hanno praticata in via pressoché esclusiva. Ogni volta che si è posto il problema

Topographia Christiana di Cosma Indicopleuste), in un punto del testo che si riferisce proprio all'Etiopia settentrionale, un ignoto commentatore ha apposto una glossa (o scolio) che menziona le «genti dei Tigrētai» (ἔθνη τῶν Τυγρητῶν) (Wolska-Conus 1968: 375). Un'opera della letteratura etiopica medievale databile alla prima metà del XV secolo, il *Libro dei misteri del cielo e della terra* (*Māṣḥafä məstirä sämay wämədr*), parlando delle lingue dei discendenti del biblico Cam (cf. Genesi 10:6-20), menziona «il ḥabäsi, ovvero (la lingua dei) Təgray» (*ḥabäsi zäwə'atomu təgray*) (Perruchon 1903: 25), ricorrendo alla nota base onomastica *ḥabäš-*, ma in una forma parzialmente divergente a causa dell'alternanza *s/š* (cf. Smidt 2014).

⁷ Lo confermano le iscrizioni in gə'əz dei secoli III-IV, che sono state trovate a nord e a sud dell'attuale frontiera fra Eritrea ed Etiopia, costituita dal corso del fiume Märäb, a riprova del fatto che non vi è rapporto fra gli attuali confini politici e le antiche e moderne aree linguistiche.

di scrivere altre lingue di cristiani, come l'amarico, il tigrino e il tigré,⁸ il ricorso al sistema di scrittura del gə'əz è apparso ineludibile. Si tratta di un alfasillabario derivato da un *abjad* arabo pre-islamico e molto probabilmente influenzato dalla scrittura indiana *brahmi*. In base agli stessi principi, nel *fidäl* etiopico ogni grafema rappresenta una sequenza CV (vedi Appendice 1). Separatamente vanno considerati i simboli per i numerali, perché adattati a partire dalla lingua greca, in cui le lettere dell'alfabeto fungono anche da indicatori delle cifre (Meier 2016: 138-153).

In realtà, l'uso del *fidäl* per la scrittura di testi in tigrino (e in tigré) è un fenomeno molto recente. Pur essendo la lingua maggioritaria in Eritrea e Tigray, per secoli il tigrino è stato usato solo come strumento della trasmissione di una ricca letteratura orale (Kolmodin 1912, 1916; Conti Rossini 1942). Di fatto, gl'inizi della letteratura scritta si datano a epoca successiva al contatto con i missionari europei (luterani svedesi e cattolici italiani) e allo stabilimento delle prime due tipografie in Eritrea. La Stamperia Cattolica fu fondata nel 1863 a Massaua (prima pubblicazione nel 1867), poi trasferita nel 1879 a Cheren (Kärän) e nel 1912 ad Asmara. La stamperia della Missione Evangelica Svedese fu fondata intorno al 1870 a Moncullo (ጄmkullu), presso Massaua, e fu poi spostata ad Asmara. Conseguentemente, si osservano due fatti rilevanti. In primo luogo, i primi testi scritti e stampati in tigrino appartennero al genere dei catechismi e delle opere di edificazione spirituale, funzionali alle esigenze di propaganda religiosa delle due confessioni. In secondo luogo, il processo di letterarizzazione del tigrino ebbe origine in Eritrea prima ancora che in Tigray, perché nei principali centri eritrei erano attive le missioni religiose e europee, che avevano promosso la stampa di testi in lingue locali (tigrino e tigré).

STORIA LETTERARIA

In Eritrea, l'impiego del tigrino come strumento della comunicazione scritta, dopo aver ricevuto il primo impulso dall'attività delle due stamperie missionarie, fu promosso soprattutto da amministratori italiani, che avvertirono tutta l'utilità di impiegare, fra le lingue della Colonia, quella storicamente più collegata ai gruppi aristocratici dominanti la regione del Ḥamasen. Qui, infatti, si trovava l'antico borgo di Asmara (Lusini 2018), che fu scelto dalle autorità italiane come sede del governo centrale della Colonia (Fig. 3), e divenne poi capoluogo dell'Eritrea federata all'Etiopia, e infine capitale dello stato indipendente.

⁸ Sebbene il clero tradizionale etiopico abbia sempre contrastato l'estensione della scrittura a lingue diverse dal gə'əz, per cui i processi sono avvenuti col concorso essenziale di gruppi di provenienza europea, cattolici e luterani.



Figure 3. Asmara nell'anno 1884 [Archivio fotografico dell'Istituto Italiano per l'Africa e l'Oriente, Roma].

In Eritrea, il movimento letterario in tigrino ha conosciuto fino ad oggi particolare vivacità e forza, in misura da giustificare sintesi storico-critiche raffinate e di valore assoluto (Ghirmai Negash 1999).

Il 'padre' della letteratura tigrina è il *dābtāra* Fəśśəha Giyorgis Abiyäzgi (1868-1931), nativo del Tigray (Yəḥa, presso 'Adwa), ma attivo soprattutto in Eritrea, che si sforzò di sintetizzare formazione religiosa tradizionale e cultura europea. Nel 1890, per iniziativa delle autorità coloniali, si trasferì a Napoli per insegnare all'Istituto Universitario Orientale, sotto la guida di Francesco Gallina (1861-1942). Nel 1895, a Roma, egli licenziò l'opuscolo intitolato *Notizie del Viaggio di un Etiopico dall'Etiopia all'Italia* (Hailu Habtu 2000), il primo testo di natura non religiosa mai pubblicato in tigrino, e in seguito scrisse una *Storia d'Etiopia*, pubblicata cinquant'anni dopo la sua morte (Yaqob Beyene 1987).

Dopo la fine dell'occupazione italiana, l'uso della lingua fu ulteriormente promosso dalla British Military Administration (1941-52), con l'istituzione ad Asmara del Tigrinya Language Council (1944), e con l'avvio delle pubblicazioni di *Eritrean Weekly News* (1942), il primo periodico mai pubblicato in tigrino, diretto da Wäldä Ab Wäldä Maryam (1905-1995), che ne fece un luogo di discussione sul futuro politico e culturale dell'Eritrea.⁹ I decenni che precedettero lo scioglimento della Federazione Eritreo-Etiopica, con il

⁹ Il periodico cessò le sue pubblicazioni nel 1953 e divenne dapprima *Zämän*, 'Time' (in tigrino e arabo), al tempo della Federazione Eritreo-Etiopica, e successivamente *Həbrät*, 'Union'. Dalla dichiarazione dell'indipendenza eritrea (1991) il principale quotidiano eritreo è *Ḥaddas Ertəra*, 'New Eritrea' (Johnson 1996).

conseguente inizio della trentennale guerra di liberazione (1962), possono essere considerati i più produttivi nella storia della letteratura tigrina. Fra gli autori di spicco, citiamo almeno Musa Aron (1930-2011), il ‘gigante’ della letteratura eritrea. Figura di raffinato intellettuale africano, mosso da sincero zelo religioso in quanto pastore protestante,¹⁰ è stato autore di uno dei maggiori romanzi originali in tigrino, *Wärqaha* (1965), in cui sono trattati alcuni dei problemi tipici della società urbana, come alcolismo e prostituzione. Nell’ultimo ventennio, la poesia eritrea di lingua tigrina ha conosciuto uno straordinario sviluppo, caratterizzandosi per ampiezza dello spettro tematico e intensità di espressione, tanto che autori come Ribka Sibhatu, grazie anche a specifici progetti accademici, hanno raggiunto una notorietà internazionale (Cantalupo, Ghirmai Negash 2005).

LA SITUAZIONE MODERNA

Come si è visto, storicamente l’area linguistica qui considerata comprende il Tigray, le cui città principali sono Mäqäle, capitale dello stato regionale, Aksum, ‘Addigrat e ‘Adwa,¹¹ ma anche l’Eritrea centrale, in particolare le tre regioni dell’altopiano, il Ḥamasen, il Sära’è e il ‘Akkälä Guzay. Tutti questi territori mostrano un certo grado di omogeneità dal punto di vista geografico, etnografico, e naturalmente linguistico. Dunque, l’area dei parlanti il tigrino è sovranazionale e lo status politico della lingua dipende da svariate situazioni e fattori (Voigt 2011). In Eritrea, il tigrino è *working and school language*, usata nei documenti ufficiali dello stato insieme all’arabo (il paese è membro osservatore della Lega Araba), ed è lo strumento della comunicazione orale anche per i due milioni e mezzo di persone che parlano le altre otto lingue praticate nel paese,¹² nonché lingua letteraria nazionale, con un’intensa produzione nell’arco di più di un secolo (Hailu Habtu 1981; Abbā Agostinos-Tädlä 1994: 1-54, 239-506).

¹⁰ Il suo esordio letterario fu segnato dall’opera intitolata *Robinson Kruso* (1957), traduzione in tigrino del notissimo romanzo di Daniel Defoe *The Life and Strange Surprising Adventures of Robinson Crusoe* (1719), un testo ricco di evidenti spunti di meditazione filosofico-religiosa.

¹¹ Città nota perché nei suoi dintorni, il 1° marzo 1896, si combatté fra Italiani ed Etiopici un’aspra battaglia, che segnò il tramonto delle ambizioni coloniali del Regno d’Italia.

¹² Ovvero il tigré e l’arabo, lingue semitiche; il beja, il bilin e il saho e il ‘afar, lingue cuscitiche; il nara e il kunama, lingue nilo-sahariane (Fig. 4).

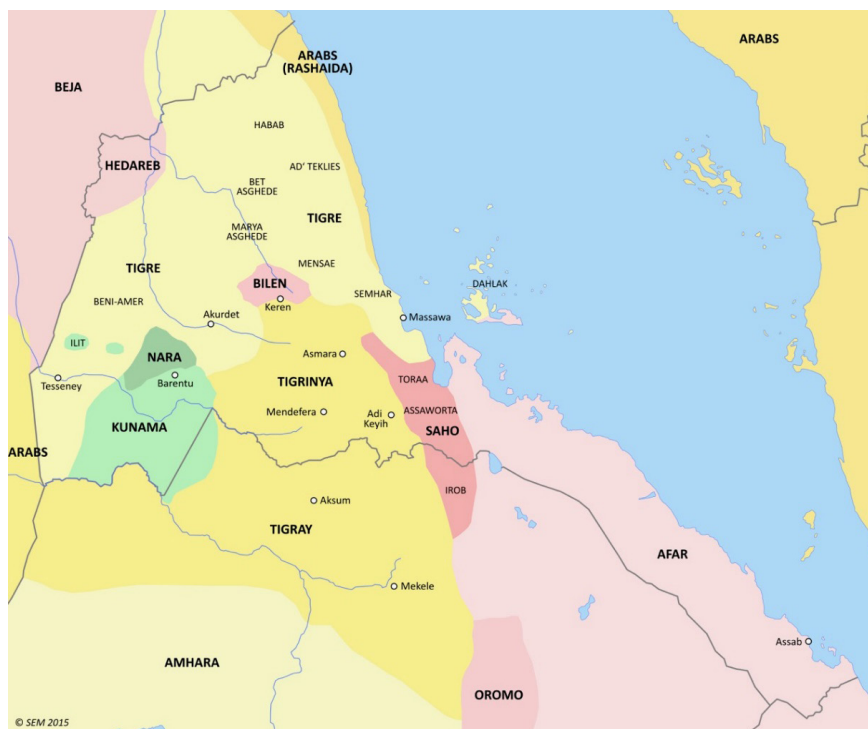


Figure 4. Distribuzione delle lingue dell'Eritrea e dell'Etiopia settentrionale [riprodotto da <https://www.easo.europa.eu/sites/default/files/public/Eritrea-Report-Final.pdf>, p. 13].

In Etiopia, un paese caratterizzato da un contesto linguistico particolarmente complesso, con più di settanta lingue ancora parlate, il tigrino è uno dei quattro idiomi maggiori. Viene dopo l'oromo, circa venticinque milioni di parlanti, l'amarico, circa venti milioni, il somalo, circa cinque milioni, e prima del sidamo, circa tre milioni, del wolaytta, del gurage e del 'afar, ciascuno con circa un milione e mezzo di parlanti.¹³ Malgrado il gran numero di persone che lo usano, e il suo radicamento in una regione come il Tigray, di primaria importanza per lo sviluppo storico-culturale di tutta l'Etiopia, paradossalmente il tigrino ha sempre stentato a emergere come lingua della comunicazione scritta. Per molti secoli, la gerarchia della Chiesa Etiopica Ortodossa *Täwəḥədo*, avente in Aksum la propria 'città santa', ha imposto come unico strumento della comunicazione scritta il *gə'əz*, lingua di millenaria tradizione, che ha espresso un'imponente letteratura religiosa, trasmessa da un monumentale patrimonio di manoscritti (Fig. 5).

¹³ Come espressione di una politica linguistica inclusiva, da marzo 2020 il governo federale etiopico ha stabilito che i documenti ufficiali debbano essere redatti, oltre che in amarico, anche in afan oromo, 'afar, somalo e tigrino.

E proprio nei codici in gə‘əz troviamo le prime e sporadiche attestazioni di elementi lessicali tigrini, dovuti a quegli amanuensi che, trascrivendo o componendo testi di indole religiosa, per motivi diversi fecero ricorso occasionalmente a parole o espressioni tratte dalla loro lingua madre.



Figure 5. Manoscritto in gə‘əz, da Gundä Gunde (Däbrä Garzen), ‘Agamä, Təgray [Parma, Biblioteca Palatina, Parm. 3838, mm. 152 x 115, Horologion, XIV secolo].

Successivamente, le possibilità di letterarizzazione del tigrino sono state condizionate dalla concorrenza dell’amarico, lingua della corte imperiale e dell’esercito, e veicolo dell’insegnamento religioso tradizionale (Lusini 2019). Come conseguenza, pur esistendo all’interno del tigrino una significativa differenziazione dialettale (Yaqob Beyene 1972; Voigt 2006, 2009; Tsehaye Kiros Mengesha 2009; Niguss Weldezzgu Mehari 2021), nessuna varietà è mai diventata lingua letteraria standard, fatto che va collegato proprio alla storia politica. Come si è visto a proposito del tigrino d’Eritrea, dove la varietà del Ḥamasen si è imposta nella seconda metà del XIX secolo, diversi fattori possono influenzare il processo di letterarizzazione di un dialetto. Di norma, la preminenza politica di un gruppo sociale o di un centro culturale favorisce la sua parlata anche come strumento creativo. In aggiunta, il contatto con le missioni europee è stato spesso responsabile dell’emergere di una varietà come lingua letteraria. Nel caso del tigrino d’Etiopia niente di tutto ciò è avvenuto,¹⁴

¹⁴ L’effetto più vistoso di questa dinamica è che a tutt’oggi il processo di normalizzazione ortografica del tigrino non può dirsi completato, malgrado l’esistenza di specifici organismi governativi di livello regionale (*in primis* la Təgray Language Academy) preposti, fra le altre cose, alla individuazione di forme di standardizzazione della lingua scritta.

malgrado ogni cittadino etiopico riconosca nel Tigray la regione che ospita le più antiche tradizioni storiche e religiose del paese. Poiché in età moderna il centro di gravità politica e religiosa si è spostato a sud, verso l'Amhara (Amara) e lo Šäwa (Scioa), l'amarico è stato riconosciuto come il vero erede del gə'əz, e si è creata la storica dicotomia fra *lassanä şəḥuf*, 'lingua della scrittura', cioè della letteratura (il gə'əz), e *lassanä nəguś*, 'lingua del re', cioè della corte e della classe dirigente del regno (l'amarico).

Nell'Etiopia moderna, a causa dell'intreccio fra lingua ed etnicità, lo status del tigrino è piuttosto oscillante. In primo luogo, osserviamo che in ambienti accademici è considerevolmente cresciuto l'interesse per lo studio di una lingua semitica non ancora perfettamente nota, ma che presenta una quantità di fenomeni particolarmente interessanti in chiave comparativa. Inoltre, negli anni Settanta e Ottanta i parlanti tigrino dell'Eritrea e del Tigray, anche se raccolti all'interno di diverse organizzazioni politiche e militari, hanno combattuto fianco a fianco contro il governo assolutista di Mängəstu Haylä Maryam (Menghistu Haylé Maryəm). L'omogeneità linguistica ha rappresentato un fattore fluidificante tra i combattenti, a partire proprio dai massimi dirigenti politici e militari, Isayəyyas Afäwäraqi (Isaias Afewerki), attuale presidente eritreo, e Mälläs Zenawi (Melles Zenawi), presidente (1991-1995), poi primo ministro etiopico (1995-2012), entrambi di madrelingua tigrina. Dopo la proclamazione dell'indipendenza eritrea (1993), divergenze politiche insanabili hanno portato allo scoppio della guerra fra i due stati (1998-2000) e a una situazione di stallo politico. In questa lunga e dolorosa fase storica, solo in Eritrea l'uso del tigrino come strumento di comunicazione nazionale e di creazione artistica non ha mai conosciuto battute d'arresto (Fig. 6).

ዓድና	Il nostro Paese
እንተ-ትዛረብ	Se venisse detto
ብቋንቋ ትግርኛ ብቋንቋ ዓረብ	in tigrino o in arabo
ብቋንቋ ኹሉ መላሽ መረብ።	in ogni lingua oltre il Märäb,
ምበለት ይመስለኒ	credo che significherebbe
ንቀረረብ	avviciniamoci
ሰብ ናብ ሰብ	essere umano a essere umano
ዝብኢ ናብ ገረብ	la iena all'albero

Fig. 6: ዓድና (Il nostro Paese) di Reesom Haile (1946-2003), in *Five Poems by Reesom Haile* [riprodotto da <https://sites.google.com/site/poetryintranslation2015>]. Traduzione di Gianfrancesco Lusini

Con la morte di Mälläs Zenawi (2012), a sua volta l'Etiopia è entrata in una fase di crisi politica che dura fino ad oggi. A partire dal 2018, con la nomina di Abəy Aḥməd 'Ali (Abiy Ahmed Ali) a primo ministro e di Sahlä Wärq Zäwde (Sahle Uork Zeudé) a presidente della Repubblica federale, e la conclusione di uno storico accordo di pace fra Eritrea ed Etiopia, il tigrino è diventato sempre più un fattore identitario per il Tigray e per i suoi dirigenti locali, che hanno vissuto con disagio la perdita del primato politico e la crescita di nuovi protagonismi, rappresentativi di altre nazionalità del paese, in particolare gli Oromo. Dopo l'esplosione del conflitto aperto tra il governo centrale etiopico e il gruppo dirigente locale del Tigray (novembre 2020), ogni questione linguistica è stata messa a tacere dal fragore delle armi e tutto lascia presagire che occorrerà non poco tempo prima che i programmi culturali riacquistino la priorità sui piani militari di entrambe le parti.

RIFERIMENTI
BIBLIOGRAFICI

- [Abbā] Agostinos-Tädlä, ofmcap. 1994. *La lingua abissina*. Asmara: Edizioni «Adveniat Regnum tuum».
- Bulakh, Maria & Kogan, Leonid. 2010. "The Genealogical Position of Tigre and the Problem of North Ethio-Semitic Unity". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 160, pp. 273-302.
- Cantalupo, Charles, Ghirmai Negash. 2005. *Who Needs a Story? Contemporary Eritrean Poetry in Tigrinya, Tigre and Arabic*. Asmara: Hdri Publishers.
- Conti Rossini, Carlo. 1942. *Proverbi, tradizioni e canzoni tigrine*. Verbania: Airoldi (Collezione scientifica e documentaria dell'Africa Italiana).
- Diakonoff, Igor Mikhailovich. 1988. *Semito-Hamitic Languages. An Essay in Classification*. Moscow: «Nauka» Publishing House.
- Ghirmai Negash. 1999. *A History of Tigrinya Literature in Eritrea: the Oral and the Written, 1890-1991*. Leiden: Research School CNWS.
- Hailu Habtu. 1981. *Aspects of Tigrinya Literature (until 1974)*. Thesis submitted for the degree of M.Phil. at the School of Oriental and African Studies. University of London.
- Hailu Habtu. 2000. "The Voyage of Däbtära Fesseha Giyorgis to Italy at the end of the 19th century". *Annales d'Éthiopie* 16, pp. 361-368.
- Johnson, Deborah. 1996. "Media History of Eritrea". *Eritrean Studies Review* 1, 1, pp. 143-154.
- Kolmodin, Johannes. 1912. *Traditions des Tsazzega et Hazzega. Textes tigrigna*. Rome: De Luigi (Archives d'Études Orientales 5, 1).
- Kolmodin, Johannes. 1916. *Traditions des Tsazzega et Hazzega. Traduction française*. Upsal: Appelberg (Archives d'Études Orientales 5, 2).
- Lusini, Gianfrancesco. 2018. "Asmara in ancient and medieval times: about the hidden story of a capital city". *Rassegna di Studi Etiopici* (3a serie) 2, pp. 183-195.
- Lusini, Gianfrancesco. 2019. "Lingua letteraria e lingua di corte: diglossia e insegnamento tradizionale in Etiopia fra Tardo Antico e Medio Evo". *AION-Sez. di filologia e letteratura classica* 41, pp. 274-284.
- Meyer, Ronni. 2016. "The Ethiopic script: linguistic features and sociocultural connotations". *Oslo Studies in Language* 8, 1, pp. 137-172 = Binyam Sisay Mendisu & Janne Bondi Johannessen (eds), *Multilingual Ethiopia: Linguistic Challenges and Capacity Building Efforts*.
- Niguss Weldezegu Mehari. 2021. *A Grammar of Rayya Tigrinya*. A Dissertation presented to

- the Department of Linguistics and Philology for the Degree of Doctor of Philosophy in Linguistics. Addis Ababa University.
- Perruchon, Jules (avec le concours de Ignazio Guidi). 1903. *Le livre des Mystères du Ciel et de la Terre. Texte éthiopien publié et traduit*. Paris: Firmin-Didot (Patrologia Orientalis I, 1).
- Renan, Ernest. 1863. *Histoire générale et système comparé des langues sémitiques. Première Partie. Histoire générale des langues sémitiques*. Paris: Michel Lévy Frères.
- Sanders, Edith R. 1969. "The Hamitic hypothesis, its origin and functions in time perspective". *Journal of African History* 10, 4, pp. 521-532.
- Smidt, Wolbert. 2014. "The Term Ḥabāša: An Ancient Ethnonym of the 'Abyssinian' Highlanders and its Interpretations and Connotations", in: Hatem Elliesie (ed.), *Multidisciplinary Views on the Horn of Africa. Festschrift in Honour of Rainer Voigt's 70th Birthday*, pp. 37-69. Köln: Rüdiger Köppe.
- Tsehaye Kiros Mengesha. 2009. *A Comparison of Wajerat Tigrigna vs. Standard Tigrigna*. A Thesis submitted to the School of Graduate Studies of Addis Ababa University in partial fulfilment of the requirements for the Degree of Master of Arts in Linguistics. Addis Ababa University.
- Voigt, Rainer. 2006. "Südtigrinische Dialekte: Das einfache und zusammengesetzte Präsens im Dialekt von May-Čäw (Tigray)", in: Siegbert Uhlig (ed.), *Proceedings of the XVth International Conference of Ethiopian Studies. Hamburg July 20-25, 2003*, pp. 893-898. Wiesbaden: Harrassowitz.
- Voigt, Rainer. 2009. "Südtigrinische Dialekte: Phonologie und Personal pronominal im Dialekt von May-Č'äw (Təgray)", in: Gideon Goldenberg, Ariel Shisha-Halevy (eds), *Egyptian, Semitic and General Grammar. Studies in Memory of H.J. Polotsky*, pp. 169-183. Jerusalem: The Israel Academy of Sciences and Humanities.
- Voigt, Rainer. 2011. "Tigrinya as National Language of Eritrea and Tigray", in: Stefan Weninger (ed.), *Semitic Languages. An International Handbook*, pp. 1170-1177. Berlin: De Gruyter Mouton (Handbücher zur Sprach- und Kommunikationswissenschaft 36).
- Wolska-Conus, Wanda. 1968. *Cosmas Indicopleustès. Topographie Chrétienne. Tome I (Livres I-IV). Introduction, texte critique, illustration, traduction et notes*. Paris: Les Éditions du Cerf (Sources Chrétiennes 141).
- Yaqob Beyene. 1972. "Le prepalatali nel dialetto tigrino del Sā'si' ('Agāmā)". *Annali dell'Istituto Orientale di Napoli* 32, pp. 359-65.
- Yaqob Beyene. 1987. *Fesseḥa Giyorgis, Storia d'Etiopia*. Napoli: Istituto Universitario Orientale.

Appendice 1. Inventario dei caratteri dell'alfasillabario del tigrino

	ä	u	i	a	e	(ə)	o	wä	wi	wa	we	wə
h	ሀ	ሁ	ሂ	ሃ	ሄ	ህ	ሆ					
l	ለ	ሉ	ሊ	ላ	ሌ	ል	ሎ					
ḥ	ሐ	ሑ	ሒ	ሓ	ሔ	ሕ	ሖ					
m	መ	ሙ	ሚ	ማ	ሜ	ሞ	ሟ					
ś	ሠ	ሡ	ሢ	ሣ	ሤ	ሥ	ሦ					
r	ረ	ሩ	ሪ	ራ	ራ	ር	ሮ					
s	ሰ	ሱ	ሲ	ሳ	ሴ	ስ	ሶ					
š	ሸ	ሹ	ሺ	ሻ	ሼ	ሽ	ሾ					
k	ቀ	ቁ	ቂ	ቃ	ቄ	ቅ	ቆ	ቇ	ቈ	቉	ቊ	ቋ
k ^h	ቐ	ቑ	ቒ	ቃ	ቄ	ቅ	ቆ	ቇ	ቈ	቉	ቊ	ቋ
b	በ	ቡ	ቢ	ባ	ቤ	ብ	ቦ					
v	ቨ	ቩ	ቪ	ቫ	ቬ	ቭ	ቮ					
t	ተ	ቱ	ቲ	ታ	ቲ	ት	ቶ					
č	ቸ	ቹ	ቺ	ቻ	ቼ	ች	ቾ					
ḥ	ሐ	ሑ	ሒ	ሓ	ሔ	ሕ	ሖ	ሠ	ሡ	ሢ	ሣ	ሤ
n	ነ	ኑ	ኒ	ና	ኔ	ን	ኖ					
ñ	ኸ	ኹ	ኺ	ኻ	ኼ	ኽ	ኾ					
˘	አ	አ	አ	አ	አ	አ	አ					
k	ከ	ከ	ከ	ከ	ከ	ከ	ከ	ከ	ከ	ከ	ከ	ከ
x	ኸ	ኸ	ኸ	ኸ	ኸ	ኸ	ኸ	ኸ	ኸ	ኸ	ኸ	ኸ
w	ወ	ወ	ወ	ወ	ወ	ወ	ወ					
˘	ዐ	ዐ	ዐ	ዐ	ዐ	ዐ	ዐ					
z	ዘ	ዘ	ዘ	ዘ	ዘ	ዘ	ዘ					
ž	ዸ	ዸ	ዸ	ዸ	ዸ	ዸ	ዸ					
y	የ	የ	የ	የ	የ	የ	የ					
d	ደ	ደ	ደ	ደ	ደ	ደ	ደ					

ğ	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ					
g	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ
ḡ	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ					
č	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ					
p	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ					
š	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ					
š	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ					
f	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ					
p	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ					

Appendice 2. Alcuni strumenti di base (grammatiche e vocabolari) per lo studio del tigrino

Alula Enkubahry. 2001. *English Tigrigna Idioms and Phrases Dictionary*. [Asmāra]: s.i.e.

Bulakh, Maria. 2019. “Tigrinya”, in John Huehnergard & Na’ama Pat-El (eds), *The Semitic Languages, Second Edition*, pp. 174-201. London and New York: Routledge.

Conti Rossini, Carlo. 1942. *Lingua tigrina. Parte prima: Elementi grammaticali ed esercizi*. Milano: Mondadori (Grammatiche e lessici delle lingue dell’Africa Italiana).

De Vito, Ludovico. 1893. *Esercizi di lettura in lingua tigrigna*. Roma: Casa Editrice Italiana.

De Vito, Ludovico. 1895. *Grammatica elementare della lingua tigrigna*. Roma: Tipografia Poliglotta della S. C. De propaganda Fide.

De Vito, Ludovico. 1896. *Vocabolario della lingua tigrigna. Introduzione e indice italiano-tigrigna del Dott. Conti Rossini Carlo*. Roma: Casa Editrice Italiana.

Francesco da Bassano. 1918. *Vocabolario tigray-italiano e repertorio italiano-tigray*. Roma: Casa editrice italiana di C. De Luigi.

Francesco da Offeio. 1935. *Grammatica della lingua tigray*. Roma: Arti grafiche G. Lolli.

Isaac Tseghay. 1997. *Mahder English-Tigrinya Dictionary*. [Asmāra]: s.i.e.

Kane, Thomas Leiper. 2000. *Tigrinya-English Dictionary*. Springfield: Dunwoody Press.

Kogan, Leonid. 1997. “Tigrinya”, in Robert Hetzron (ed.), *The Semitic Languages*, pp. 424-445. London –New York: Routledge.

Leslau, Wolf. 1941. *Documents tigrigna (Éthiopien septentrional). Grammaire et textes*. Paris: Klincksieck (Collection linguistique publiée par la Société de Linguistique de Paris 48).

Mason, John. 1996. *Tigrinya Grammar*. Lawrenceville NJ: The Red Sea Press.

Mauro da Leonessa. 1928. *Grammatica analitica della lingua tigray*. Roma: Tipografia Poliglotta Vaticana.

Praetorius, Franz. 1871. *Grammatik der Tigrīñasprache in Abessinien, hauptsächlich in der Gegend von Aksum und Adoa*. Halle: Verlag der Buchhandlung des Waisenhauses.

Tekie Tesfay. 2012. *Advanced Dictionary English-Tigrinya*. Asmara: Hdri Publishers.

Tesfay Tewolde Yohannes. 2002. *A Modern Grammar of Tigrinya*. Rome: U. Detti.

Tsehaye Teferra. 1979. *Reference Grammar of Tigrinya*. Unpublished PhD Dissertation. Washington DC: Georgetown University.

Ullendorff, Edward. 1985. *A Tigrinya (Təgrəñña) Chrestomathy. Introduction – Grammatical Tables – Tigrinya Texts – Letters – Phrases – Tigrinya-English Glossary – Select Bibliography*. Stuttgart: Franz Steiner (Äthiopistische Forschungen 19).

Voigt, Rainer. 1977. *Das Tigrinische Verbalsystem*. Berlin: D. Reimer (Marburger Studien zur Afrika- und Asienkunde 10).

Voigt, Rainer. 2011. "Tigrinya", in Stefan Weninger (ed.), *Semitic Languages. An International Handbook*, pp. 1153-1169. Berlin: De Gruyter Mouton (Handbücher zur Sprach- und Kommunikationswissenschaft 36).

Yimesgen Hailegiorgis. 2011. *Concise Tigrinya Figure of Speech Dictionary*. Asmara: s.i.e.

Zemicael Teclé. 2012. *Deutsch-Tigrinisches Wörterbuch*. Bearbeitet von Freweyni Habtemariam, Mussie Tesfagiorgis, Tedros Hagos und Tesfay Tewolde Yohannes. Wiesbaden: Harrassowitz (Äthiopistische Forschungen 78).

Transcription and orthography in two endangered languages of Ethiopia: Ts'amakko and Ongota

GRAZIANO SAVÀ

Università degli studi di Napoli "L'Orientale"

ABSTRACT

This paper touches upon the process of mapping spoken languages onto writing systems. Case studies relating to two endangered languages of Ethiopia, Ts'amakko and Ongota are presented. The discussion concerns two kinds of mapping, transcription for descriptive purposes and orthography for literary and literacy purposes. It is shown that transcription is more scientific and precise but less readable than orthography, that is more user-friendly for the wider public and the community of speakers.

KEYWORDS

Ts'amakko; Ongota; transcription; orthography; endangered.

1. TS'AMAKKO AND ONGOTA

Ts'amakko (also Tsamay and similar) and Ongota (also ſongota, Birale and similar) are found in the Bena-Tsemai district (*woräda*) of the South Omo Zone, which is one of the administrative sections of the Southern Nations Nationalities and Peoples Federal State of Ethiopia. Ts'amakko is a Cushitic (therefore Afroasiatic) language and belongs to the East branch of this group. The classification of Ongota is uncertain (see below).

Ts'amakko and Ongota are two endangered languages. They are part of the 28 languages that are declared in danger or extinct by the UNESCO in the Atlas of the World's Languages in Danger (<http://www.unesco.org/languages-atlas/index.php>). In particular, Ts'amakko is considered as definitely endangered, while Ongota is considered critically endangered. In terms of number of speakers Ts'amakko, according to the Atlas, counted 8621 speakers in 1998, while the Ongota speakers, were 8 in 2007. The official numbers by the Ethiopian government are controversial. The 2007 Population and Housing Census of Ethiopia (Central Statistical Agency of Ethiopia 2012) reports 17.390 Ts'amakko speakers in the South Omo Zone. However, according to the previous census of 1994 the Ts'amakko mother tongue speakers were 7.820 (Central Statistical Authority of Ethiopia 1996). Therefore, it is hard to believe that in ten years the number of Ts'amakko speakers was more than doubled. The statistics for Ongota are also to be considered with great care. According to the last 2007 census, the number of Ongota mother tongue speakers was 469 out of a population of 897. However, the last counting of the present writer in 2018 was a bit more than 100 for the ethnic group members with 7 speakers.

What is a speaker of Ongota should be better defined and should be done in relation to the Ts'amakko language. In fact, Ongota is so endangered because the people switched to Ts'amakko in everyday conversation. Those who still know Ongota are only a group of elders that use it scantily in situations in which they do not want to be understood, as a sort of secret language. Therefore, in everyday life, Ongota has a very little role.

The start of language switch was probably due to the decision of the present last speakers of Ongota to stop teaching it to their children and to teach them Ts'amakko instead. They already were bilingual due to the contacts with neighbouring Ts'amakko. Indeed, the Ongota live within the Ts'amakko territory, along the Wäyto River, inside the valley of this river. The Ongota were hunter-gatherers, fishermen and bee-keepers, but due to the contact with pastoralist Ts'amakko, they began agriculture and small-scale cattle breeding.

The alliance with the Ts'amakko seems to be one of several in the history of the Ongota. And each time there were if not a real language shift, at least heavy bilingualism and language influence. The result is that it is hard to find proper classification of Ongota, which presently is not classified, and which could be an isolate.

2. TS'AMAKKO DESCRIPTION, DOCUMENTATION TRANSCRIPTION AND ORTHOGRAPHY

Ts'amakko was described in *A Grammar of Ts'amakko* prepared by the writer as the accomplishment of a PhD at Leiden University (Savà 2005). The only noteworthy previous description, with comparative aims, is an article by Hayward (1989). The present paper focuses on the transcription used in Savà (2005).

The present paper will also deal with the Ts'amakko transcription system adopted in a documentation project which was part of a wider project on some Afroasiatic languages called *CorpAfroAs*, a *Corpus for Spoken Afroasiatic Languages: Prosodic and Morphosyntactic Analysis* (funded by the French Agence Nationale de la Recherche. Mettouchi, Vanhove & Caubet 2015).

An orthographic system was adopted for the preparation of a storybook in Ts'amakko. This is also described in the following paragraphs.

Finally, an orthographic system was designed and adopted for a literacy project in Ts'amakko carried out by the Summer Institute of Linguistics. This system will also be discussed.

2.1 GRAPHEMES AND PHONOLOGICAL INVENTORY OF CONSONANTS IN A GRAMMAR OF TS'AMAKKO

To start with, here is the chart of the Ts'amakko consonant phonemes in which phonemes are represented by IPA characters:

	Bilabial		Alveolar		Palatal		Velar		Uvular		Pharyng.		Laryng.	
	-v	+v	-v	+v	-v	+v	-v	+v	-v	+v	-v	+v	-v	+v
Obstruents														
Stops	p	b	t	d	tʃ		k	g						ʔ
Fricatives			s	z	ʃ	ʒ			χ		ħ	ʕ		h
Glottalic		ʙ	ts'	d'	tʃ'		g'	q'						
Non-obstruents														
Glides		w				j								
Lateral				l										
Trill				r										
Nasals		m		n		(ɲ)								

In the following chart the phonological transcription is “normalized” by spelling conventions that replace some IPA characters:

- /tʃ/ = <c>
- /tʃʰ/ = <cʰ>
- /ʃ/ = <š>
- /ʒ/ = <ž>
- /χ/ = <x>
- /j/ = <y>

Below is the chart showing the transcription of Ts’amakko consonantal phonemes as they appear in Savà (2005)¹:

	Bilabial		Alveolar		Palatal		Velar		Uvular		Pharyng.		Laryng.	
	-v	+v	-v	+v	-v	+v	-v	+v	-v	+v	-v	+v	-v	+v
Obstruents														
Stops	p	b	t	d	c		k	g						ʔ
Fricatives			s	z	š	ž			x		ħ	ʕ		h
Glottalic		ɸ	tsʰ	dʰ	cʰ		gʰ	qʰ						
Non-obstruents														
Glides		w				y								
Lateral				l										
Trill				r										
Nasals		m		n		(ɲ)								

The representation of some phonemes still requires the use of diacritics or special IPA characters. As can be seen, the choice was to indicate palatalization of the fricative sibilants with the haček (ˇ), while it was deemed superfluous to add an haček to the other sibilants <c> and <cʰ>. The apostrophe <ʰ> is used to mark ejectives, while for implosives the special IPA charac-

¹ As conventionally common, the // contain phonemes, the [] contain phonetic realisation and < > contain graphemes.

ters for these sounds are adopted. Specific IPA characters represent also the voiceless pharyngeal fricative /ħ/, the voiced pharyngeal fricative /ʕ/ and the voiced laryngeal stop /ʔ/, while the voiceless uvular fricative is simply <x> a simplification of IPA [χ]. The palatal nasal <ɲ> appears rarely only in loanwords.

All consonant phonemes can appear geminated. Gemination is represented by doubling the character.

2.1.1. CONSONANTAL GRAPHEMES AND ALLOPHONIC REALISATION

Several phonemes have allophonic realisations due to their position in the word or for free variation:

- all stops and glottalised obstruents are partially released word-finally. The only exception is /p/, which appears as [f], as we will see below;
- the voiced palatal fricative /ʒ/ is affricated as [dʒ] in postconsonantal and geminated position;
- the voiceless uvular fricative /x/ is trilled [χ̤] before high vowels;
- the voiceless pharyngeal fricative /ħ/ can also be trilled [ħ̤] and it is pronounced with a particularly powerful airflow;
- the voiced pharyngeal fricative /ʕ/ is often glottalised [ʕʔ] when in word-initial position and geminate;
- the voiceless alveolar ejective affricate [tsʼ] can be optionally pronounced as a fricative [sʼ];
- the ejectives /b/ and /g/ can be freely devoiced into [b̥] and [g̥] respectively;
- the position of the tongue in the articulation of the alveolar implosive /d/ is apical [ɖ] or laminal [ɗ];
- the uvular ejective /qʼ/ is quite unstable. It can be optionally pronounced either as an affricate ejective [qχʼ] or as a voiceless implosive [q̥] or a voiced implosive [ɢ];
- another problematic phoneme is what in the grammar is transcribed as <p> but that, in fact, corresponds to a phoneme that is articulated between bilabial stop and labio-dental fricative [f] or something in between. The variants partially depend on the position in the word. In word-initial position /p/ is pronounced as a plain stop [p], an aspirated stop [p^h] or a bilabial fricative [ɸ]. Intervocally /p/ appears as [f] or [ɸ]. In most of the cases both allophones are possible. In pre-consonantal position /p/ is realised as /f/. As for postconsonantal position, the /p/ can only be preceded by /m/, /r/ or /l/. When preceded by /m/

it is pronounced as [p^h]. When preceded by /r/ or /l/ it is pronounced as [f]. Final /p/ appears always as [f]. When geminated, /p/ is articulated as [p^h:] morpheme internally and [p^h:] or [f:] across morpheme boundaries. This brief account of the realisation of /p/ does not aim to be exhaustive but to show that this phoneme in Ts'amakko is problematic in deciding how to represent it in transcription. The decision in the grammar at the level of analysis is to leave the phonemic status of this phoneme as for its stricture parameter. At the level of transcription, it is decided to write it as <p> in initial, postnasal and geminated root-internal position and to write it as <f> in intervocalic, pre-consonantal, postconsonantal (except nasal), final and geminated across morpheme boundary position.

In all these cases, the grapheme corresponds to the representation of the phoneme regardless of the allophonic realisation. As we saw, an exception is /p/ that is actually represented by two graphemes according to its position.

Different decisions are taken if the phoneme changes due to some morphological processes. In these cases, phonological rules apply that cause a modification of the phoneme. In some cases, this modification is represented by the relevant grapheme, in some cases not. The decision depends on the fact that modification may correspond to an element already represented graphically in the transcription inventory. If there is no corresponding grapheme, the grapheme representing the phoneme is used. If the modification corresponds to an already existing grapheme, this is used. See some examples:

- the /t/ assimilates the voicing of a preceding /b/, /d/ or /g/. It becomes [d] and is transcribed as such: <d>;
- the implosives /ɓ/, /ɗ/ and /ɠ/ are devoiced before /t/. /ɓ/ and /ɠ/ become ejectives [p'] and [k'] respectively but are represented graphically as <ɓ> and <ɠ> because these ejectives are not present as phonemes and have no relevant character in the transcription system. /ɗ/, instead, not only is devoiced, but it loses glottalisation becoming [t]. In this position, therefore, it is represented as <t> since this grapheme is present in the transcription inventory;
- for the same reason, when /ɗ/ changes to [n], in contact with the suffix -ni of first person plural unmarked paradigm², it is transcribed as <n>;
- the alveolar nasal /n/ assimilates the place of articulation following a velar, becoming [ŋ], and a bilabial, becoming [m]. However, /n/ remains <n> because <ŋ> is not used else in the transcription system and in order to simplify the transcription avoiding the use of the IPA

² The Unmarked is one of the two main verbal paradigms in Ts'amakko. It is characterised by the fact that it is neutral in terms of aspect as it can be used to describe both perfective and imperfective actions.

character ŋ. A parallel solution was adopted in the case of the allophone [m] even if there is a grapheme <m> in the system;

- a different point regards the use of <ʔ> word initially. Since, according to Ts'amakko syllable structure, there are no syllables with no onset but all the syllables must be minimally CV, those words that appear vowel-initial actually carry an initial /ʔ/. This is therefore constantly transcribed initially.

2.1.2. VOWELS

The situation of the vowels is quite simple. Ts'amakko has five cardinal vowels, /a/, /e/ /i/, /o/ and /u/ with their long counterparts /aa/, /ee/, /ii/, /oo/ and /uu/ graphically represented in the same way. Lengthening as a consequence of a phonological rule occurs when a nominal is followed by the case clitic =*ma* (Savà 2020) or, optionally, by a locative case suffix. Vowel lengthening also indicates that a name is a possessor. See examples:

manne “house”

maanne = *ma* “to the house”

pašo “field”

paš-ilo or *paaš-ilo* “in the field”

beze “male name”

ħaarko beeze “Beze’s hand”

2.1.3. TONE

The tone system of Ts'amakko is based on a high tone and a low tone and has limited function load. It is not the case to make a full explanation of it in the present paper, as it suffices to state that, with some exception, tone is not marked in the Grammar of Ts'amakko. This is because the grammar provides information about the tonal characteristics of almost all word classes and verb paradigms and tone placement can be predicted from this information. An exception are those nominals that do not follow the default patterns HL or HHL. In these cases, tone is marked. The graphic rendering of tone is acute accent <´> for high tone and grave accent <`> for low tone.

2.1.4. A TRANSCRIPTION ALPHABET FOR A GRAMMAR OF TS'AMAKKO

For the sake of the present article, the consonants and the vowels as they are transcribed in *The Grammar of Ts'amakko* are put in the alphabetical order as follows:

a
b
b̄
c
c'
d
d̄
e
f
g
ḡ
h
h̄
i
k
l
m
n
ɲ
o
p
q'
r
s
š
t
ts'
u
w
x
y
z
ž
ʔ
ʕ

That makes 35 characters. Consider that all characters can also appear doubled to represent consonantal gemination and vocalic lengthening. Moreover, there is no distinction between majuscules and minuscules.

2.1.5. SAMPLE TEXT

In order to better exemplify the transcription system discussed above, here is a sample text extracted and adapted from *A Grammar of Ts'amakko*. Notice that punctuation marks are not used, except for the question mark, and each line corresponds to a sentence or a coherent sequence of sentences. Moreover, in *A Grammar of Ts'amakko* the texts are interlinearised with grammatical and lexical glosses, that are left over in the text below. An English translation follows:

maakke garrilo ?aaka maakke gubalatte
q'arra garro ?ardo bitami na gubale maarte bitamti
maarte gubalatte gurti ba woq'ošiba dāli
maarte daltinnay garro kiyānāy kaayu ka dāli kiyi ba ?ardilo turditte sorto ki
?adda gubale gore buskaḃti
gore zingatte sa?ate lākkīyay bukaḃe ba gubale sa?ate salaḥ ki xafti
gore kiyankinnay ?ine ka buskaḃti ba ?ato ?aakkama baḃay? kiye
gaaḥko ḡondami ba deelloma ?oladi ?ise ka kiyiti
gore kesse bukaḃi gaaḥko moo ḡondamu ba ka delay gubalatte kiye
?ise kiyannay ?ardo moo ki dālada nūnu bukaḃanku? kiyiti
tannu ?ise ka bolḡomisi
tannu garro kiya nay boḡolkonu q'ole c'oxinda
?ombottanne kīnkōyay ?axxe c'oxonkibba ganda xumbi c'oxankibba nūnu
?awkose boḡolko nata ?aḡīma šeeḡonki
boḡolko ?ombotto ?ugisi nay ?ombotto dookko ?akka ?uggisi nay ?ombotto
dookko ?akka ?uggisi nay ?ombotanne q'arū šeeḡe xumbi ?ugisi kiyankinnay
?ombotanne ketta xumbi rawti nay bolte takka ki ?aq'i
bolte tettakka gaaraḡtema beḡi
dawra boḡolko ḥaarkoyay ḡabbikka kup bayi ba q'eedda loq'a boḡoltekkā
kup bayiti nay garaḡte nata bo?ti
?asa tannu gubalatte mala dayi? gubale ka parti
maakke garrilo ?aaka maakke gubalatte kettay

The tale of the squirrel and the rabbit

One day the squirrel bought an ox and the rabbit bought a heifer. The heifer of the rabbit mated, got pregnant and gave birth. After the heifer gave birth, the squirrel said: "Mine gave birth", and put some placenta into the buttock of the ox. The rabbit made people gather. The people gathered at eight in the morning and the rabbit came at ten. The people said: "You made gather us

and where did you disappear?”. She said: “A stone broke and I spent the day sewing it”. The people who gathered said to the rabbit: “How does it come that a stone breaks and you sew it?”. After that, she was made queen. So, the squirrel said: “Milk cattle for the queen!”. They milked ten containers, the whole neighbourhood milked and brought to the place where the queen lived. He made the queen drink the milk containers. He made her drink one milk container and after that, he made her drink another milk container. Eventually, he made her drink all the milk containers they had brought to the place. After she had finished all those milk containers a small drop remained. “It is taboo! A queen does not take anything with the hands. She bends, licks and swallows”. When the queen bent, her belly blasted. So, then, what happened to the rabbit? The rabbit died. That was the tale of the squirrel and the rabbit.

2.2. TRANSCRIPTION FOR A DOCUMENTATION OF TS’AMAKKO

A more scientific, IPA-like, transcription was used for Ts’amakko in the context of the project *CorpAfroAs*, a *Corpus for Spoken Afroasiatic Languages: Prosodic and Morphosyntactic Analysis*. In fact, each text was transcribed two times. One transcription is strictly phonetic and the second one is more broadly phonemic. In order for the phonetic transcription to reflect the spoken language more faithfully, various IPA characters, signs and diacritics were used that are not possible to discuss in the present paper. The phonemic transcription is more readable and normalized and is discussed here. The characters are still IPA and the inventory similar to the one of *A Grammar of Ts’amakko*. The only differences are that there is no ejective /c’/ [tʃ̰] but only /c/ [tʃ] and the uvular /q’/ is not an ejective.

Transcriptions, annotation and translation were ordered in parallel tiers with the use of the ELAN-Corpa programme. This programme is a version created by Christian Chanard and his collaborators at the French CNRS research center LLACAN (Langues, Langage et Cultures d’Afrique Noire) of the ELAN produced and maintained by the Max Planck Institute in Nijmegen. The difference is that ELAN-Corpa has a special semi-automatic function of morpheme glossing.

From the screenshot below one can see that tier tx@SP1 (text of speaker 1) hosts the broad phonetic transcription and boundaries separate intonation units. The signs “/” and “//” are used to set the boundary of minor and major intonation units respectively, but this does not have an effect on the transcription. The characters used in the transcription are IPA characters that represent all possible nuances of the speech. The tier mot@SP1 (the French word “mot” for “word” is used to name this tier) has still a phonetic transcription with phonetic characters and shows all phonological modifications but it normalizes allophones to phonemes. On this tier boundaries separate

words. In the following tier, mb@SP1 (morpheme boundary of speaker 1) each segment hosts a morpheme, before any phonological modification. The following two tiers are glossing tier, one lexical and grammatical, one syntactic, and the last one is the translation tier. On top of everything there is a tier assigning a reference code to each intonation unit:

	00:00:26.500	00:00:27.000	00:00:27.500	00:00:28.000
ref@SP1 [312]	TSB_NARR_001_023		TSB_NARR_001_024	TSB_NARR_001_025
tx@SP1 [310]	ellexaye:: /		405	la:gǣŋki //
mol@SP1 [748]	ʔelleka	ke /	405	la:gǣŋki //
mb@SP1 [1095]	ʔelele	=ka ke /		la:g -g -anki //
ge@SP1 [1094]	togeth CONT PL /			retur SEM IPFV. //
rx@SP1 [1094]	ADV.M	FOC PRO /		V der.V TAM. //
ft@SP1 [...]	Together			they gathered [the animals]

It is to note that consonant length is indicated by doubling the consonantal character. Vocalic length is indicated by the IPA sign [:]. Moreover, the duration of the pauses was marked from 200 ms.

Below is the list of characters of the phonemic transcription and part of a sample text:

- a
- b
- β
- tʃ
- d
- ɕ
- d
- e
- f
- g
- g
- h
- ħ
- i
- j
- k
- l
- m
- n
- o

p
q'
r
s
ʃ
t
ts'
tʃ
u
w
χ
z
ʔ
ɣ

gelzakko ʔa:ka garro

gelzakko ʔa:ka garro leʔe ʔellema ka:rink ka:rinko goʃaɗa

gelzakko gelzakko ma:re goʃi

garro leʔe goʃi

gelzakko kaʔʔiba baj garro ʔato leʔeja zow

baj gelzo ma:reja zow kaji

ʔelleka ke la:gʃanki

ʔelle la:gʃennaj q'ajnakaja garro leʔeja zeji pugadiki ʔa:ga

agarro ko pugadi kiʔa:ga

q'ajnakaja likke ʔasa q'ole goʃaɗe goʃe

kuba ma:reja koba leʔejaj bajina garro ko pugadi kiʔa:ga

baj garro ʔato mo: ɕiʔtiba q'ajto χumbika pugadaj bajinaj

gelzakko kijanaj agarro kijana ʔanoʔane...

The squirrel and the baboon

The baboon and the squirrel used to tend the cattle together according to an agreement. The baboon was tending female calves. The squirrel tended cows. The baboon got up and “Friend squirrel. You go with the cows”. “Friend baboon. Go with the calves” [the squirrel] said. Together they gathered [the animals]. They gathered [the animals] and the day after the squirrel went with the cows, got satiated and went back home. The squirrel came back home being satiated. The day after right in the same way, they tended cattle. One [went] with the calves. The other one with the cows and the squirrel went back home being satiated. “Friend squirrel, what is that you eat and always get satiated?” He said and the squirrel said Who, me?...

2.3. ORTHOGRAPHY FOR A TS'AMAKKO FOLKTALE BOOK

There is an on-going project for the publication of some Ts'amakko folktales by the present writer and Dr Pavel Mikeš, the present Czech Ambassador in Addis Ababa³. The aim is to publish a booklet with a dozen of folktales for the wider public.

A proper Ts'amakko orthography to be taken as reference for similar projects was designed. Compared to the transcription used in *The Grammar of Ts'amakko* and in the *CorpAfroAs* documentation project, more readable solutions were adopted. IPA characters were avoided, except for the <ʔ> of the voiced pharyngeal fricative. As for the rest:

- the character of the voiceless laryngeal stop <ʔ> is represented by an apostrophe <'> and is not marked word initially;
- all glottalised, ejectives and implosives, are marked by the apostrophe besides the main character. This solution avoids using the IPA characters for the implosives;
- <ts'> was simplified to <s'>;
- the IPA character <ħ> for the voiceless pharyngeal fricative is replaced by <hh>.

Moreover:

- the haček is also avoided. <ž> is, therefore, represented by <j>, that, as we saw, corresponds to the affricated realisation of the voiced palatal fricative /ž/ in postconsonantal and geminated position. As for <š>, it is represented by the English-like <sh>;
- the distribution of <p> and <f> is the same as in *The Grammar of Ts'amakko*;
- all characters, consonantal and vocalic, can appear doubled to represent gemination and lengthening.

a A

b B

b' B'

³ Dr Pavel Mikeš, besides being a diplomat, is a great scholar with a deep knowledge of Ethiopia. He has a particularly strong relation with the Ts'amakko community, that he visited already in the early nineties. He is the one who introduced the present writer to the Ts'amakko people and facilitated his linguistic research work among them. For this, the present writer still feels deep gratitude.

c	C
c'	C'
d	D
d'	D'
e	E
f	F
g	G
g'	G'
h	H
hh	Hh
i	I
j	J
k	K
l	L
m	M
n	N
o	O
p	P
q	Q
r	R
s	S
sh	Sh
t	T
s'	S'
u	U
w	W
x	X
y	Y

z Z
, ,
ʃ ʃ

All 34 characters can be doubled to indicate consonantal gemination and vowel lengthening.

2.4. SAMPLE TEXT

Also in this case, a sample text is presented below:

Maakke gudurkilo aaka maakke gaarmilo

Garmo qooshi. Gudurkokka qooshu. Shambo aaka abbayo qooshi. Abbayo aaka shambo qoshebbā lo'o bog'inki. Lo'o ji'ankinnay lo'o ji'tilo ʃagankinnay garmo gaarkomi ka'iba

ufund'eka i'una shambose gudurkilo shambo gudurkilo kiyānay "abba ula guddo garmo na" kiyina gudurko kiyānay "ei ussa garmokka kuyo" kiyi. "Saankoka ji'nanki saanko garmokka kuyo" kiyi. Saanko ji'ankinnay joome ka'inki. "Abba ano intayikka ato intawu" kiyi shambo. Qaru garmoka ammake i'i donnay shukuyina gudurko intayi. Gudurko intayi ooda. Oodānay garmokkana miinatte guddonu ufund'e ka'inay garmokkanay miinatte guddonu ka'i. Suurema nunnu ridu. Abbayo suuratte lig'i aaka garmo asa abbayo ki bog'i. Shambose gudurkilo sori zowu. Garmo haysamma gudurko bog'i. Maakke garmilo aaka makke gudurkilo hayissayay d'ikkad'i.

The story of the hyena and the lion

The lion was hunting in the same field where the hyena and his son were hunting. The hyenas killed a cow and started eating it. While they were eating it the lion climbed on a tree and the hyenas could see him. The hyena child said "Father, over there is a lion, let's be careful". And the hyena father said "No, that is not a lion, it is a termite hill" and kept on eating the meat. While they were eating the father repeated "I am sure it is not a lion but a termite hill". After they got satiated, they left but the child said "Father I do not want to go first, you go first". So, the hyena father walked first. The lion followed their movements, started walking towards them and stopped waiting for them in a small path into the forest. The hyena father entered that path and the lion killed him. The young hyena ran away. So, the lion killed the father hyena. In this way ends the tale of the hyena and the lion.

2.5. TS'AMAKKO ORTHOGRAPHY FROM A SIL (SUMMER INSTITUTE OF LINGUISTICS) LITERACY PROJECT

SIL Ethiopia is involved in a literacy project in the Ts'amakko area. Training is organised in two locations, Luqa, the village where the present writer did his linguistic field research, and Birale, close to the Ts'amakko main town Wäyt'o. No further news on this project could be collected, besides the orthography adopted for the production of primers and other didactic material (provided by the local collaborator Haylu Berhanu Golla).

Compared to the orthography of the folktales presented above, this orthography is characterised by the following features:

- implosives and ejectives are marked by a <h> following the main character. Exceptions are <ts>, that indicate the ejective alveolar affricate, and <c>, that indicates the ejective palatal affricate /tʃ/. Its non-ejective counterpart is with an <h>, e.g., <ch>. This is a solution adopted in several orthography of languages in the area as the ejective counterpart of this affricate is much more common than the pulmonic one. Due to the higher occurrence of /tʃ/, the simpler character <c> is used;
- the voiced pharyngeal fricative /ʕ/ is represented by <v>. This is possible because there is no phoneme /v/ in Ts'amakko (as in all the languages of the Ethiopian area);
- the voiceless counterpart of /ʕ/, i.e., /ħ/, is represented by a simple <h>. This is possible because /ħ/ is not considered among the phonemes of Ts'amakko;
- there is a character combination <ny> for the palatal nasal /ɲ/;
- there is no <p>, only <f>;
- the SIL Ts'amakko orthography includes majuscule letters. In case of double characters, such as, for example, the implosive <bh>, only the first main character is majuscule, i.e., <Bh>;
- questions are indicated with the lengthening of the final vowel.

Here is the list of characters, minuscule and majuscule, in alphabetic order:

a	A
b	B
bh	Bh
c	C
ch	Ch

d	D
dh	Dh
e	E
f	F
g	G
gh	Gh
h	H
i	I
j	J
k	K
l	L
m	M
n	N
ny	Ny
o	O
q	Q
r	R
s	S
sh	Sh
t	T
ts	Ts
u	U
v	V
w	W
x	X
y	Y
z	Z
‘	‘

All characters can be doubled in order to represent consonant gemination and vowel lengthening. However, in those cases of double character, such as, for example, the implosive <dh>, only the first main character is doubled, i.e., <ddh>.

Unfortunately, it is not possible to provide a sample text.

3. ONGOTA

Besides being a very endangered language and a possible isolate, Ongota is largely understudied. There are only two reference articles as for description: Fleming et al. (1992/93) and Savà and Tosco (2000). The language has also been object of documentation (see below). Part of the documentation work was the creation of a quadrilingual alphabet for the community.

In the following, I will overview the transcription systems adopted in Savà and Tosco (2000) and the documentation of Ongota and the orthographic systems (Latin and Ethiopian) used for the alphabet.

3.1. ONGOTA DESCRIPTIVE ARTICLE

Savà and Tosco 2000 is the result of a fieldwork conducted by the present writer and Mauro Tosco in Jinka with the late Mole Sagane. As for the transcription system adopted, below is the chart of the Ongota consonantal phonemes in IPA characters:

	Bilabial/ labiodental		(Post) Al- veolar		Palato-al- veolar		Velar		Uvular		Pharyng.		Laryng.	
	-v	+v	-v	+v	-v	+v	-v	+v	-v	+v	-v	+v	-v	+v
Plosives														
Oral		b	t	d			k	g	q					ʔ
Glottalised				d̥				g̥						
Nasal		m		n										
Affricate			ts		tʃ	dʒ								
Fricative	f		s	z	ʃ				χ		ħ	ʕ		h
Trill				r										
Approximant														
Central		w				j								
Lateral				l										

In the transcription system used in the article, the following IPA characters were replaced:

/tʃ/ = <c>

/ʃ/ = <š>

/χ/ = <x>

/j/ = <y>

Here is an adaptation of the chart of Ongota consonant phonemes from Savà and Tosco (2000:66). The representation of the phonemes corresponds to the graphic representation:

	Bilabial/ labiodental		(Post) Al- veolar		Palato-al- veolar		Velar		Uvular		Pharyng.		Laryng.	
	-v	+v	-v	+v	-v	+v	-v	+v	-v	+v	-v	+v	-v	+v
Plosives														
Oral		b	t	d			k	g	q					ʔ
Glottalised				t̚				g̚						
Nasal		m		n										
Affricate			ts		c	j								
Fricative	f		s	z	š			x		ħ	ʕ		h	
Trill				r										
Approximant														
Central		w				y								
Lateral				l										

The transcription solutions are similar to those adopted in the *Grammar of Ts'amakko*. In part this is also due to the fact that the phonological systems of the two languages are similar and, understandably so, due to the close relationship that they have. The palate-alveolar fricative sibilant is marked by the haček (ˇ), while there is no haček on <c>. The apostrophe <'> marks ejectives and implosives are represented by IPA characters. IPA characters are also used for the voiceless pharyngeal fricative /ħ/, the voiced pharyngeal fricative /ʕ/ and the voiced laryngeal stop /ʔ/. The voiceless uvular fricative is simply <x> a simplification of IPA [χ]. As in Ts'amakko, all phonemes can appear geminated. Gemination is represented by doubling the character.

3.1.1. GRAPHEMES AND ALLOPHONIC REALISATION

Some phonemes have optional allophones that are not marked graphically:

- Voice opposition can be optionally neutralised. See examples of [h] ~ [ʔ] and [d] ~ [t]:

gaddahhuni [gaddahhuni ~ gaddaʔuni] “big”

gidata [gidata ~ gitata] “you (PL)”

- In the section on Ts’amakko we have seen the complexity of realisation of p/f. In Ongota the only (optional) alternation attested is between [f] ~ [p^h] word-initially:

oxoni faʔo [faʔo ~ p^haʔo] “to kindle the fire”

- The palate-alveolar affricate /j/ optionally becomes a fricative.

janta [dʒanta ~ ʒanta]

It is to notice that Ongota /j/ ([dʒ]) corresponds to Ts’amakko /ž/ ([ʒ]), that has an allophone [dʒ].

Here are some of the possible grapheme modifications due to phonological rules:

- A nasal preceding a plosive assimilates the point of articulation of the plosive. This modification is graphically represented:

tagamá “sleep (SG)!” *tagánta* “sleep (PL)!”

- Due to sibilant harmony, the /s/ of the causative suffixes -san, -is and -as becomes a palato-alveolar š if in the verbal stem there is a palate-alveolar consonant. This modification is represented graphically:

ka = cóq “I shot” *ka = cóqšan* (instead of **cóqsan*) “I made shoot”

Although the imperative plural suffix -ta causes devoicing assimilation to a preceding voiced plosive, the resulting modification is not graphically represented, i.e.,

yeqadá “hiccup (SG)!” *yeqadťá* “hiccup (PL)!” [yeqattá]

3.1.2. VOWELS

As for the vowels, the situation is like in Ts'amakko: there are five cardinal short and long vowels. The long vowels are indicated by doubling the vocalic grapheme. There are occurrences of initial vowels. Therefore, the situation is not as in Ts'amakko, where initial vowels are always preceded by the character of the glottalic stop <ʔ>. The initial glottal stop can be optionally heard, but in view of its irregular presence it is never marked.

3.1.3. ACCENT

There is no systematic marking of accent, which is indicated sporadically in the description when it causes contrast between two segmentally identical words or if it has grammatical meaning. It is marked by an acute accent.

An important grammatical situation in which the accent plays a role and is marked is the distinction between past and non-past. In the following example the accent falls on the penultimate syllable in the past verb form and on the last syllable in the non-past verb form:

ka = múxi "I laughed"

ka = muxí "I laugh/will laugh"

3.1.4. A TRANSCRIPTION ALPHABET FOR SAVÀ AND TOSCO (2000)

Below is an alphabet reconstructed from the description of Ongota by Savà and Tosco (2000):

a
b
c
d
ḍ
e
f
g
g̣
h
ḥ
i
j

k
l
m
n
o
q
r
s
š
t
ts
u
w
x
y
z
?
ʃ

Unfortunately, Savà and Tosco (2000) does not contain texts.

3.2. THE TRANSCRIPTION FOR THE DOCUMENTATION OF ONGOTA

A two-year postdoc grant for the documentation of Ongota was awarded in 2007 to the writer by the Hans Rausing Endangered Languages Documentation Programme. The project produced recordings, some of which transcribed, annotated and translated, of Ongota speech samples.

The transcription, annotation and translation methodology is the same as the one used for Ts'amakko in the project *CorpAfroAs*, a *Corpus for Spoken Afroasiatic Languages: Prosodic and Morphosyntactic Analysis* (see above paragraph 2.2).

3.2.1. A TRANSCRIPTION ALPHABET FOR THE DOCUMENTATION OF ONGOTA

The following points are to notice compared to the transcription of Ongota in Savà and Tosco (2000):

- consonant length is marked by doubling the character. Vowel length is indicated by the IPA mark “:”;

- initial vowels are always preceded by <?>;
- <j> indicates the voiced palate-alveolar central approximant, instead of <y>;
- <dʒ> indicates the voiced palate-alveolar, instead of <j>;
- according to a new analysis, <ts'>, <tʃ'> and <q'> represent ejective sounds and therefore carry the apostrophe “'”. Therefore, <tʃ'> corresponds to <c>;
- <ʃ> corresponds to <š> and;
- <χ> is the equivalent of <x>.

The following list of characters can be extracted from the “mot” tier:

a
b
d
dʒ
d'
e
f
g
g'
h
h̃
i
j
k
l
m
n
o
q'
r
s
ʃ
t
ts'
tʃ'
u
w
χ
z
?
ʔ

A text extracted from the documentation was published in *Ethnorêma* by the present writer (Savà 2017). The article, and the whole journal, can be freely consulted and downloaded.

3.3. TWO ORTHOGRAPHIES (LATIN AND ETHIOPIAN) FOR AN ONGOTA ALPHABETIZATION PROJECT

A booklet with a didactic alphabet was realised in the context of the ELDP documentation of Ongota. It consists of one page per letter. On the page there is the single letter in Latin and in Ethiopian script, or *fidäl*, a figure and the corresponding word in Ongota and in Ts'amakko, written in Latin and Ethiopian script, plus the word in Amharic and in English. The name of the language is written under each word. See an example:



As for the Latin script, compared to the transcription in Savà and Tosco (2000) diacritics and IPA characters have been avoided. As a consequence:

- the implosives are indicated by a <h> after the main character as in the SIL Ts'amakko orthography;
- the <š> is represented by <sh>;
- the voiceless laryngeal stop <ʔ> is reduced to ', but it does not appear word ini-tially;
- the voiced laryngeal fricative <ʕ> changes to <v>, as in the SIL system for Ts'amakko (see above paragraph 2.5.);
- the voiceless laryngeal fricative <ħ> is represented by hh;
- besides the <c> in Savà and Tosco (2000), there is an ejective marked as <c'>.

One may notice that the name of the Ongota language actually has a voiced pharyngeal fricative at the beginning, i.e., *ʕongota* [ʕoŋgota]. The spelling Ongota is used as a simplification.

As for the *fidäl* script, the typical syllabary system with change of vowel by modification of the basic character carrying the vowel /ä/ is adopted. Reference is made to the use of the script in Amharic. Moreover, the following remarks are in order (for the Latin transcription of words in *fidäl* it applies the system used in Savà and Tosco 2000, see above):

- initial vowels are represented by the voiceless laryngeal stop character ለ with relevant vocalic modification. The voiceless laryngeal stop is represented by the same character in the middle of the word. For example: ለለ *iiʔa* “hand”;
- the example above also shows that long vowel length is not marked. The same for consonantal gemination (see *hobatto*, below);
- the implosive /d/ is represented by a vocalized ደ preceded by the “sixth order” of the same consonantal character. For example: ደደላ *düla* “flour”. The same system is adopted for the other implosive /g/. For example: ግጊናግ *gĭnano* “mosquito”;
- as for the back fricatives, the system makes use of the Ethiopian characters that are used in Amharic even if they do not show difference in pronunciation. All of them represent indeed the voiceless laryngeal fricative /h/ in Amharic. Starting from /h/, this phoneme is represented by ሀ. For example: ሀባቶ *hobatto* “washing”. The voiceless pharyngeal fricative is represented by ሁ. For example: ሁንቻ *hanca* “tree”. The voiceless uvular fricative /χ/ is represented by the other Amharic “h”, i.e., ሐ. For example: ሐና *xoona* “sheep”;
- the uvular stop /q/ corresponds to the Amharic velar ejective ቀ. For example: ቀሮማ *qoroma* “rhinoceros”;
- the alveolar affricate /ts/ is indicated by the character of the alveolar ejective ጸ even if it is not ejective. For example: ጸናፋ *tsanafa* “six”.

Here is an alphabetic list of Latin and Ethiopian characters extracted from the pages of the booklet:

a	ለ
b	ቦ
c	ቸ
c'	ጮ

d	ᵈ
dh	ᵈːᵈ
e	ᵉ
f	ᶠ
g	ᵍ
gh	ᵍᵍ
h	ᵂ
hh	ᵂᵂ
i	ᶦ
j	ᶜ
k	ᵏ
l	ᶫ
m	ᵐ
n	ᵑ
o	ᵒ
q	ᶞ
r	ᵣ
s	ᵑ
sh	ᶜᶜ
t	ᵀ
ts	ᶜ
u	ᵁ
w	ᵂ
x	ᶞ
y	ᶜ
z	ᶜ

4. CONCLUSION

The decisions that are relevant to what kind of writing system to adopt for traditionally unwritten languages such as Ts'amakko and Ongota depend on the purposes of the writing system. Scientific transcription tends to use IPA characters and diacritics. However, there is a more phonetic and a more phonemic transcription style, depending on the kind of document that is created. The phonetic one is very detailed, using a large array of IPA characters and diacritics since it has to reflect the spoken language in as many details as possible. It has been mentioned but not treated in the present paper. The phonemic transcription is more regular and readable, even if it tends to use IPA characters and diacritics. Readability is the main principle for the designing of orthographic systems for literacy and literary aims. We have seen that, except for the voiced pharyngeal fricative represented by <ʕ>, no IPA characters are used and no diacritics. Other solutions, mainly the use of double characters, are adopted. In two cases a common character representing a consonant not present in the phonemic inventory of the language has been used for something else. It is the case of Ts'amakko and Ongota <v> for the voiced pharyngeal fricative /ʕ/ in the orthographic systems described.

REFERENCES

- Central Statistical Authority of Ethiopia. 1996. *The 1994 Population and Housing Census of Ethiopia: Results for Southern Nations, Nationalities and People's Region*. Addis Ababa: Central Statistical Authority.
- Central Statistical Agency of Ethiopia. 2012. *The 2007 Population and Housing Census of Ethiopia, Statistical SNNPR*. Addis Ababa: Central Statistical Agency.
- Fleming, Harold, C. et al. 1992/93. "Ongota or Birale: A moribund language of Gemu Gofa (Ethiopia)". *Journal of Afroasiatic Languages* 3(3), pp.181-225.
- Hayward, Richard, J. 1989. "Comparative notes on the language of the S'aamakko". *Journal of Afroasiatic Linguistics* 1, pp.1-53.
- Mettouchi, Amina, Vanhove, Martine & Caubet Dominique. 2015. *Corpus-based Studies of Lesser-described Languages: The CorpAfroAs Corpus of Spoken AfroAsiatic Languages*. Amsterdam: Benjamins.
- Savà, Graziano & Tosco Mauro. 2000. "A sketch of Ongota, a dying language of Southwest Ethiopia". *Studies in African Linguistics* 29(2), pp.59-135.
- Savà, Graziano. 2005. *A Grammar of Ts'amakko*. Cologne: Köppe.
- Savà, Graziano. 2017. "An Ongota ethnotext on beehive". *Ethnorêma* 13 <https://www.ethnorema.it/wp-content/uploads/13-09-In-altre-lingue.pdf>
[Site consulted on 25/06/2021].
- Savà, Graziano. 2020. "Case clitics in Ts'amakko (East Cushitic, Ethiopia)". *Essays on African languages and linguistics in honour of Maarten Mous*. Leiden: African Studies Centre, pp.159-173.

Wolof language and literature: an introduction

EMILIANO MINERBA

Università degli studi di Napoli “L’Orientale”

ABSTRACT

Wolof language is knowing a period of rapid increment of its status and prestige, which is making it, from being one of the several local languages of Senegal, the second vehicular language of the country alongside with French. To this increasing status, also a rapid evolution of Wolof literature is accompanied: this particular development has started before the independency of Senegal, with the activity of Muridiyya brotherhood at the beginning of the 20th. This paper aims at being an introduction to the Wolof language from these two aspects: the socio-linguistic one, where the extents and limitations of the employment of Wolof in Senegal will be outlined, and the literary one, when the rapid developments in Wolof literature – and particularly poetry – from the beginning of the last century will be presented.

KEYWORDS

Wolof; language; poetry; prosody; wolofal.

Wolof is a West-African language spoken by more than five million people, the large majority of whom live in Senegal. It is increasingly spreading in Gambia and Mauritania, where the number of speakers is currently about 18,000¹. A member of the Atlantic subgroup of the Niger-Congo language family, the largest in Africa, Wolof is closely related to Pulaar and Sereer; besides these two languages, it has long lived together with other linguistic realities historically present in the Malian and Senegambian areas.

HISTORICAL BACKGROUND

The history of Wolof, currently the most widely spoken African language in Senegal (but not the only one: other languages, such as Sereer, Pulaar, Mandingo, Joola and Sóninke are found as well), can be read on the one hand in terms of its coexistence with the other languages of the region, and on the other in its relationship with Islam. The first attestations of this language date back to the empire of Jolof, born in the 14th century. Among the five kingdoms that composed the empire, four (Jolof, Waalo, Baol and Cayor) were inhabited by Wolof speakers: this division is partially reproduced in the Wolof dialects (Baol, Cayor, Jolof, but also Léebeu and Jandeer). In the Jolof empire, Islam was practised, but exclusively by the *buur* (emperor) and the aristocratic class; the great majority of the population was animist (Da Costa e Silva, 1996: 612).

Islam was already long established in other parts of West Africa, despite being limited to the nobility in the states that had developed in this area (Gentili, 2012: 76). A change of perspective took place only between the 18th and the 19th century, when new movements of religious renewal gained strength in all the region, very often thanks to the action of Sufi brotherhoods such as the Qādiriyya and Tiġāniyya, which, though originating from outside West Africa, flourished rapidly in this zone. In the middle of the 19th century, many different languages were spoken in the territory of present-day Senegal. The Sufi brotherhoods and the movements they led brought new religious, social and political ideas, promoting not only their religious language – Arabic – but also the languages spoken by their members.

The rise of Wolof is due in particular to the influence of a third brotherhood, the Murīdiyya, born at the end of the 19th century thanks to the work of Ahmadu Bàmba, the member of a great family of marabouts. Bàmba is still remembered today in Senegal as one of the greatest historical personalities of the country: he was a very popular religious leader, and succeeded in ex-

¹ Ethnologue: <https://www.ethnologue.com/language/wol> (last visited on 20/06/2021)

panding his brotherhood to include the majority of the Senegalese people. A peculiarity of the Murīdiyya was their use of Wolof, instead of Arabic, in communication: although Bamba himself wrote all his literary works in Arabic, he encouraged his first disciples to present the precepts of their brotherhood in Wolof poems (Ngom, 2015: 342). The prestige of the Wolof language thus increased progressively, together with the popularity of the Murīdiyya. Nowadays Wolof is often used instead of French as a vehicular language between non-Wolof people, and, though not recognised as an official language, it has definitely gained pre-eminence over the other African languages of Senegal, especially in cities such as Dakar, where it is the *lingua franca* for Senegalese immigrants from all the country (Robert, 2011: 2).

WOLOF SOCIO-LINGUISTIC STATUS

The case of the Wolof language in Senegal is a paradigmatic example of how the status of the superimposed European language undergoes fundamental changes with the end of the colonial period. Under the French rule, the French language was imposed by coercion, functional in the assimilationist perspective of the colonisers, and its diffusion was deliberately in contrast to the local languages, whose existence was considered as an obstacle to remove. In nowadays Senegal, on the contrary, the status of French is rather tied to its instrumental and professional functions, which derive mostly from its recognition as the only official language of the country; its acquisition is no longer imposed, but is an option that Senegalese young people can choose depending on their life aspiration (Schiavone, 2007: 6-7). As aforementioned, this redimensioning of the status of French has been accompanied, since the independence of Senegal, by the rise of the prestige of Wolof. The relation between Wolof, French and the other languages of Senegal is hardly describable with the classic diglossia model. A first reason of this is the lack of genetic closeness between French and Wolof. Ferguson's model of diglossia concerns two linguistic varieties that the speakers of a certain community could think of as a unique idiom (Calvet, 1994: 93), which can be the case of Classic and Vernacular Arabic, but certainly not of French and Wolof. Moreover, the perceived "superiority" of French, which should be the H language in this hypothetical diglossic model, is not a fact in Senegal: in fact, French is still established as "*la langue d'intégration aux structures de l'État*", Wolof is undoubtedly the first language for "*l'intégration en ville*" (Calvet, 1994: 101). From this state of affairs it derives that the employment of either Wolof or French as vehicular languages for the many non-Wolof Senegalese citizens is not a question of "fight for existence" between the two idioms, but of pacific co-existence of both of them in the linguistic mosaic of Senegal. To explain this coexistence Calvet (1994: 101) presents the example of a man

from Ziguinchor, the main town of Casamance (a region of Senegal where the dominant languages are Diola and Manding), which moves to Dakar to work in public administration: in all likelihood, this man will employ Diola at home, French at his office and Wolof in its errands in the city.

The vehicular role taken by Wolof in modern Senegal, alongside the equivalent function of French, seems also to be recognised by the young people of the country. A survey carried out in 1996 in Saint-Louis, a city where the Wolof people are the majority – 74,7% of the total population in 1988 – remarked that the informants recognised two varieties of both French and Wolof: *français correct* and *français débrouillé* (in the sense of “simplified French”) on the one side, and *wolof* and *vrai wolof* on the other (Auzanneau, 2006: 4). ‘Correct French’ is the standard variety as regulated from the *Académie Française*, which can be learnt only by formal teaching and is exclusively employed in formal and official situations, whereas ‘simplified French’ is a variety of the same language, influenced by Wolof, which is employed by non educated people with the only aim of achieving a successful communication (Auzanneau, 2006: 5). In the same manner, ‘true Wolof’ is thought of as the “pure” variety, not influenced by the European languages, which is, in the opinion of the informants of Auzanneau’s survey, still spoken in the Wolof countryside; despite its prestige, however, the variety employed in the urban space does not reject French and English loanwords (Auzanneau, 2006: 7-8). Urban Wolof is therefore employed not to express a cultural belonging, but to allow communication with other people in the same space, and shares therefore with *français débrouillé* a vehicular function. It is remarkable that, while Auzanneau’s informants felt the need to distinguish two varieties for both French and Wolof, they didn’t do the same for the other languages spoken in the zone (primarily Pulaar), which were grouped together in the broad category of *langues locales* (Auzanneau, 2006: 4).

WOLOF LITERATURE: A GENERAL LANDSCAPE

Wolof literature reflects the huge variety of historical contexts related to this language. It originated as oral literature: even today, there are songs that are traditionally performed on specific occasions. The relation between music and ritual is seen in non-Islamic ceremonies (for example, invocations for rain and rites of passage), as well as in Islamic contexts (Islamic new year, recitation of *ḍikr* in the Sufi communities). In addition, certain chants are heard on non-traditional occasions, such as songs in praise of the *lâmb* fighters (*lâmb* is a traditional martial art in the Wolof world, practised today as a sport), and also children’s songs (Penna-Diaw, 2016: 37).

The traditional literature includes prose genres, above all tales, which nowadays are often published in collections (an example is Diouf, 2009a). Theatrical art also has its place in Wolof literature, for example in the tradition of *simb*, “the play of the false lions”, where one or more actors dance while acting as lions, and using particular dresses and make-up².

Other genres recently introduced in Wolof literature (but not in Senegalese literature as a whole, since the literary production in French in this country has been really vast) are the novel, Western theatrical genres, and free-verse poetry.

The history of Wolof poetry, which constitutes the main topic of this paper, is interesting and worth being investigated, especially in respect of its relation with cultural change. Wolof poetry has known two main stages: in the earlier one production was exclusively oral, and in the later one written poetry appeared, heavily influenced by Arabic models. It would be exaggerated, however, to see in this second phase a complete rupture with the preceding one, as I aim to show in this brief historical sketch.

WOLOF POETRY AND CULTURAL CHANGES

Within the landscape of Wolof literature, poetry offers a privileged perspective to understand how the Wolof perceive their own cultural identity along the different influences they have been exposed to in their history. In fact, despite the high prestige attributed to Arabic, and later French, literary models, Wolof poetry is characterised by an uninterrupted continuity, from its archaic stages to modern productions.

In its earliest form, Wolof poetry consisted mainly of what is still today called *woy* (‘chant’, ‘song’). The term contains a strong reference to music: *woy* is an artistic form which embraces the two Western categories of music and poetry in a single dimension. The Wolof language, in fact, has no specific terms corresponding exactly to our “poetry” and “music”: these two concepts could be differentiated only by the use of Arabic loanwords (Penna-Diaw 2016: 37). The musical dimension is perceivable, expectedly, in the performance of the *woy*, which, on the other hand, presents a textual artistry based on very rich and articulated, yet unmetrical, versification.

The *woy* which survived in the Wolof tradition were mainly related to a particular context of execution. This is also the main criterion for categorising

² Youtube: <https://www.youtube.com/watch?v=vNJjTXjb4gI> (last visited on 20/06/2021)

this poetic form in genres and subgenres. For example, Cisse (2006: 110-113) identifies the following genres:

- *woyi taajabóon*, religious songs executed during the *tamxarit* (the Wolof name for the first Islamic month of the year, *muḥarram*);
- *woyi gàmму*, religious songs executed during the *gàmму* (the Wolof name for the third Islamic month of the year, *al-rabiʿ al-awwal*; *gàmму* is also the name of the festivity of *mawlid*, celebrated in the same month);
- *woyi ndëpp*, exorcism songs;
- *woyi jat*, apotropaic songs, employed to avert bad luck in various circumstances;
- *woyi njong*, circumcision songs. These are divided into:
 - *ngomaar*, songs performed before the circumcision of young men in order to encourage them;
 - *kasag*, songs performed during the circumcision rituals;
 - *woyi njam*, tattooing songs, performed in a tattooing ritual which signals the entrance of a young woman into her adult life;
 - *woyi céet*, marriage songs, divided into:
 - *tàgg*, praise-poetry centred on the genealogy of the praised person, performed during marriages and other familial festivities;
 - *yab*, songs for the bride when she leaves her parents' home to join her husband;
 - *laabaan*, songs praising the bride's virginity, lost after her first night of marriage;
 - *xaxar*, songs performed during the ceremony in the groom's home to welcome the new bride;
 - *woyi làmb*, performed during *làmb* fights – the traditional Wolof martial art. These can be divided into:
 - *bàkk*: praise poem for a fighter by his *géwel* (griot);
 - *bàkku*: self-praise poem that a fighter performs for himself;
 - *woyi pecc*, performed during communal meetings. These can be divided into:
 - *taasu*, songs performed individually by women;
 - *band*, songs performed individually by men;
 - *gumbe* and *ndawrabbin*, songs for big festivities and events;
 - *woyi baawnaan*, rain songs;
 - *woy mbey*, agricultural songs, used to create rhythm and coordinate collective work in the fields. These can be divided into:
 - *kañ*, songs for exhorting people to work;

- *kañu*, songs for praising one’s own good agricultural skills, compared to the other workers;
- *taaxuraan*, songs for celebrating the harvest;
- *woyi boroom gaal*, songs performed by pirogue drivers to compete with each other or to praise the water beneath them;
- *woyi yalwaan*, songs performed by the *ṭālibs* (the children living in the Islamic schools, *daara*) to ask for economic aid;
- *woyi nax/beetal*, lullabies;
- *woyi mbëggeel*, love songs;
- *woyi jaloore*, epic songs;
- *woyi seede*, songs of moral exhortation.

It is clear from Cisse’s list of genres that there is no stylistic criterion in this categorisation. No genre is distinguished for having a particular length, or special poetic structure, or particular features. In fact, despite this categorisation, all the genres of *woy* seem to share the same style. The texts can be properly described as poems, since they are composed in verses; such verses are not defined by a constant metre, but by one or more stylistic devices, which the composer can select and combine *ad hoc*. As a first example, we can quote a *woyu lām̄b*, where the performer, a fighter, asks his griot to praise him (Cisse, 2006: 49):

- | | |
|----------------------------------------|----------------------------------------|
| 1. Lees bàkk ma! | 11. Lees, praise me! |
| 2. Lees ngala bàkk ma! | 12. Lees, please praise me! |
| 3. Yaa ma daan bàkk | 13. You were the one who praised me |
| 4. Géwél yu ndaw yi fi nekk | 14. The lesser griots are here |
| 5. Ak géwél yu mag yi fi nekk | 15. And the greater griots are here |
| 6. Yaa di seen njiit | 16. You are their leader |
| 7. Boo jiiitoo ñu topp ciy mbalañ-faññ | 17. If you lead them, they will follow |
| 8. Mbalañ-faññ, mbalañ-faññ | you dancing |
| 9. Lees ngala bàkk ma! | 18. Dancing, dancing |
| 10. Yaa ma wara bàkk | 19. Lees, please praise me! |
| | 20. You are the one who must praise me |

The reader will notice that the verses of this composition (which are marked by lineation in the transcription and by pauses during the oral performance) are not subjected to a constant metrical pattern; nevertheless, many linguistic devices are employed to mark linguistically the verse boundaries, so that this versification cannot be defined as a free-verse form. Verses 1 to 3 present a polyptoton of the verb *bàkk* (“praise”): that is, the same verb occurs once in each line, but in different flexion forms. Another word which is repeated once in each of these verses is *ma* (“me”, 1st singular person object pronoun). Verses

1 and 2 are bound together by the anaphora of *lees* and the homeoteleuton in *bàkk ma* (“praise me”). Verses 4 and 5 are bound by a syntactic parallelism which becomes an actual repetition: verse 5 is identical to verse 4, with the only two exceptions of the occurrence of the introductory conjunction *ak* (“and”) and the substitution of *mag* (“great”) with *ndaw* (“little”). The ending of these two verses *nekk* ([ne:k:ə]) binds them in a consonantal rhyme with the word *bàkk* ([ba:k:ə]) which structures the preceding lines. Verses 6 to 8 are bound together by two repetitions. The first repetition, involving verses 6 and 7, concerns the root *jiit* which appears first in the noun *njiit* (‘leader’, verse 6) and then in the verb *njiitu* (‘lead’; occurring as *njiitoo* in verse 7). The second repetition concerns the verb *mbalañ-faññ* (‘turning around dancing’), which occupies the end of verse 7 and all of verse 8. This group of three verses is connected to verses 4-5 by the reference to the figures of the griots (*gewel*), which is the main theme of verses 4 to 8. At the end of the poem, a couplet of verses goes back to the first linguistic device employed for versification in this composition (the polyptoton of *bàkk*) and uses it to close the text.

In this example it is evident that the versification, despite not being metrical, is not free. Free verse does not coincide with merely unmetrical verse: it is a form of versification where the verse boundaries are disposed in a way that cannot be expected in view of the linguistic organisation of the text (Gasparov, 1996: 285; Hartman, 1980: 61-80). This is not the case with the Wolof *woy*: here, the verse boundaries are always introduced by one or more linguistic devices, even if not in a regular and constant way, as it happens in metrical verse. To this particular way of versification, which is neither metrical nor free, I give the name of *distinct verse*. The adjective ‘distinct’ here refers to the fact that in this form of versification the verse-boundary is always marked by one or more linguistic devices: the verses, therefore, are somehow always distinct from each other. Distinct verse is not an exclusive feature of Wolof poetry; on the contrary, it is a versification form that can be observed in many different literary traditions, both inside and outside Africa. In the context of West Africa, distinct verse is part of the stylistic tradition developed across the centuries by the griots.

Common to several cultures of West Africa, the Wolof *gewel*, the Pulaar *gawlo*, the Mandinka, Fali and Bambara *jali* and other similar figures are represented by the French word *griot*, meaning the bards and performers of those societies which were part of the ancient Empire of Mali (Leymarie 1999: 5). As in other societies, the *gewels* occupy a special social class among the Wolof, and interact with the other members of the population in clearly defined ways. The *gewels* were a group of the *ñeeño*, that is, that set of classes of artisans compelled to practise endogamy. They were not, however, a pres-

tigious group among the *ñeeño*, since they were put at the lowest level of the artisanal classes, together with the *ràbbs* (weavers). The *gewels* themselves were divided into several subgroups with different functions (Leymarie 1999: 26-28):

Les griots étaient divisés en bardes bouffons (bawleck [bawlekk]), qui chantaient et récitait les généalogies; généalogistes et marchands (jëfleck [jëflekk]), parmi lesquels les gawlo et les géwél; en musiciens, plus prestigieux, distingués selon leur style musical; et en sableck [sablekk], aux fonctions mal définies.

Despite their inferior status, the *gewels* played an important role in traditional Wolof society: not only did they help to preserve the Wolof cultural patrimony, in particular genealogies, but they were also messengers and counsellors of the sovereign. Moreover, the griots had a monopoly on the practice of music, which was discouraged also for the nobles (*géer*; see Leymarie 1999: 32).

It is worth quoting a further excerpt from Leymarie's (1999: 122-123) work:

L'acquisition du répertoire est longue et débute à un très jeune âge. Une vieille griote lébou se touchait la poitrine pour indiquer qu'elle avait commencé à absorber la "tradition" au sein de sa mère. Avant d'entreprendre un véritable apprentissage, l'enfant géwél bénéficie d'une immersion constante dans la musique, les chants et les récits oraux. Il assimile par mimétisme, en observant les membres de la famille. Les connaissances sont fréquemment transmises par des personnes du même sexe que lui : le père enseigne à ses fils, la mère à ses filles. [...] Traditionnellement, il était impensable, pour un enfant gewel, de ne pas se consacrer à l'étude des chroniques historiques, des chants de louanges, d'un instrument de musique ou de généalogies.

From the situation described by Leymarie it appears plausible that Wolof *woy* and its parallel poetic forms in the rest of West Africa developed as stylistic traditions that were passed down within the *griot* classes of each population. The chain of inheritance was between father and son, or mother and daughter; if one considers that, in most cases, the *gewel* families were specialised in just one activity (such as the performance of genealogies or the mastering of a particular drum), one could consider that each of the skills that composed the Wolof verbal art of *woy* was connected to its past by a vigorous continuity. This continuity also connects the *woy* with the protagonist of the second period of Wolof literature, namely the Murīd *wolofal*. The major innovation in this genre is represented stylistically by the introduction of metres, in particular the Classical Arabic metres. As mentioned above, the protagonists of this literary change were the poets of the rising Murīdiyya brotherhood.

The Murīd poets, under the encouragement of Ahmadou Bamba Mbakke, the founder of this new brotherhood, gave birth to this new poetic genre, whose

name has the literal meaning of “to make Wolof”. Scholars have proposed two definitions of *wolofal*, focusing on different aspects. One is that shared by Fallou Ngom (2015: 335), according to whom *wolofal* is a manuscript tradition, not limited to literary works, of Wolof texts written in ‘*ajamī* (that is, an *ad hoc* adaptation of the Arabic script). Other views, more centred on the stylistic aspects of Murīd production, such as that of Lô (1993: 2), define *wolofal* as a genre of Wolof poetry which employs Arabic metres for poems composed in Wolof. The latter definition seems to be more fitting for modern times. More than fifty years after the death of Muusaa Ka in 1967, the youngest member of the first generation of Murīd poets, *wolofal* poems are more and more frequently being published using Latin script, or more precisely the Wolof standard orthography. The genre has lost its religious specificity, since the poems of the 20th century also focus on secular topics (for example, Jóob’s *Njool-Céytu* is an elegy for the famous Senegalese anthropologist Cheykh Anta Diop). The only feature that has been retained, despite many important developments, is the employment of Arabic metric forms.

Despite the recent secularisation of the *wolofal* genre, however, it is undoubted that this formation is due to the Murīds’ intellectual activity. Ngom (2016: 3) argues that the use of *wolofal* as a mass-communication strategy has been key to the success of Murīdiyya. As mentioned above, Bamba himself encouraged his disciples, *in primis* Moor Kayre (Mor Kairé, 1869-1951), Samba Jaara Mbaay (Samba Diarra Mbaye, 1870-1917), Mbaay Jaxate (Mbaye Diakhaté, 1875-1954), and Muusaa Ka (Moussa Ka, 1889-1967), to compose and diffuse *wolofal* poems (Ngom, 2015: 342). It must be noted that the Wolof were not the first in Western Sudan to engage in the reception of Arabic prosody and script. Old Tashelhit or Medieval Berber, Songhay and Kanuri are believed to be the first West African languages to have been written in ‘*ajamī* between the 10th and the 16th centuries, followed by Fulfulde, Hausa, Wolof, and Yoruba (Ngom, 2016: 8). From the 18th century in Futa Jallon, Pular literature developed forms of religious poetry in which Arabic metres were employed (see Barry, 2004; Seydou, 2001; Sow, 1965). Fuuta Tooro, by contrast, seems to have had a lower production of manuscripts, at least until the Islamic revolution of al-Hajji Umar (Humery, 2014: 173-174); this could explain the “delay” between Wolof culture, which was in strict contact with it, and other cultures in developing the ‘*ajamī* scripts. However, given the extent of the familial and intellectual networks among marabouts in Western Africa, it is unlikely that the Wolof *sērīñ* and the Pulaar *cerno* did not have any knowledge of these traditions.

As mentioned above, the first main Murīd *wolofal* poets were four in number: Mor Kairé, Samba Diarra Mbaye, Mbaye Diakhaté and Moussa Ka. Of these, the first three were Tiḡānī scholars who had joined Murīdiyya, and started composing *wolofal*, during Bamba’s exile in Mauritania. Moor Kayre, from Xombol in Kajoor, focused on hagiographies and topographic Islamic

knowledge (how to locate the right place and the topographic requirements for building a mosque). His relative, Samba Jaara Mbay, hailed from the Islamic centre of Kokki. He later relocated to the city of Saint-Louis where he spent the rest of his life. He was an Arabic-trained scholar who also engaged in commercial activities in Saint-Louis, and was versed in Islamic sciences, especially metaphysics, when he encountered Bamba in the early 1900s. Thus, his poems regularly dealt with metaphysics, including the transhistorical voyage of the souls of the Murīds that began from the Qurʾānic narrative of the primordial Day of the Covenant (the Day of Alastu). Mbay Jaxate was from Kër Makala, which was founded by his distinguished Tiġānī father, Qāḍī Majaxate Kala (Bamba’s teacher of metrics). Mbay Jaxate lived for the remainder of his life in the village of Xuru Mbakke in the vicinity of Diourbel. He lived there farming, teaching, and composing ‘*ajamī*’ poems. Jaxate focused on internal social criticism, satire, and the code of ethics of Murīd discipleship (Ngom, 2016: 22).

Muusaa Ka, the youngest of the four Murīd ‘*ajamī*’ pioneers, is from the village of Ndilliki-Kanka in Bawol, which corresponds roughly to the present-day region of Diourbel. His father, Usman Ka, who was a Qāḍirī Qurʾānic teacher, taught him the Qurʾān, Islamic law, and Ṣūfism. He also studied with his father’s half-brother, Shaykh Ka. Muusaa Ka was a relative of Bamba to whom his father later entrusted him. He lived in Ceyeen-Jolof with Bamba during the first house arrest of his master (1907– 1912), performed peripatetic learning in Senegambia, and lived with Bamba again in Diourbel during the second house arrest (1912– 1927). Muusaa Ka drew on Bamba’s extensive writings, multiple Islamic sources, and the works of his three senior colleagues, and he became an unmatched Senegalese ‘*ajamī*’ hagiographer, historian, and poet. His productivity and eclecticism accelerated and broadly popularised Murīd ‘*ajamī*’ poetry and the odyssey and teachings of Bamba among the Senegalese masses. His work spans religious, hagiographic, historical, and secular domains. Muusaa Ka captures best the voices and the collective memory of Murīds and offers the most radical alternative narrative to the prevailing discourse on the Muridiyya (Ngom, 2016: 23).

Far from being an attempt to carry out an “Arabisation” of Wolof literature, the Murīd tried to build a genuinely Wolof Islamic literary tradition. That is why, leaving aside the script and the metres employed, *wolofal* places itself in direct continuity with the preceding tradition of *woy*. This is evident, for example, in Muusaa Ka’s masterpiece, *Xarnu bi* (“The century”), an elegy for Ahmadu Bamba’s death, whose death is compared with the other catastrophes and uncertainties of the century. Metrically, this poem is a *tasmīṭ murabbaʿ*, that is, a composition of quatrains following the rhyme pattern *aaax*, *bbbx*, *cccx*, ..., *zzzx*. Besides sharing the same rhyme all along the poem, the last

lines of all the quatrains also present a common refrain, since they all end with the expression *xarnu bi* ('the century'). The metre of this *tasmīt* is *rağaz mağzū*⁷. Like the other Arabic metres, *rağaz mağzū*⁷ is quantitative, thus based on a specific succession of short and long syllables. Its metric pattern can be represented as $\underline{\text{u}} \text{ } \underline{\text{u}} \text{ } \text{—} \text{ } | \text{ } \underline{\text{u}} \text{ } \underline{\text{u}} \text{ } \text{—}$. One could quote, as an example, the first stanza of the poem (Kamara, 2008: 121):

Sëriñ bi noo ngi deeti ñaan
Fàqir dafay nangoo dagaan,
Nangul nu lepp lu nu ñaan,
Ndax Yàlla naatal xarnu bi.

Master, here we are again begging,
 since it is licit for a poor man to beg,
 Grant us everything we ask you
 so that God makes wealth the century.

In this case we have a plain application of the Arabic metric pattern of the *rağaz mağzū*⁷ to the Wolof language. There are some particularities regarding syllabification, due to the different prosodies of the two languages. For example, the apparently monosyllabic word *lepp* [leppə] must be counted as bisyllabic, *lep-pə*, the first syllable being long and the second short. This appears to be due to the Murids' effort to adapt the Arabic metres; for the rest, the application of the metre is no particular problem. Other stanzas, however, reveal that Muusaa Ka employs other stylistic devices taken from the large patrimony of the *woy*, such as syntactic parallelism and repetition (Kamara, 2008: 123):

Xiif tax na ñenn mag ñi yooy,
Xiif tax na gor su ndaw di jooy,
Xiif tax na yenn tool yi booy,
Sëriñ bi geesul xarnu bi.

Starvation has made some of the oldest men dull
 Starvation has made the smallest of men weep
 Starvation has made some fields barren
 Master, look back on the century.

One can also find examples of polyptoton, so common in the *woy*. In the following stanzas the polyptoton of the word *baax* ("to be good") occupies seven verses, and overlaps with a similar use of the word *waa* ("people"), which involves four lines (Kamara, 2008: 131):

Lu baax du baax ba jàll baax
Waaye Sëriñ bee weesu baax,
Baax a ngi yem fi xotti baax,
Baaxaale gaayi xarnu bi.

What is good is not good beyond what is good
 But the Master has gone beyond the good
 There is a limit to good, but he has broken it.
 Good were with him the companions of the century.
 The people who loved you are now good
 The people who opposed you are no longer good
 The people who thank you are the most good
 Oh people! To him we owe the century.

Waa ju la sopp dal di baax,
Waa ju la weddi dootu baax,
Waa joo gërëm muy gën di baax,
Waayoo ko war na xarnu bi.

This bond with tradition can be explained by the explicit aim of the Murīdiyya to give Wolof a recognised Islamic literature, which would be part of the Islamic network without renouncing its particularities. Stylistic analysis offers evidence of this bond between *wolofal* and the literary tradition it is born within.

CONCLUSIVE REMARKS: A GLANCE AT MODERNITY

Although *wolofal* poetry was written even in the earliest phases of its development, it has never lost its connection with orality. These compositions are intended to be performed in the presence of an audience rather than being read, and their execution needs a high level of mastery³. Thus, the continuity with the preceding tradition is striking. Nowadays, the composers of *wolofal* differ more and more from the first Murīd composers: many of them are not religious poets, and there are authors, such as the composer Ousmane Loo, who aim at renewing *wolofal* in order to make it even closer to the original Wolof poetic models.

Looking at today's poetic landscape, the main difference between *wolofal* and traditional songs seems to lie not in the medium of transmission but in the creative possibilities of the poet. The *wolofal* composer Ousmane Loo, in an interview I conducted with him in Dakar on 17/09/2019, explained that whereas traditional songs have their *raison d'être* in their content and the occasion of their performance, in *wolofal* the poet can freely choose the content, since everything can be the topic of a poem if it is expressed respecting the metrical forms. It is interesting that this is the point of view expressed by Loo, who is an extremely innovative poet, not only in his thoughts (his poetry is not at all religious) but also on the stylistic side (as already mentioned, he introduces many variations in the Arabic metres employed in *wolofal*). Nowadays, *wolofal* is undoubtedly one of the most interesting literary phenomena in Senegal.

³ A good example of performance is available here: <https://www.youtube.com/watch?v=dz3NJvAc5oY&t=57s>. It is the poem *Xarnu bi* ("The Century") by Moussa Ka, one of the most famous poems of the genre.

REFERENCES

- Auzanneau, Michelle. 2006. "Quelques indices pour l'étude de la dynamique sociolinguistique de Saint-Louis", in: Akura Sarr, Edris Makward, Amadou Tidiane Fofana, C. Frederick (eds.), *The histories, Languages, and cultures of west Africa: Interdisciplinary essays*, The Edwin Mellen Press, pp. 58-69. Available at the Url: https://hal.archives-ouvertes.fr/hal-00927295/file/Quelques_indices_pour_la_etude_de_la_dynamique_sociolinguistique_de_Saint-Louis.pdf
- Barry, Alpha Ousmane. 2004. "Mode d'expression poétique et stratification sociale dans l'état théocratique du Fouta Djallon", in: *Semen. Revue de sémio-linguistique des textes et discours*, 18. Available at the Url: <https://journals.openedition.org/semen/2294>.
- Cisse, Momar. 2006. *Parole chantée ou psalmodiée wolof. Collecte, typologie et analyse des procédés argumentatifs de connivence associés aux fonctions discursives de satire et d'éloge*, Dakar: PhD Thesis discussed at the University Cheikh Anta Diop.
- Da Costa e Silva, Alberto. 1996. *A Enxada e a Lança. A África antes dos Portugueses*, Rio de Janeiro: Editora Nova Fronteira.
- Diouf, Jean-Léopold. 2009a. *Beykat bi ak Jinne ji. Le cultivateur et le djinn*, Paris: L'Harmattan.
- Gasparov, Mikhail Leonovic. 1996 [1984]. *A History of European Versification*, Oxford: Clarendon Press.
- Gentili, Anna Maria. 2012. *Il leone e il cacciatore. Storia dell'Africa sub-sahariana*, Roma: Carocci.
- Hartman, Charles O. 1980. *Free Verse. An essay on Prosody*, Princeton: Princeton University Press.
- Humery, Marie-Ève. 2014. "Fula and the ajami writing system in the Haalpulaar of Fuuta Tooro (Senegal and Mauritania): a specific 'restricted literacy'", in: Mumin, Meikal & Versteegh, Kees (eds.), *The Arabic Script in Africa. Studies in the Use of a Writing System*, 173-198. Leiden: Brill.
- Jóob, Libaas. 2016. *Njool-Céytu*, Dakar: Éditions Papyrus Afrique.
- Kamara, Sana, Sèriñ Muusaa Ka. *Melokaani Roytéef*, Éditions Papyrus Afrique, Dakar, 2008.
- Leymarie, Isabelle. 1999. *Les griots wolof du Sénégal*, Paris: Servedit.
- Lô, Mamadou. 1993. *Un aspect de la poésie "Wolofal" mouride. Traduction et analyse de quelques titres de Sèriñ Mbay Jaxate*, Dakar: Master Degree dissertation discussed at the University Cheikh Anta Diop.

Ngom, Fallou. 2015. "Murid Ajami sources of knowledge: the myth and the reality", in: Kominko, Maja (ed.), *From Dust to Digital. Ten Years of the Endangered Archives Programme*, 331-376. Cambridge, Open Book Publishers.

Ngom, Fallou. 2016. *Muslims Beyond the Arab World. The Odyssey beyond the Arab world. The Odyssey of 'Ajami and the Muridiyya*, Oxford: Oxford University Press.

Penna-Diaw, Luciana. 2016. *La Musique des Wolof au Sénégal. (Régions du Kajoor, Saalum et Waalo)*, Paris: Peeters.

Robert, Stephane. 2011. "Le wolof", in: Emilio Bonvini, Joëlle Busuttill & Alain Peyraube (eds.), *Dictionnaire de langues*, 23-30. Parigi: Quadriga/P.U.F.

Seydou, Christiane. 2001. "Poésie religieuse et inspiration populaire chez les Peuls du Fouta Djallon", in: *Journal of African Cultural Studies*, 14(1), pp. 23-47.

Sow, Alpha Ibrahima. 1965. "Notes sur les procédés poétiques dans la littérature des Peuls du Fouta-Djallon", in: *Cahiers d'Études africaines* 19, pp. 370-387.

The future of minority languages in Nigeria

GIAN CLAUDIO BATIC

Università degli studi di Napoli “L’Orientale”

ABSTRACT

Nigeria exhibits an extraordinary linguistic diversity, both in terms of genetic affiliation and sociolinguistic status. A large proportion of the 520 (and counting) Nigerian languages are spoken by minority groups. In most cases, these groups are subject to a process of linguistic and ethnic conversion that will lead to a reduction in linguistic diversity and the consolidation of two main vehicular languages: Hausa and Nigerian Pidgin. This paper will discuss the notion of minority language and the idea of language endangerment, and consequently the factors that seem quintessential in determining the sociolinguistic framework of tomorrow’s Nigeria.

KEYWORDS

Nigeria; minority languages; identity; endangerment.

1. INTRODUCTION

The region that extends within the borders of what is now Nigeria is an extraordinary linguistic microcosm. Of the four linguistic phyla on the African continent, three are present in Nigeria (Afro-Asiatic, Niger-Congo, and Nilo-Saharan). With some 520 living languages, Nigeria alone holds 24% of the linguistic heritage of the African continent. This particular situation is the consequence of a turbulent history of contacts and migrations that began thousands of years ago, when the Chadic groups (descendants of the Proto-Chadic group, a population of Afro-Asiatic origin settled around 6,000 - 5,000 BCE on the northern shores of what was then Lake Megachad) came into contact with populations speaking Niger-Congo languages. The theatre of contact between the Chadic-speaking groups and the Niger-Congo speaking groups was north-central Nigeria. One branch of the Chadic family occupied the region west of Lake Megachad (3,500 BCE), which eventually led to a series of migrations of the Niger-Congo groups southwards, thus drawing a distribution of the two main phyla that would be maintained in the following millennia.

Most of the languages spoken in Nigeria are so-called minority languages. In the last few decades, the scientific community has produced an abundant literature on small mono or bilingual communities, trying to identify the relevant factors that contribute to determine the degree of fragility and the risk of extinction of these languages. In this paper, I will discuss the Nigerian case, focusing on three fundamental aspects: the notion of minority language, language endangerment, and the relationship between language and identity. Finally, I will outline a hypothesis of a not too distant future, trying to project the current trends and forces into the linguistic framework of tomorrow's Nigeria.

2. MINORITY LANGUAGES

The state of minority languages in Nigeria is often described within a binary system where the notion of 'minority' is opposed to that of 'vehicular' (or dominant) language. Hence, minority languages would be all those languages whose speakers represent a minority within a given country or region *vis-à-vis* the number of speakers of another language in the same country or region. Such a view would virtually group together all those languages with a number of speakers inferior to that of, let's say, Hausa, Igbo and Yoruba. Moreover, the notion of minority language is often accompanied by the idea that being a minority entails a certain degree of linguistic endangerment, i.e. minority languages are on the path towards extinction due to the pressure exerted by vehicular and other dominant (e.g. national) languages. This picture

is completed by the fact that the languages spoken today in Nigeria are more than 500, an astonishingly high number if we consider that the total number of the world's languages is (near to) 7,139 (Eberhard, Simons and Fennig 2021). The combination of these three notions – the opposition minority-dominant, the relative small number of vehicular and national languages, and the state of endangerment inherent to minority languages – might lead us to think that the majority of Nigerian minority languages are in a state of extreme fragility, which will eventually result in a massive loss of linguistic diversity. To this we could also add the concern as well as the rhetoric of a certain activism within and outside the academia that has nourished a sense of urgency towards the risks of language endangerment.

The notion of 'minority language' is misleading. As we have seen, all languages are doomed to be classified as 'minority' against vehicular or national languages. A binary division, I will argue, is over-simplistic: it does not capture the state of things. First of all, languages can be small and yet have a status of *lingua franca*, as in the case of inter-village communication languages. Then we have languages dominant at state or region level that nevertheless coexist with a national or vehicular language. It would be hard to group a regional language with 1,000,000 speakers together with a local (e.g. mono-village) language of 2,000 speakers, but it would be equally difficult to treat a regional language in the same way we treat a vehicular-national language such as Hausa or Yoruba. Hence the label 'minority' fails to grasp the differences between languages of different numerical consistencies and geographical extensions, neglecting the fact that in most cases being a minority language is a relative condition. Nevertheless an absolute state of minority – i.e. a situation where the language is not spoken by anyone else except for the community that uses it as a mother tongue, often in a single village – does exist and is very common, although it is just one among many. This is the kind of scenario where endangerment becomes a fully-fledged reality.

If talking of minority languages is not so descriptive of what is going on, what term should we use? Or better: how could we regroup languages of different sizes, used in different contexts and spoken at local, state or regional level? The British-German linguist Conrad Max Benedict Brann, concerned with the issue of framing the different types of Nigerian languages for educational planning, proposed a macro-sociolinguistic model (Brann 1977, 1993, 1994).

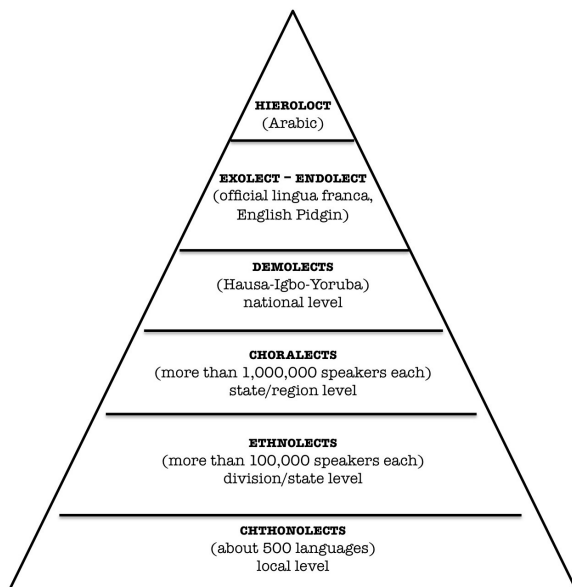


Figure 1– Macro-sociolinguistic model (Brann 1977)

Brann’s model takes the form of a pyramid (see figure 1 above) in which five main types of languages are distinguished: chthonolects, ethnolects, choralects, demolects, and exolects/endolects. Chthonolects (‘languages of the soil’, Brann uses this term to avoid ‘vernacular’) form the base of the pyramid and are spoken mostly by L1-speakers in small communities. At the second level of the pyramid we find the ethnolects, i.e. those languages spoken by ‘larger minorities’. Larger minorities are defined by Brann in numerical terms (100,000 speakers each), but we could also consider them as dominant minorities, that is groups that for historical, demographic and socio-political reasons constitute the majority and at the same time exert political control over other communities present in the area. This is the case, for example, of the Tangale in northeast Nigeria, a group of more than 200,000 people whose traditional ruler, the *Mai Kaltungo*, is also the paramount chief of the other groups of the area (southern Gombe State). At the next level Brann posits what he calls choralects, or regional languages. These languages are “spoken by a majority in any one larger administrative division, but also as a second language by minorities” (Brann 1977: 322). Apart from being used in a larger area and by a larger number of people, choralects have also L2-speakers, whereas ethnolects (with some exceptions) and chthonolects have only L1-speakers. Going up further the pyramid, we find the demolects or vehicular (or national) languages. This position is occupied by those languages whose number of speakers is in the tens of millions. They function at national level and are used in federal mass communication as well as in written primary,

secondary and tertiary education. Not surprisingly, demolects constitute a very small club and for the foreseeable future its members will still be the ‘Big Three’, i.e. Hausa, Igbo, and Yoruba. Going up one level we encounter the continuum exolect-endolect, that is English and Nigerian Pidgin English. Although both languages are widely used, they play different roles: English is the official language used in federal mass communication, administration, and education, whereas Nigerian Pidgin English is *de facto* a national *lingua franca*. Finally, at the very top of the pyramid, Brann posits the hierolect or ‘sacred language’, which in Nigeria is represented by Arabic. This language is of little importance to us, since its function is primarily one of participation and association, and not one of communication.

3. LANGUAGE ENDANGERMENT

Let us leave aside for a moment the concept of ‘minority’ to focus instead on the issue of endangerment. While there is a global trend that sees small communities abandoning their own languages, in Africa the situation seems to be different. Scholars agree on the fact that multilingualism (Vigouroux and Mufwene 2008) and urbanization (Lüpke 2015) play an important role in keeping indigenous languages alive. The assessment of language endangerment/development using the Expanded Graded Intergenerational Disruption Scale (EGIDS) (Lewis and Simons 2010) confirms this African peculiarity, with Nigeria making no exception: the absolute majority of Nigerian languages are classified as ‘vigorous’¹ (i.e. over 300 languages, Eberhard, Simons and Fennig 2021). Several scholars do not share this view, pointing out that the lack of data may distort the assessment of linguistic vitality and underestimate the real danger of extinction (cf. Essegbey 2020: 834 ff.). Nevertheless, the African case shows – at least to a certain extent – that an important number of indigenous languages do not face an immediate threat of extinction, which makes possible to argue against the automatic association between minority languages and endangerment.

Several studies, however, have shown the precarious state of African languages from the point of view of language endangerment (among others, Brenzinger 1998, Batibo 2005). While it is true, as we have seen before, that more than half of Nigerian languages are classified as ‘vigorous’, it is worth mentioning that the remaining languages (except, of course, vehicular and institutional languages) are at risk of extinction (about 125 languages). Of the

¹ “This is the level of ongoing oral use that constitutes sustainable orality. Intergenerational transmission of the language is intact and widespread in the community. The language use and transmission situation is stable or gaining strength.” (Lewis and Simons 2010: 112).

2,154 languages spoken in Africa, 520 are found in Nigeria. To give an idea of the areal situation, of the 890 West African languages listed on *Ethnologue*, 445 languages are labelled as ‘vigorous’, 124 as ‘in trouble’ and 45 as ‘dying’ (Eberhard, Simons and Fennig 2021)². Moreover, even if we were to take the data reported in *Ethnologue* as faithfully representing the reality of things, the definition of ‘vigorous’ provided by the EGIDS should not lead us to be too optimistic. On the one hand it is important to remember that things can change very quickly, especially when intergenerational transmission of the language has broken down. On the other hand, the shift from monolingualism to bilingualism observed in recent decades in many communities exposed to the pressure of vehicular languages indicates a clear trend towards the erosion (and possibly disappearance) of local languages. To these two aspects we could add the fact that, in most cases, communities do not oppose the adoption of the vehicular language as a mother tongue.

Languages can be ‘in trouble’ or ‘dying’, but what does threaten them? What is the cause of a condition of fragility that will eventually end up in language death? There are two main scenarios that determine the disappearance of a language: the first involves the occurrence of a conflict or disaster that causes the disappearance of a language community. The second scenario – which describes almost all cases on the African continent – is one in which a community abandons its own language and adopts another. As dramatic as this may seem in terms of loss of cultural heritage, the linguistic shift occurs peacefully and, in most cases, without any particular underlying tensions or actions of cultural resistance. To use Edward’s words, “it has always been natural in our sublunar realm for societies and their languages to falter, to decline and to pass from the scene. [...] the general pattern is a robust and enduring one” (Edward 2010: 14).

If we look at Brann’s categorisation, it will be fairly easy to identify the type of language most prone to erosion. The languages most at risk of extinction – and here there is no surprise – are chthonolects. Although the absolute number of speakers alone is not sufficient to assess the vitality of a language, there is no doubt that small communities of only L1 speakers are destined to disappear under the pressure of a changed demographic and cultural context. In Nigeria (and generally in West Africa), monolingualism is a rare condition. In the urban context, a typical speaker is competent in one or more vehicular languages and a choralect; in the rural context, communities tend to add a vehicular language to their chthonolect. While in the former case the languages coexist in a situation of substantial parity (subject to the distinctions inherent to official, administrative and educational uses), in the latter bilingualism is

² Compare this with the survey published by Batibo in 2005: 485 languages are attested in Nigeria, of which 73 are classified as ‘highly endangered’, 55 as ‘(nearly) extinct’ and 363 as ‘less endangered’ (a more factual label than ‘vigorous’) (Batibo 2005).

the result of the insufficiency of chthonolect to meet the communicative (and cultural) needs imposed by the areal context³.

One could object asking the following: why do these small communities not add a vehicular language to their 'local' language, assigning different functions to them but without giving up their chthonolect? This is a fundamental question, as it closely touches on the issue of identity and its relationship with language.

4. THE IDENTITY FACTOR

The issue of the relationship between identity and language is at the core of a prolific scholarly literature. In this section, I will present some aspects relevant to the Nigerian context by attempting to frame the notion of identity in relation to the language(s) spoken by a group.

Language as an identity mark

Language is often regarded as the essential feature of a cultural identity. The close association between language and identity, however, finds a limit in all those cases where the fact of speaking the same language does not act as an identity factor. Newman (1969/70), for example, describes the case of the Tera communities in northeast Nigeria: although linguistically indistinguishable, they can be subdivided into two groups identified with two distinct historical origins, i.e. two different migrations (one group is said to have arrived in the area migrating from the east, while the other originated in the north, in the Bole area). The Tera language and its adoption represent, so to speak, the linguistic output of the coexistence of the two groups, which, however, did not translate into overlapping identities. However, the case of Tera tells us that the opposite is also true: different linguistic groups can share the same identity (e.g. in terms of oral traditions, privileges, and obligations). The Bole-speaking village Kafarati recognises itself (and is recognised) socio-politically in one of the two Tera groups. Again, the linguistic distribution alone does not tell us much about culture and socio-political ties.

Another case of clear dissociation between language and culture (i.e. history and identity) is the one described by Blench (2015) about the Yangkam group (Plateau State, Central Nigeria). Blench highlights the paradoxical situation of the Yangkam where "members of the ethnic group are very proud of their history and identity, but do not associate them with retention of the language" (2015: 151).

³ The community will eventually shift again to a situation of monolingualism once the chthonolect fades away.

Language distribution is associated with migration phenomena. Specifically, the degree of kinship and the geographical distribution of the languages of the same family should provide an indication of the migrations and subdivisions that have taken place from a proto-group. Consequently, one might think that a language classification diagram is also a faithful representation of the separations (and therefore, migrations) that have taken place between language groups. The migration model underlying African societies is quite different: the system of group reproduction described by Kopytoff (1987), for example, offers a much more complex picture. In Kopytoff's model, the creation of a new group is illustrated by the following stages: 1) the separation of segments (clans, family units, individuals) from the metropolis; 2) their migration into the empty political space, i.e. the 'African frontier'; 3) the formation of a new political centre; and 4) the inclusion, over time, of other segments from other metropolises, regardless of linguistic affiliation. Each segment (clan, household, individual) that aggregates in the newly occupied frontier makes a contribution in terms of culture and identity. The result is a group that has codified all the components both linguistically and socio-politically. On the linguistic level, this product of synthesis can be seen in the lexicon and in the acquisition/transfer/deletion of linguistic traits. Consider the following scenario: the first segment to occupy a space of the frontier speaks language A; in the course of time, other segments will be added: a segment speaking language A' (genetically related to language A), and two other segments speaking language B and C (genetically unrelated). Imagine also that the resulting group speaks language A'': a language derived in its structure from A, but influenced (in terms of acquisition/transfer/deletion of linguistic traits) by A', B, and C. This scenario exemplifies the model of linguistic reproduction and strongly departs from the linearity of internal classification diagrams⁴. As for the socio-political codification, it will take place, for instance, through clan organisation and the setting of obligations, rights, privileges, following modalities determined by the prestige and consistency of the components, as well as taking into account possible situations of conflict.

⁴ This scenario exemplifies a model. I would add that to complicate the non-linear picture of the model there may be return migrations (segments decide to return to the metropolis of origin) and splits within the group (segments decide to occupy other space in the frontier). Moreover, some segments may keep their language of origin (e.g. in order not to break the link with the metropolis of origin, because of their socio-political position, etc.).

Changing and shifting identities

The discussion above has shown that identities are not socio-cultural 'blocks', but rather multi-layered realities resulting from a process where different segments (i.e. micro-groups, clans, families, or individuals) with different geopolitical backgrounds merge to form a socio-political and cultural complex. Therefore, an essentialist approach to identities in the West African context would miss the complexity of societies. An identity can be considered as the set of practices and ideas that allow a community to define itself, i.e. identity as culture as well as an ideology on culture. However, we should note that identities may not only change, but also shift or disappear.

Cultural shifts are gradual, yet they can occur relatively quickly. They involve, for example, the adoption of a new way of dressing, a change in the religious paradigm, and the transmutation of the value scale.

In northern Nigeria, for example, Hausatisation is a steady process of religious, cultural and political expansion that sees many communities giving up their cultures to adopt that of the dominant group. Hausatisation (and the Islamisation that accompanies it) has erosive effects on the transmission of musical knowledge and non-Islamic rites and beliefs, even acting in the modulation of the values codified in the oratures.

The need to identify with the dominant culture inevitably also involves adopting the language that expresses it, since the language of the group is no longer competitive in communicative and cultural terms. In most cases cultural shifts have no dramatic outcomes.

5. THE FUTURE

A prediction of the future status of minority languages – and therefore of their numerical consistency – must necessarily proceed from an accurate set of data on the current state of the languages and their diachronic analysis. Unfortunately, our knowledge of the sociolinguistic situation of the non-dominant Nigerian languages is rather limited: of some languages we know a lot, of many languages we know too little, and the assessment of the degree of penetration within a group of a vehicular language as a mother tongue is often superficial. To this we must also add the scarcity of information and data on the relationships between non-dominant languages within the same area.

The two languages that contribute most to the erosion of linguistic diversity are Hausa and Nigerian Pidgin (NP),⁵ in the north and south of the country

⁵ Nigerian Pidgin, sometimes called Nigerian Pidgin English, is classified as an Indo-European language belonging to the West African Creole English group. Some scholars use the term Anglo-Nigerian Pidgin (Mann 1993, Simire 2004). Ofolue (2010) employs the

respectively. The sociolinguistic profile of the two languages is quite different: Hausa is a demolect directly attributable to a dominant group, while NP is the output of an exolect-endolect continuum. NP is ideologically neutral, rapidly evolving, and with a very low degree of stigmatisation (in contrast, for example, to Ghanaian Pidgin, see Ofolue 2011). The peculiarity of NP is its use in conjunction with other languages that are usually qualified as vehicular, namely Igbo and Yoruba. This phenomenon highlights the limits of the two southern Nigerian demolects: although they are firmly established in large geographical areas and have tens of millions of speakers, Igbo and Yoruba lack the driving force necessary to transcend their traditional perimeter of diffusion. The consequence is that NP also impacts on demolects such as Igbo and Yoruba, which may one day be demoted to the rank of non-vehicular demolects.

Hausa, on the other hand, possesses the driving force that Igbo and Yoruba lack. The Hausa language has established itself as a *lingua franca* throughout the north and even in the southern parts of the so-called Middle Belt (a large region that roughly occupies the central part of the country from east to west). It is likely that Hausa has somehow met its geographical limits, especially south of the Middle Belt, and that the next stages of expansion will be its affirmation and consolidation as L1 in areas where it is already present as a vehicular language.

So what is the future for minority languages? What will be the linguistic layout of the country in 50 years' time? In all likelihood, Brann's pyramid will remain unchanged in its structure, but will see the thickness of the chthonolect level considerably reduced. On the long run, minority languages spoken by monolingual communities will be replaced by a vehicular language (which will then become L1). Generally, speakers of chthonolects consider their language as lacking in prestige and show a certain indifference towards its disappearance. Moreover, minority language communities often lack political leverage, and the kind of resistance they can exert to language assimilation is almost non-existent.

The current policy approach to the protection of linguistic diversity is consistent with the projection outlined so far. It is interesting to note that often the only bastion of language protection is represented by the local 'language boards', voluntary associations with no official support set up within the communities with the aim of promoting (i.e. maintaining and defending) the local language.

term *Naija*, while speakers identify it with the expression 'broken English'. Nigerian Pidgin is different from Nigerian English, which is the variant of English widespread in Nigeria (although some lexical overlaps are inevitable, cf. Blench 2005). Although Nigerian English is establishing itself as a mother tongue among an affluent élite living in urban centres such as Lagos or Abuja (Adeyanju 2009), its position in terms of *lingua franca* is still very weak and will not be discussed here.

An important change that has taken place in Nigeria concerns the ability of minority groups to escape the control of dominant groups, thus escaping the sphere of influence of metropolises. The areas that Jungraitmayr & Leger (1993) define as areas of “ethnic and linguistic compression” (clusters of ethno-linguistic groups in relatively isolated areas) are increasingly rare. Whereas in the past groups migrated to remote areas to escape the control of metropolises or because of conflict, settling in areas where the dominant languages had not yet penetrated, today this isolation is no longer possible (nor, in most cases, desired): population density, commercial and infrastructural development, and the downsizing of conflicts have neutralised the effects of distance, helping to accelerate the dynamics of cultural and linguistic assimilation.

On the basis of what has been discussed so far, it is possible to formulate a prediction of the Nigerian linguistic layout of the future in which the total number of languages will decrease, while the number of Hausa and NP speakers will increase. Of course, we could reason in terms of what will survive and what will die, but there is another way of analysing the issue. There is, in fact, a difference between the languages present in a certain region and the languages (or the language) actually spoken by the population. So far, we have treated languages as discrete units, elements that may or may not be there, live or die, be used or not. However, in a situation of high language density, people do not tend to speak in discrete units, i.e. in abstract blocks distinct from each other. The notion of ‘code-switching’, for instance, does not help to describe the fluidity of language use: identifying code-switching practices may be useful at the descriptive level, but it does not give us any indication of *which language* is actually being used by the speaker. The notion of ‘translanguaging’ disseminated by Garcia & Wei (2014) may be more useful in illustrating the process at work. Multilingualism is seen as an obsolete term for what is an integrated and fluid system in which it is no longer possible to separate languages once they occur in speakers’ utterances. Migration to large urban centres coupled with the high growth rate recorded in Nigeria favour the phenomenon of translanguaging, and in particular of what Otsuji & Pennycook (2010) identify as ‘metrolingualism’.⁶

Urban centres absorb people (just as the Kopytoff’s frontier absorbs segments from metropolises), and thus speakers, but not necessarily new languages in the sense of language communities proper. In this context, Brann’s pyramid – which is built on discrete units for the purpose of language planning – is no longer a suitable model to describe the state of affairs.

Certainly, the standardised languages (e.g. Hausa, Igbo, Yoruba, and English) will be somewhat shielded from the forces of translanguaging, but the vehicular code in use will assert itself independently of the standard lan-

⁶ For a discussion of the scope and theoretical implications of the notions of ‘translanguaging’, ‘fluidity’ and ‘superdiversity’ in the African context see Wolff 2018.

guages. Translanguaging is the result of flows emanating from language communities, i.e. from languages in use by relatively large groups of speakers. Chthonolects, as we have said, are destined to be replaced by more functional or more prestigious languages, and therefore their role in the construction of translanguages is rather limited: the migratory flows expressed by chthonolect-speaking communities have little or no impact within the translanguaging process. The situation is different for ethnolects and choralects, which are often already present in the large urban centres and have an important numerical consistency and diffusion.

6. CONCLUSION

The absence of language protection policies, demographic growth, the pressure of prestigious or highly functional vehicular languages, the gradual disappearance of refuge areas, and the fact that in a large number of cases language no longer symbolises identity: these are the main factors behind the gradual disappearance of minority languages, i.e. those languages identified by Brann as 'chthonolects'. Internal migrations towards the large urban centres will stimulate the emergence of two main vehicular languages, Nigerian Pidgin and Hausa. In densely linguistic urban contexts, vehicular codes could take on the features of 'translanguages', i.e. fluid languages made up of flows, exchanges and grafts derived from vehicular and areal languages. In this context, while it is almost certain that the 'languages of the soil' will have no place, national (demolects) and sub-national (ethnolects) languages will contribute to the formation of larger vehicular codes.

REFERENCES

- Adeyanju, Dele (ed). 2009. *The Sociolinguistics of English and Nigerian English*. München: Lincom Europa.
- Batibo, Herman M. 2005. *Language Decline and Death in Africa*. Clevedon, Buffalo, Toronto: Multilingual Matters.
- Blench, Roger. 1998. "The status of the languages of central Nigeria", in: Matthias Brenzinger (ed.) *Endangered Languages in Africa*, 187-206. Köln: Rüdiger Köppe Verlag.
- Blench, Roger M. 2005. A dictionary of Nigerian English, <https://www.rogerblench.info/Language/English/Nigerian%20English%20Dictionary.pdf> (accessed 1 April 2021).
- Blench, Roger. 2015. "Endangered languages in West Africa", in: Matthias Brenzinger (ed.), *Language Diversity Endangered*, 140-162. Berlin, Boston: De Gruyter Mouton.
- Brenzinger, Matthias (ed.). 1998. *Endangered Languages in Africa*. Köln: Rüdiger Köppe Verlag.
- Brann, Conrad Max Benedict. 1977. "Language planning in Nigeria: Problems and prospects". *Africa* 3, pp. 317-336.
- Brann, Conrad Max Benedict. 1993. "Democratisation of Language Use in Public Domains in Nigeria". *The Journal of Modern African Studies* 31(4), pp. 639-656.
- Brann, Conrad Max Benedict. 1994. "The National Language Question: Concepts and Terminology", *Logos* 14, pp. 125-134.
- Eberhard, David M., Gary F. Simons & Charles D. Fennig (eds.). 2021. *Ethnologue: Languages of the World*, 24th edition, Dallas: SIL International, Online version: <http://www.ethnologue.com>.
- Edwards, John. 2010. *Minority Languages and Group Identity*. Amsterdam and Philadelphia: John Benjamins.
- Essegbey, James. 2020. "Language endangerment, documentation, and revitalization", in: Reiner Vossen & Gerrit J. Dimmendaal (eds.), *The Oxford Handbook of African Languages*, 833-842. Oxford: Oxford University Press.
- Garcia, Ofelia & Li Wei. 2014. *Translanguaging: Language, Bilingualism and Education*. Basingstoke, Palgrave Macmillan.
- Igboanusi Herbert & Lothar Peter. 2004. "Oppressing the oppressed: the threats of Hausa and English to Nigeria's minority languages", *International Journal of the Sociology of Language* 170, pp. 131-140.
- Jungraithmayr, Herrmann & Rudolf Leger. 1993. "The Benue-Gongola-Chad basin. Zone of ethnic and linguistic

- compression". *Berichte des Sonder-forschungsbereichs* 268 (2), pp. 161-172.
- Kopitoff, Igor. 1987. "The internal African frontier: The making of African political culture", in: Igor Kopitoff (ed.), *The African Frontier. The Reproduction of Traditional African Societies*, 3-84. Bloomington and Indianapolis: Indiana University Press.
- Lewis, Melvyn Paul & Gary Simons. 2010. "Assessing endangerment: Expanding Fishman's GIDS". *Revue roumaine de linguistique* 55(2), pp. 103-120.
- Lüpke, Friederike 2015. "Ideologies and typologies of languages endangerment in Africa", in: James Essegbey, Brent Henderson, and Fiona McLaughlin (eds.), *Language Documentation and Endangerment in Africa*, 59-106. Amsterdam and Philadelphia: John Benjamins.
- Mann, Charles C. 1993. "The sociolinguistic status of Anglo-Nigerian Pidgin: an overview". *International Journal of the Sociology of Language* 100/101, pp. 167-178.
- Newman, Paul. 1969/70. "Linguistic relationship, language shifting, and historical inference". *Afrika und Übersee* 53, pp. 217-223.
- Ofolue, Christine I. 2010. "Towards the standardisation of Naija: Vocabulary development and lexical expansion processes", in: *Proceedings of the Conference of Nigerian Pidgin* (University of Ibadan, 8-9 July, 2009), Ibadan: University of Ibadan.
- Ofolue, Christine I. 2011. "Nigerian Pidgin and West African Pidgins: A sociolinguistic perspective", paper presented at the *SPCL Conference*, (University of Ghana, Legon, Accra, Ghana, 2-6 August, 2011).
- Otsuji, Emi & Alastair Pennycook. 2010. "Metrolingualism: fixity, fluidity and language in flux". *International Journal of Multilingualism* 7 (3), pp. 240-254.
- Simire, G. O. 2004. "The status of kom and the time-aspect-modality system of Anglo-Nigerian Pidgin". *Global journal of humanities* 3: 27-32.
- Vigouroux, Cecile B. & Salikoko Sangol Mufwene. 2008. "Colonization, globalization and language vitality in Africa: an introduction", in: Cecile B. Vigouroux & Salikoko Sangol Mufwene (eds.), *Globalization and Language Vitality: Perspectives from Africa*, 1-31. London, Continuum.
- Wolff, H. Ekkehard. 2018. "Multilingualism, translanguaging, and linguistic superdiversity: An Africanist's perspective on 'language'". *Nordic Journal of African Studies* 27 (2), pp. 1-21.

Linguistic ambiguity in Timor-Leste: local languages between pride and shame

CAROLINA BOLDONI
Universidade Nova de Lisboa

ABSTRACT

Emerging from centuries of Portuguese colonisation and a violent Indonesian military occupation lasted 24 years (1975-1999), Timor-Leste became the first sovereign state of the 21st century, in 2002. The Democratic Republic of Timor-Leste adopted Portuguese and Tetun Dili as the two national languages and this decision has been under debate since the birth of the nation, not just among local politicians, but also among foreign consultants and scholars. Based on an ethnographic fieldwork lasted 18 months, conducted between 2017 and 2018, this paper aims to address the ambiguities regarding the linguistic diversity in Timor-Leste. Local indigenous languages have been undergoing a process of governmental recognition (being 20 the local languages mentioned in the Constitution as ‘mother tongues’) and other international institutions have developed projects aimed to their safeguard. However, often both in national institutions as well as to a more grassroot level, the local linguistic diversity is considered more as a burden than as an asset. By focusing on these local ambiguities, the paper aims to discuss the interaction between linguistic prestige and status, as well as economic relations embedded in the Timor-Leste national linguistic policies.

KEYWORDS

East Timor; Timor-Leste; indigenous languages; nation building; national languages; shame.

INTRODUCTION

“UNESCO knows nothing about Timor-Leste”.¹ These were the strong words used by Taur Matan Ruak, President of the Democratic Republic of Timor-Leste, in 2017, commenting on the EMBLI (Mother Tongue Based Multilingual Education Programme). EMBLI is an international plan coordinated by UNESCO, which aims to strengthen the use of native local East Timorese languages as teaching languages in the primary schools throughout the country.² Taur Matan Ruak’s words were broadcast on national TV, during the 8 PM news that most East Timorese watch. I was watching the TV me too, together with the family who hosted me in Daralata during my doctoral ethnographic fieldwork.³ The strong words pronounced by the President of the Republic of Timor-Leste really struck me.

Why was the President of Timor-Leste so angry with UNESCO? What is the role of the local native languages in the country? And how are they perceived? What is EMBLI trying to implement and why the governmental structure seems against this plan? Discussing these questions is helpful to understand the profound complexity of the country’s language policies and politics.

Timor-Leste is a quite linguistically diverse country: Ethnologue registers around 20 alive languages⁴ in a geographical area as big as Sicily and with a population of 1,3 million inhabitants.⁵ I started to get interested in the East Timorese context while I was attending my undergraduate course in Portuguese language, already in 2010. That small territory off the northern

¹ The original sentence in Tetun Dili was “UNESCO la hatene buat ida kona ba Timor-Leste”.

² Cf. <http://timor-leste.gov.tl/wp-content/uploads/2017/01/Eng-Executive-Summary-for-EMBLI-Endline-Assessment-V4.pdf>, last accessed on 05.08.2021.

³ I conducted my fieldwork between January 2017 and April 2018 in Venilale, subregion of the Baukau region, eastern district of the country. I lived with two different families of the area, the first in Daralata (February-August 2017), a very small village in the South, and the second one in Waikulale (between September 2017 and April 2018), a northern village.

⁴ Cf. <https://www.ethnologue.com/country/TL/languages>, last accessed on 05.08.2021.

⁵ Cf. <https://data.worldbank.org/indicator/SP.POP.TOTL?locations=TL>, last accessed on 05.08.2021.

coasts of Australia caught my eye precisely because of its linguistic diversity as well as because of its recent and troubled history. Former Portuguese colony, Timor-Leste has suffered a 24-years military invasion from Indonesia, acquiring its national independence only in 2002. Currently, the Constitution of the Republic of Timor-Leste recognises Portuguese and Tetun Dili or Tetun Prasa⁶ as the two official national languages; English and Bahasa Indonesia as work languages, and the local native languages are considered as part of the national heritage and identity that the State declares to preserve and develop.⁷ Tetun Dili is the most widely spoken language of the country, being often the second language of most of East Timorese people, especially outside of Dili, the capital city (Williams-van Klinken 2015). In fact, the people who speak Tetun Dili as their first or best language generally live in Dili. Public schools use both Portuguese and Tetun Dili as means of instruction. However, the number of the East Timorese speakers of Portuguese is attested around 10% of the total population (Simões 2015). The media use interchangeably both official languages, as well as Bahasa Indonesia. This last language is well known and understood by most of the population, especially the ones who were born and lived during the 24 years of the Indonesian military occupation (the so-called *jerasaun foun*, young generation).

In Daralata, in 2017, the language used for the daily communication was Makasae, mother-tongue of most of the population of the area. However, Makasae is but one of the local languages spoken in Venilale. In the subregion, in fact, there are at least other two local languages spoken: Kairui-Midiki and Waima'a. While Makasae belongs to the Papuan linguistic *phylum*,⁸ the other two belong to the Austronesian linguistic family.⁹ On TV the common languages spoken are Tetun Dili and Bahasa Indonesia, language spoken by the characters of soap operas, films, and advertisement on TV. There are two editions of the news, the main one in Tetun Dili and the second one in Portuguese. At the local primary school of the village, pupils and teachers used to speak Makasae with each other, but then Tetun Dili and Portuguese were used as vehicular and teaching languages during the lessons. This linguistic situation is not present only in the Venilale subregion, but it is quite common in the rest of the country too (Walter 2016).

⁶ The official language is commonly known as Tetun and even the Constitution refers to this language as Tetun. However, linguists have registered the presence of different varieties of Tetun (Williams-van Klinken 2010). Namely, Tetun Terik or Belu and Tetun Dili or Prasa. And this is the reason why I distinguish the official language (Tetun Prasa) from the Tetun Terik, a variety spoken in the eastern part of the country.

⁷ Cf. <http://www.unesco.org/education/edurights/media/docs/7aa8c8cd63d2e3ec8a6546d6ba1f4071161ce516.pdf>, pp. 11-12, last accessed on 05.08.2021.

⁸ Cf. <https://www.ethnologue.com/language/mkz>, last accessed on 05.08.2021.

⁹ Cf. <https://www.ethnologue.com/language/wmh> and <https://www.ethnologue.com/language/krd>, last accessed on 05.08.2021.

With this short, yet emblematic, depiction I want to point out the linguistic richness of Timor-Leste, by some considered linguistic confusion (McAuliffe 2009). This work aims to present some of the ambiguities within the national policies concerning languages in Timor-Leste, by pointing out the social and historical values and meanings attached to the languages spoken in the country. In Timor-Leste, in fact, Portuguese, Tetun Dili, Bahasa Indonesia, English, and the other local languages are far from being neutral ways of communication. Status and prestige are often embedded in them, and I suggest that discussing the historical context as well as social configurations is necessary in order to understand the values attached to all these languages.

This paper wants to discuss the current multilingual situation in Timor-Leste, by linking it to the recent past of the territory, as well as to current governmental plans, aimed to develop the use of foreign languages throughout the country. I will discuss both the role of Portuguese, Tetun Dili and the other local languages in the country, stressing the entanglements between old and new generations, as well as opposite ways of imagining both the imagined national community as well as the linguistic diversity of the country. In the first section I present the historical context, from the Portuguese colonisation to the Restoration of the national independence of 2002, focusing on the connections between languages, politics, and identity. The second part of the paper analyses instead the different perspectives existing with regard to the choice of having Portuguese as a national language. Finally, I focus on the role of native languages, by showing the ambiguities and the contradictions of some of the governmental decisions, related to the language teaching in schools.

HISTORICAL CONTEXT

Timor-Leste was a Portuguese colony until 1975. By that time, Portuguese had been an elitist language, spoken by a tiny East Timorese minority of so-called *mestiços* and *assimilados*¹⁰ (Costa 2001, 59-60; Donzelli 2012, 138, Leach 2003, 140). Gavin Jones points out that by the 1970s the literacy rate in Timor-Leste was estimated to be around 10%. The school system was coordinated by the Church and not by the colonial administration. The first and only secondary school of the country was opened only in 1952 and the number of students enrolled was between 200 and 800 during the 1960s. The number of schools present in the country in the aftermath of the end of the colonial era (1974-75) were only 47. There were no universities; in fact, the *mestiços* and *assimilados* who wanted to attend them, had to go to Portugal

¹⁰ These are two common Portuguese expressions that were used in the Portuguese colonies to refer to people born from relationships between native women and Portuguese men, as well as to local people who managed to attend Portuguese schools.

to do so (Jones 2003, 41-43). The Catholic Church translated the Bible into Tetun Terik (Anderson 1993; Costa 2001), and into other local languages too (such as in Makasae, for example), creating a written and high variety of the language. Despite the differences between Tetun Dili and Tetun Terik linguistic varieties, however, the decision of translating the Bible was crucial for the use of this language as a *lingua franca* throughout the country, including the rural areas and not only the urban centres.

With the Portuguese Revolution on 25 April 1974 (*Revolução dos Cravos*) and the destitution of the dictatorship in Portugal, the Portuguese authorities in Timor-Leste made possible the creation of the first local political parties.¹¹ However, violent tensions between the parties rose in 1975. The Portuguese governmental representatives left the territory during the summer of 1975 and the Fretilin party declared the national independence of Timor-Leste on 28 November 1975. However, the radical ideas conveyed by the party were interpreted as communist and considered as threatening internationally especially by Indonesia and Australia, neighbour countries, as well as by the USA. Indonesia used the pretext of the violence resulting from the political tensions as well as of the allegedly 'radical' left-wing political ideas of Fretilin to occupy the territory on 7 December 1975. The Indonesian military occupation in Timor-Leste lasted 24 years, until 1999.

During the Indonesian occupation, the language used by the administration, in the schools and by the media was Bahasa Indonesia,¹² while for day-to-day interactions the common languages used by the population were Tetun Dili and other local languages. It was during the Indonesian military occupation that the Catholic Church deliberately offered the mass in Tetun and not in Bahasa Indonesia. On the one hand, this choice has been interpreted as an active means of resisting Indonesian integration (Anderson 1993; Budiardjo; Liong 1984, 119-124). On the other, however, it helped the formation of a common sentiment of belonging related to Tetun, as a national language.

In addition, during the Indonesian military occupation, the armed East Timorese resistance used Portuguese as a means of communication, to avoid being caught and understood by the Indonesian army. This is among the reasons why Portuguese was chosen as one of the two official languages of the country (Ruak 2001). Moreover, Portuguese has been used as a common language of the Portuguese-speaking countries, during their fight against Portugal. The Portuguese-speaking world is composed by former Portuguese colonies that, together with Portugal, nowadays shape the CPLP (*Comunidade*

¹¹ The main parties formed were ASDT (Pt., *Associação Social-Democrata Timorense*), UDT (Pt., *União Democrática Timorense*) and APODETI (Pt., *Associação Popular Democrata Timor-Leste*).

¹² As Donzelli points out, the Indonesian administration prohibited the use of Portuguese (2012).

dos Países de Língua Portuguesa). I argue that this suggests that there might be a connection between all the former Portuguese colonies, namely the African one (Angola, Moçambique, São Tomé e Príncipe, Cape Verde, Guiné-Bissau,) and Timor-Leste, that had all fought against the Portuguese colonial power between the 1960s and the 1970s (Boldoni 2014). I suggest that it is relevant to consider the bonds between all the former Portuguese colonies, and not just the bond between Timor-Leste and Portugal, to understand the reasons why the national East Timorese government chose Portuguese as national language in 2002. The common language is a means through which the Portuguese-speaking world stays connected, not just a submission/recognition of the Portuguese *metropole*. Many important figures for the national independence, in fact, had connections with other former African Portuguese colonies.¹³ The ideological, political, social, and linguistic connection among the former Portuguese African colonies and Timor-Leste is still strong nowadays (Boldoni 2014).

By the beginning of the 1980s the schools in the country were more than 300 and the number grew to 600 in the 1990s (Jones 2003, 41-43). Clearly, the rapid growth of the schools during the Indonesian occupation as well as the improvement of the road system had to do with the Indonesian integrationist programme. However, some of my interlocutors, used to stress the fact that important services such as electricity, infrastructures, the mail services, and schools were developed only during the Indonesian occupation and not during the Portuguese colonisation. During colonial times, in fact, these services were present but only in Dili and only for the administrative elite (Boldoni 2014, 72-73).

The intense diplomatic activities of figures such as José Ramos-Horta and Bishop Carlos Filipe Ximenes Belo, who were awarded the Nobel Peace Prize in 1996, led to the realisation of the East Timorese national independence. In 1999, the UN created a regional military force to intervene in the region, with the aim of turning Timor-Leste into an independent nation. Timor-Leste finally gained national independence from Indonesia in 2002.

¹³ José Ramos-Horta, son of a Portuguese deported to Timor, was exiled to Mozambique for his anti-colonial activities during the 1970s. Alkatiri, an Arab descendant (his ancestors arrived in Timor around 1800 from Yemen), also left the country in the 1970s to study in Angola, where he contacted the local anti-colonial front (Niner 2007, 114). By the end of 1975, when Indonesia had already occupied the East Timorese territory, Alkatiri and Ramos-Horta, together with Rogério Lobato and Roque Rodrigues left Dili flying to an external delegation of Fretilin based in Mozambique (117). Not to mention other Timorese political leaders of the 1970s, who studied in Lisbon and came back to the country as anti-colonial leaders – because in Lisbon they were in contact with anti-colonial theorists from the other Portuguese African colonies.

With the Restoration of Independence of the country,¹⁴ on 20 May 2002, the national government chose Portuguese and Tetun Dili as national official languages. However, despite the official recognition of Portuguese as one of the two official languages of the Republic of Timor-Leste, and even though Timor-Leste is part of the CPLP,¹⁵ Ethnologue counts that Portuguese in Timor-Leste is a mother-tongue for less than 6000 people.¹⁶ Portuguese is in fact rarely used in common and day-to-day conversations: it is mostly used as a formal and prestigious language by government institutions, mostly in Dili. During both my fieldworks in the country, I could observe that Portuguese was spoken by a very small minority of the population, namely old people who could learn it before 1975 (Leach 2003, 140). Therefore, due to historical reasons, currently the *jerasaun foun* (young generation, born after 1975) speak Tetun and Bahasa Indonesia fluently, while the *jerasaun tuan* (old generation, born during the Portuguese colonisation) is relatively more fluent in Tetun and Portuguese – in addition to the native languages that have been used for domestic and daily conversations until the present day.

PORTUGUESE AS THE NATIONAL LANGUAGE

As mentioned, Portuguese was crucial in the definition of the nation, and this is intertwined with the Catholic presence throughout the territory. The current national narrative describes the Portuguese presence as having allowed the unification of the different kingdoms under the same jurisdiction, the same faith, and the same language, leading to the creation of the East Timorese nation (Arthur 2020; Boldoni 2020; Leach 2008, 145-146; Ruak 2001). The Catholic Church and missionaries used to administer the school system (Jones 2003), spreading both literacy and the Gospel, as mentioned before. From such a perspective, then, Catholicism as well as the Portuguese presence define the East Timorese identity as unique towards the neighbouring countries, especially towards Indonesia. As Indonesia is the largest Muslim country in the world, and the Indonesian territory of West Timor is mainly Protestant, Catholicism and the choice of Portuguese as a national language makes Timor-Leste a unique country within Southeast Asia. Due to

¹⁴ The East Timorese government considers the self-proclaimed independence of 28 November 1975 as the official date of the national independence. Therefore, 28 May 2002 corresponds to the Restoration of the Independence of the country, after 24 years of illegal military occupation of the Indonesian army.

¹⁵ *Comunidade do Países de Língua Portuguesa* (Pt., Community of the Portuguese Language Countries). Cf. <https://www.cplp.org/id-2778.aspx>, accessed on 08.12.2019.

¹⁶ Cf. <https://www.ethnologue.com/country/TL/languages>, last consulted on 4/8/2021.

its peculiar socio-historical identity, Timor-Leste deserved to be an independent country. However, this kind of narrative is not considered inclusive by many East Timorese citizens, namely the young generations (*gerasaun foun*).

The East Timorese political leadership which has been in charge since 2002 is the same that used to oversee the armed resistance during the Indonesian occupation (Leach 2008; Magalhães 2015, 34-36). Just to give an example, Taur Matan Ruak was the Chief of the Resistance against Indonesia during the 1990s. With the Restoration of Independence, he was appointed as Major General of the National Army (F-FDTL) from 2009 to 2011. From 2012 to 2017 he was the President of the Republic of Timor-Leste and since 2018 he has been the Prime Minister of the country. He justifies the choice of Portuguese as the national language for historical reasons, pointing out that Portuguese was the language used by the armed resistance during the independence fight against Indonesia (Ruak 2001).

However, other relevant perspectives should be taken into consideration. During my fieldwork, I discussed these issues with many young East Timorese people. Most of my interlocutors were well-educated and all had had the chance to study abroad, mainly in anglophone countries (UK, USA, Australia, New Zealand). They often questioned many central governmental choices, such as the choice of Portuguese as the national language or the inference of the Catholic Church within governmental political activities. They interpreted the choice of Portuguese as a national language as a way to look back to colonialism and a choice that does not fit with the practical needs of the young generations of Timor-Leste, within Southeast Asia. However, speaking Portuguese is among the criteria to be hired within the governmental structure, NGOs and other national institutions. In a country where a very small minority speak Portuguese fluently, this linguistic criterion seems to willingly exclude the youngest people from the institutional and political life of the country.

Aurora Donzelli reconfigures the choice of Portuguese as one of the two national languages as if it was “suspended between an ambiguous display of colonial nostalgia, a pragmatic choice aimed at minimising Australian and Indonesian political-economic influence and the attempt of constructing a new sense of national identity” (Donzelli 2012, 138). The young generation (*gerasaun foun*) feels misrecognised by the official perspective that interprets Portuguese as part of the national heritage and identity, besides feeling excluded from “symbolic sources of power” (Leach 2003, 141). As Leach suggests, the contraposition between ‘elder’ and ‘younger’ reflects a major ‘fault line’ in the national politics, between the political establishment – mainly constituted by the ‘old generation’ – and ‘younger’ ideas opposed to it (Leach 2008). He highlights the importance of languages in the tensions existing between the young and old generations, as well as of different ideas of the national narrative and identity, linked to the colonial legacy (2003; 2008).

Choosing Portuguese as one of the two national languages has not been a neutral nor a consensual choice in East Timor. The disagreement related to this choice led to different ways of conceiving the public schooling teaching, as I present in the following section.

THE PLACE OF LOCAL LANGUAGES

During my fieldwork I spoke only Tetun Dili with my interlocutors. I tried to learn the local languages that my hosts used to talk, but without improvements. People were surprised about the topic of my research, my willing to better understand local languages and traditions. In Dili, many of my interlocutors often asked “What can you learn over there?”, in the rural areas of the country. When I met for the first time my hosts and the local Mayor, in their welcoming presentation of Venilale, they all listed the different languages spoken within the territory, as well as the name of the villages. They described the differences between places and villages, highlighting the internal divisions and boundaries (both physical and cultural) among them. Differentiation and diversity were foundational aspects of the ‘community’ itself. However, at the same time, my interlocutors often found hilarious that a foreigner researcher wanted to learn their *dialects* – as they used to name them – and often described their language as uneducated and not suitable for an educated woman – as I was to their eyes.

Both the local population and the governmental apparatus often attach stigma and shame on the native local languages in Timor-Leste. These languages are often considered inferior, belonging to uneducated people. The short video broadcast by the news that I referred to in the introduction of this paper, continued with Taur Matan Ruak explaining that East Timor cannot compete with other neighbouring nations, if the children of the country only speak “dialects”. He then stressed that children in school should learn Portuguese, English and Bahasa Indonesia, not Fataluku.¹⁷

In Daralata, my hosts wanted me to teach Portuguese at the local primary school: the reception of the teachers was incredibly warm, despite my uneasiness in that role. To them it was not relevant that I was not Portuguese, nor that I had never taught Portuguese to children. Some of the teachers did not understand Portuguese themselves, hence their request to help. But above all, my identity represented what the teachers and my hosts considered prestigious: to them, I was a white European woman, who had studied Portuguese,

¹⁷ Fataluku is one of the native languages spoken in Timor-Leste, namely in the eastern region of the country, Lospalos, where the EMBLI Project has been implemented in the last years. Cf. <https://www.laohamutuk.org/educ/CAFE/PolicyPt.pdf>, last accessed 09.08.2021.

and who had come to East Timor to get in touch with the local culture. It was my chance to give back the hospitality that the community was giving me and at the same time, they felt lucky to have a “Portuguese teacher” in Daralata.

Before starting the lessons, the children would sing the national anthem, in Portuguese; they would then pray the Hail Mary and Our Father in Tetun Terik, used by the Church for the liturgical texts; they spoke Makasae among them and with the teachers. However, the lessons were in Tetun and Portuguese. As soon as I started the teaching, I realised that most of the children in the class (between 8 and 9 years old) did not understand Tetun either. How was it possible to teach them a foreign language, such as Portuguese, using an unknown language, that was Tetun Dili?

To this very question, UNESCO provided an answer, by coordinating the EMBLI Pilot Educational Programme. This pilot project started to be implemented in 2012 and continues to the present day in 10 schools where 99% of the educational population speaks the same local language: four schools in the region of Oé-cusse, where the mother-tongue is Baikeno, three in Lospalos, where most of the population speak Fataluku, and three in Manatuto, where the language spoken is Galolen. EMBLI has raised several questions and critiques, especially among the governmental institutions, highly skeptical towards the project. The project prepares teachers and materials for teaching, using local languages. The project aims to provide schooling education in the common language of the children, and, at the same time, it helps keeping local languages alive. Tetun and Portuguese are introduced gradually, since they are considered as foreign languages. At the end of 2015, an external evaluation of the project measured its success in reaching its educational goals. The report showed that pupils in EMBLI schools demonstrated better reading skills and were more successful in the various subjects.¹⁸

However, many are the concerns related to this policy, such as that the project might minimise the effectiveness of learning the official languages. Another concern is related to the fact that teaching in different languages throughout the country might lead to regional divisions that threaten national unity. On the other side, however, supporters of this policy argue that EMBLI promotes social inclusion and tolerance for difference in a multilingual and multicultural society and therefore promotes unity building.

At the same time, in the attempt of strengthening the national identity, as well as increasing the number of speakers of Portuguese, the East Timorese government has implemented the CAFE educational programme.¹⁹ CAFE schools are public schools in which Portuguese is used as teaching language,

¹⁸ Cf. <http://timor-leste.gov.tl/wp-content/uploads/2017/01/Eng-Executive-Summary-for-EMBLI-Endline-Assessment-V4.pdf>, last accessed on 10.08.2021.

¹⁹ CAFE is the acronym for Centros de Aprendizagem e Formação Escolar, meaning School Learning and Training Centres.

since pre-school to the secondary level. The main mission of the CAFE project is to improve the quality of teaching and promote fluency in the Portuguese language for children. CAFE is the result of a bilateral agreement, signed between Portugal and East Timor in 2009. There are currently 13 CAFE schools, one in each region of the country, all of which located in urban areas. According to the NGO La'o Hamutuk,²⁰ less than 3% of the total number of students attending all public schools attend CAFE schools.²¹

Unlike the project promoted by UNESCO, the CAFE schooling approach argues that the use of Portuguese as the teaching language at school, since the nursery, will help children learning Portuguese as a primary language. Unfortunately, the majority of pupils in Timor-Leste do not have access either to the EMBLI or to the CAFE schools, but to schools that resemble the model provided by the public school of Daralata, especially in the rural areas of the country.

CONCLUSIONS

Despite the fact that local indigenous languages have undergone a process of governmental recognition in Timor-Leste, both governmental institutions as well as other social actors seem to consider the local indigenous languages as part of an undeveloped and backward past. This work aimed to present some of the ambiguities related to the national linguistic policies, highlighting the values and meanings embedded within languages. Historical as well as political reasons are deployed by the different political stances to justify the use of a language or another; status, prestige and shame are other important aspects that led to public disagreement as well as governmental decisions. Languages are fundamental within the nation building that East Timor has been undergoing, since they are powerful identity markers. For all these reasons, many are the ambiguities and contradictions embedded in the governmental decisions related to the linguistic policies in the country.

²⁰ La'o Hamutuk, funded in 2000, is an East Timorese NGO that analyses and reports on development processes throughout the country. Literally, 'la'o hamutuk' means walking together in Tetun Dili.

²¹ Cf. <https://www.laohamutuk.org/educ/CAFE/18CafePt.htm#summ>, last accessed 10.08.2021.

REFERENCES

- Anderson, Benedict. 1993. "Imagining East Timor". *Arena Magazine* 4, pp. 7-33.
- Arthur, Catherine. 2019. *Political Symbols and National Identity in Timor-Leste*. Belfast: Palgrave Macmillan.
- Boldoni, Carolina. 2014. *Ami Timoroan. Il contesto Timorese fra il 1975 e il 1999*. Unpublished MA thesis in Cultural Anthropology. Venice: Università Ca' Foscari.
- Boldoni, Carolina. 2020. *Heritage Politics in Timor-Leste: Ambivalent Perspectives from Venilale*. Unpublished PhD. thesis. Lisbon: ISCTE and Universidade Nova de Lisboa.
- Budiardjo, Carmel, & Liem Soei Liong. 1984. *The War Against East Timor*. London: ZED Books.
- Costa, Luís. 2001. "O tétum, factor de identidade nacional". *Revista Camões* (14): 59-64.
- Donzelli, Aurora. 2012. "The Fetish of Verbal Inflection: Lusophonic Fantasies and Ideologies of Linguistic and Racial Purity in Postcolonial East Timor" in: *Language, Ideology, and the Human: New Interventions*, by Dusan Radunović & Sanja Bahun (eds.), 131-157. London; New York: Routledge.
- Jones, Gavin W. 2003. "East Timor: education and human resource development". In *Out of the Ashes: Destruction and Reconstruction of East Timor*, by James J. Fox & Dionisio Babo Soares (eds.), 40-52. Canberra: ANU Press.
- Leach, Michael. 2003. "Privileged Ties: young people debating language, heritage and national identity in East-Timor". *Portuguese Studies Review* 11 (1): 137-150.
2008. "Difficult memories: the independence struggle as cultural heritage in East Timor". In *Places of pain and shame: dealing with 'difficult heritage'*, by William Logan & Keir Reeves (eds.), 158-176. London; New York: Routledge.
- Magalhães, Fidelis. 2015. "Past, Present and Future: Why the Past Matters". In *A New Era? Timor-Leste after the UN*, by Sue Ingram, Lia Kent & Andrew McWilliam (eds.), 41-58. Canberra: ANU Press.
- McAuliffe, Peter. 2009. *Modern-Day Babel: Language Politics and Building the Rule of Law in East Timor*. *Hibernian Law Journal* 8(1), pp. 119-154.
- Niner, Sara. 2007. "Martyrs, heroes and warriors: the leadership of East Timor". In *East Timor: beyond independence*, by Damien Kingsbury & Michael Leach (eds.), 113-128. Clayton: Monash University.

- Ruak, Taur Matan. 2001. "A importância da língua portuguesa na resistência contra a ocupação indonésia". *Revista Camões* (14): 1-2.
- Simões, Fernando Dias. 2015. "Law and Language in Timor-Leste: Bridging the Divide". *ISEAS–Yusof Ishak Institute* 37 (3): 381-405.
- Walter, Stephen. 2016. *The Mother Tongue-Based Multilingual Education (EMBLI) Pilot Program Endline Assessment*. Dili: Government's Press.
- Williams-van Klinken, Catharina. 2010. "Two Tetuns: Language change before your very eyes", paper presented at the 6th International East Nusantara Conference on Language & Culture Kupang, 4-6 August 2010.
2015. *Mapping the mother tongue in Timor-Leste. Who spoke what in 2010?* Dili: Dili Institute of Technology.
- Assessment-V4.pdf [09.08.2021]
- East Timor population's data <https://data.worldbank.org/indicator/SP.POP.TOTL?locations=TL> [09.08.2021]
- East Timor's Constitution <http://www.unesco.org/education/edurights/media/docs/7aa8c8cd63d2e3ec8a6546d6ba1f4071161ce516.pdf> [09.08.2021]

Web references

- Ethnologue website <https://www.ethnologue.com/country/TL/languages> [09.08.2021]
- EMBLI's Assessment <http://timor-leste.gov.tl/wp-content/uploads/2017/01/Eng-Executive-Summary-for-EMBLI-Endline->

This book was conceived during the closing event of the DiM project, developed within the framework of the *Erasmus plus KA204 - Strategic Partnerships for Adult Education programme*. It is the fruit of a collective effort, aimed at grasping the complex and changeable set of phenomena connecting language and identity. Its fourteen chapters intend to offer food for thought on some of the currently most debated questions for linguists in the global village, and are divided into three thematic sections: 1) multilingualism, minority languages and the eternal dichotomy between orality and writing; 2) lexicography and L2 teaching; 3) the role of linguistics in particularly complex multilingual contexts. The book was published thanks to a grant obtained in 2018 by Regione Friuli Venezia Giulia.

ILARIA MICHELI, coordinator of the scientific committee of the AtRA series, is Associate Professor in African Linguistics at the University of Trieste.

FLAVIA AIELLO is Associate Professor in Swahili at the University of Naples “L’Orientale”.

AMELIA (LIA) PENSABENE is a teacher of English at the CPIA Avellino, where she was coordinator for the DiM project.

MADDALENA TOSCANO is a former researcher and teacher of Swahili language at the University of Naples “L’Orientale”.



Euro 18,00