Typecraft
Software for the documentation of minority languages

Ota Ogie and Dorothee Beermann

NTNU, Trondheim, Norway

Africa hosts 2000 of the 6000 languages of the world.

**Although spoken by millions of people
these languages have received relatively little
scientific attention.**

**We would like to show today one way to make
these languages more known and to add to
their international recognition.**

First the tool:
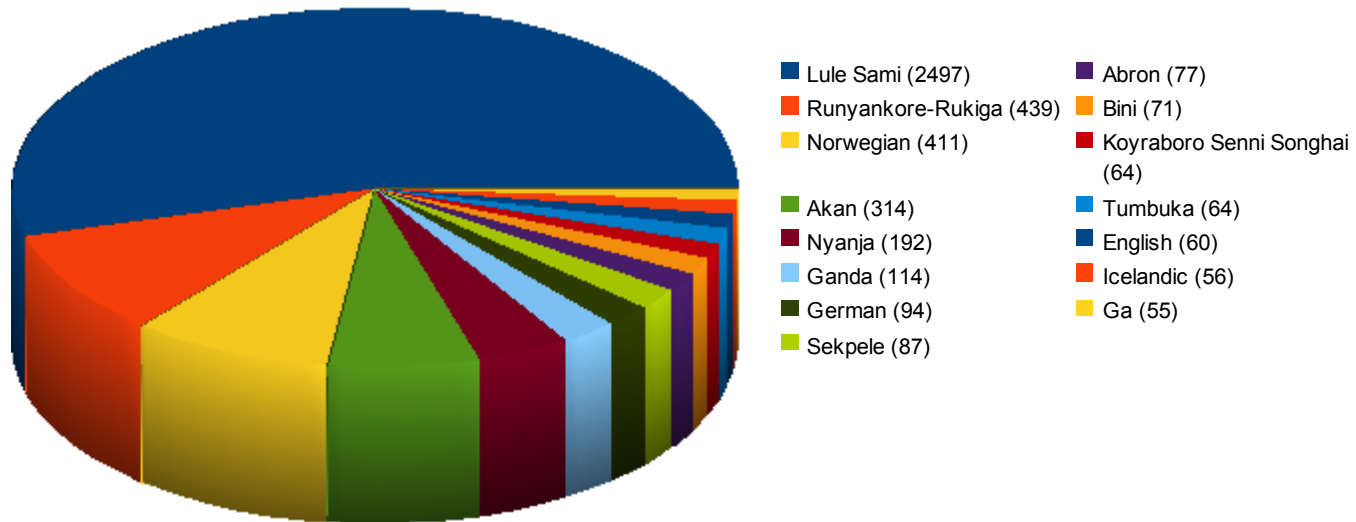TypeCraft and what it does

Second the matter at hand:
Hausa, Igbo and Yoruba are the three main languages
of Nigeria, but also Èdó is a language of Nigeria.

Studies like Ota's PdD thesis can enhance the
 recognition of the language and the people that speak it.

**Science can boost a language and thus support the
people that speak it !**

Some facts...

Top 15 TypeCraft languages

(by number of phrases)

- Lule Sami (2497)
- Runyankore-Rukiga (439)
- Norwegian (411)
- Akan (314)
- Nyanja (192)
- Ganda (114)
- German (94)
- Sekpele (87)
- Abron (77)
- Bini (71)
- Koyraboro Senni Songhai (64)
- Tumbuka (64)
- English (60)
- Icelandic (56)
- Ga (55)

TypeCraft

With TypeCraft the user adds *linguistic annotation* to
written material which is stored in a relational
*database* from where it can be retrieved using
multiple views.

Annotations are created on *multiple tiers*.

Texts as well as annotations are in *Unicode*. Annotated data can be
exported to standard text editors (WORD, Open Office and LaTex)
as well as to XML format. Archiving can be achieved through the
export of *XML with embedded DTD*.

TC has been designed for projects on *minority languages*, joining
linguists and native speakers. It is intuitive to use and
allows wide and *distributive usage.*

TC uses PostgreSQL as database format and is written in Java.
It can be freely used online with Mozilla Firefox, although at
present a login is required. TC runs on a server owned by NTNU.
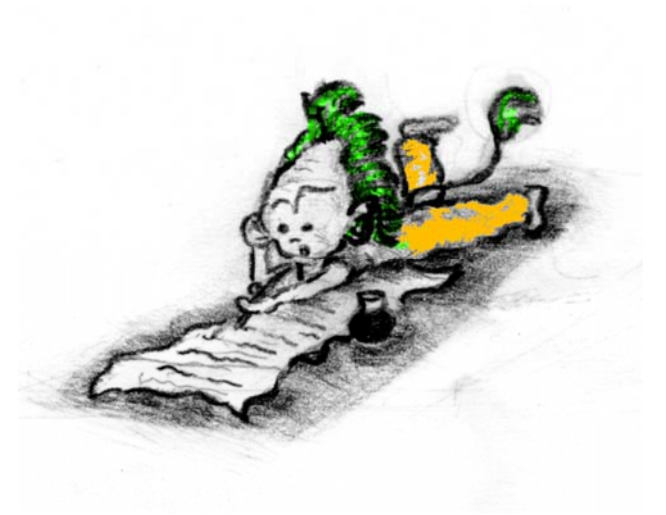
Science of language and language use

You think it is normal to learn in your own language.

They do not!

In 1976 the Nigerian Educational Council suggested that also children speaking Ẹ̀dó and others of the 9 major languages should be allowed to be taught in their own language in addition to the languages of national integration.

Ota Ogie has worked on serial verb constructions (SVC)
in Ẹ̀dó and other African languages.

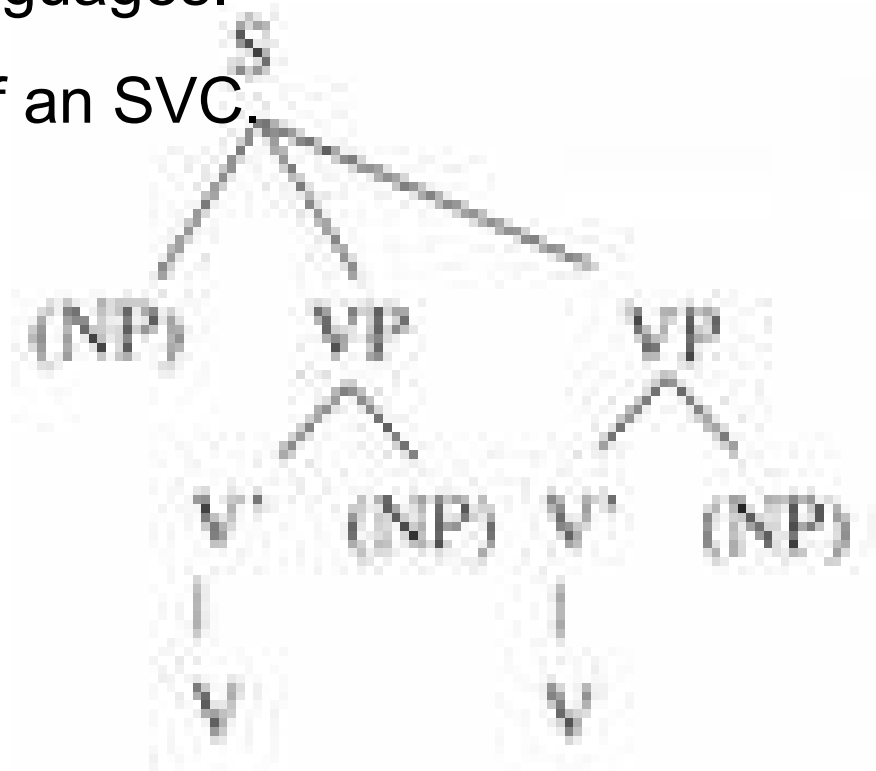To the right you see the skeleton of an SVC.

Òzó lé ízẹ̀ ré

Òzó	lé		ízẹ̀	ré
òzó	lé		ízẹ̀	ré
Ozo	cook.PAST		rice	eat.PAST
PN	V		N	V

"Ozo cooked rice and ate"

Generated in TypeCraft.

To the left you see the annotated version of and SVC
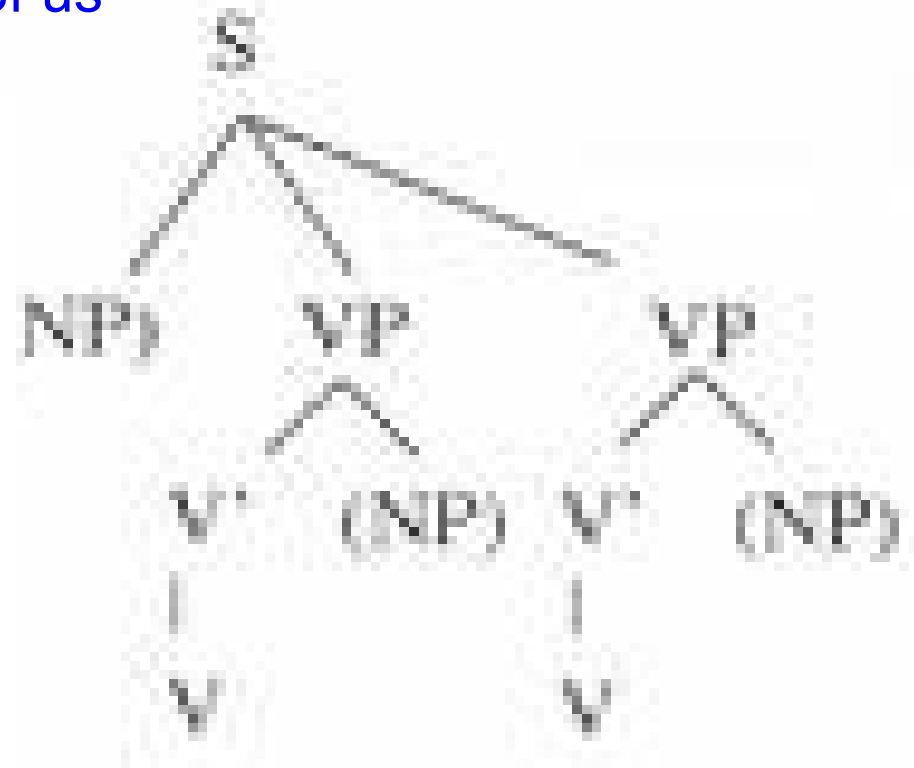from Ẹ̀dó , a language of Nigeria spoken by 1 Million peop

Through the creation of a universal meta-language,
part of which you see here in the background,
Ẹ̀dó becomes accessible to all of us

Gbomɔ kpatuuu kpɔɔ

Gbòmɔ̀	kpá!túú		kpɔ́ɔ̂	
Gbòmɔ̀	kpá!tú	ú	kpɔ́	ɔ̂
person	*do*	NEG.AOR	*pluck*	NEG.AOR
CN	V		V	

"*A person doesn't prosper without effort - no pain no gain*"

Generated in TypeCraft.

...and if we take good care of this structural data
it will be preserved for many generations to come and with in
our knowledge of the culture that created this language

Ẹ̀dó has been documented over hundreds of years, but documents are either out of print or not easily accessible. Ota's thesis is the first one that will make annotated data from Ẹ̀dó and other African languages directly accessible to researchers, teachers and students all over the world

Ẹ̀dó a language of Nigeria ISO 639-3 (bin)

Ethnologue
Languages of the World

How did TypeCraft help with linguistic research on Ẹ̀dó and with the effort to make the language more known?

positive-declarative -----achievement- V+modifier construction

Ọ̀zó rhùlẹ́-rè làọ́ òwá

Ọ̀zó	rhùlẹ́rè		làọ́	òwá
òzó	rhùlẹ́	**rè**	làọ́	òwá
ozo	*run*	PST	*enter*	*house*
PN	V		V	N

"Ozo ran into the house"

Generated in TypeCraft.

meaning and function of morphemes

meta-information about sentences

grammatical categories

How did TypeCraft help with linguistic research on Ẹ̀dó and with the effort to make the language more known?

In order to be understood not only by native speakers language has to be marked with symbols that represent the meaning and function of every unit. This can be done with TypeCraft

It offers a set of predefined symbols

Words, morphemes and whole sentences can be marked-up which allows for in-depth annotation.

Data produced in TypeCraft can be easily exported to other formats and to other applications.
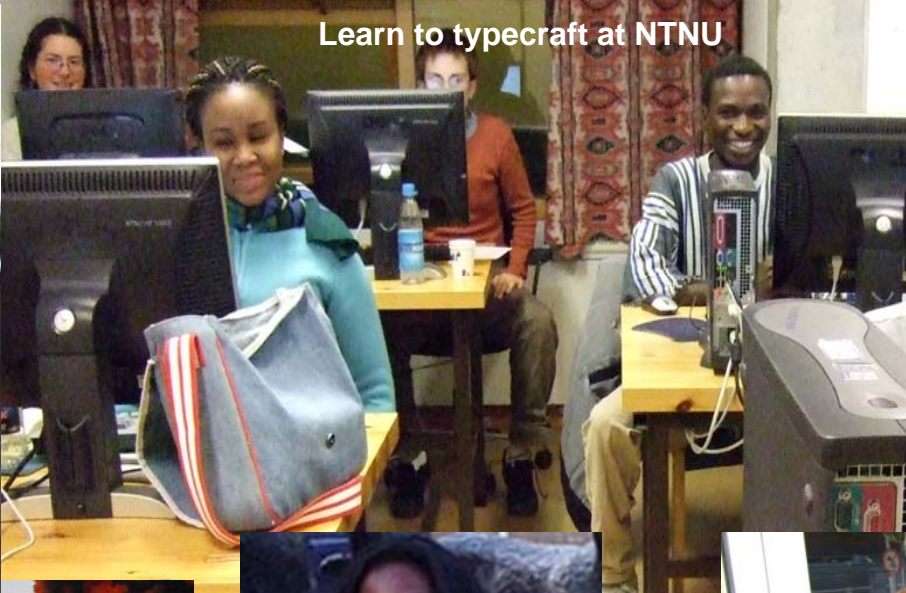
TypeCraft
The Natural Language Database

Learn to typecraft at NTNU

University of Malawi
CENTRE FOR LANGUAGE STUDIES
P.O. Box 108, Zomba
Tel.: 01 527 460/ 01 525 709
Fax: 01 525 587

Represent your project

Svenn Egil

Ben

Ota

Dorothee

Pavel

You meet other people that care for languages and the people that speak them