NTNU
Innovation and Creativity

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

In 1976 the Nigerian Educational Council suggested that also children speaking Òdó 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 Èdó and other African languages.

To the right you see the skeleton of an SVC.

```
Ozó lé ízè ré
Ozó lé ízè ré
ózó lé ízè ré
Ozo cook:PAST rice eat:PAST
PN V N V
```

"Ozo cooked rice and ate"

To the left you see the annotated version of an SVC from Èdó, 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, Ẹdó becomes accessible to all of us.

Gbọmọ kpáltúú kpọọ

Gbọmọ kpáltú ú kpọ ọ

person  do  NEG.AOR  pluck  NEG.AOR

CN  V  V

“Á person doesn't prosper without effort - no pain no gain”

...and if we take good care of this structural data it will be preserved for many generations to come and with our knowledge of the culture that created this language.
Èdó 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 Èdó and other African languages directly accessible to researchers, teachers and students all over the world.

Èdó a language of Nigeria ISO 639-3 (bin)
How did TypeCraft help with linguistic research on Èdó and with the effort to make the language more known?

Generated in TypeCraft.

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

“Ozo ran into the house”

meta-information about sentences
grammatical categories

meaning and function of morphemes

Generated in TypeCraft.
How did TypeCraft help with linguistic research on Èdó 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.
You meet other people that care for languages and the people that speak them.

Learn to typecraft at NTNU

Represent your project