Nwed Usem Ibibio: Towards the Empowerment of a Non-Major Language

Moses Ekpenyong
Department of Computer Science, University of Uyo, Nigeria
mosesekpenyong@uniuyo.edu.ng, gmail.com
Contents

- Nigerian Languages
- Language Diversity: How has Nigeria fared
- Why must we empower our Languages?
- What Language Policies do we Propose?
- Challenges we Identified
- Proposed Solution
- Ibibio
- Empowering Ibibio
- The Ibibio Language
- Ibibio Lexicographic Materials
- Ibibio Speech Resources
Nigerian Languages

- **Nigerian Languages (Urua & Udoh, 2014)**
  - 3 major and over 450 minors
  - The over 400 languages are struggling to survive. They are either:
    - Moribund (those no longer used, not transmitted, threatened)
    - Retreating (dying in an area, but flourishing in another area), inter-country boundaries
    - Under developed (no orthographies, written lit, meta lang)
    - Developing (have fairly developed orthography, lit. tradition and meta-language in course)
    - Pidgin (Connell 1994) - spoken by the younger illiterate ones especially when people migrate from one language region to another
Language and Diversity: How has Nigeria Fairied

- Before 1914, there had flourished local literatures in the various languages spoken in what is now called Nigeria.
- In the North, Arabic literary scholarship was the vogue.
- Much poetry blossomed, whether in Arabic or Ajami (the Hausa version of the Arabic language)
Language and Diversity: How has Nigeria Faired

- In the South, before 1914, traditional literature held sway.
- Believes and deeply rooted traditions were the order of the day
- People were proud of the language they spoke and upheld moral values
- But, something happened …
  - Did we suddenly lose direction?
  - Do we blame it on Religion?
  - The civil war?
Language and Diversity: How has Nigeria Faired

- Were did we miss it! The Civil War?
  - We have wars everyway and in my opinion, it’s a necessary evil.
  - Every war yields its own literature; the Nigerian civil war of 1967-1970 could not be different.
  - Due to the agitation for resource control
    - Two warring factions emerged: Biafra vs. Nigeria
    - But because of language barriers, Biafra could not be sustained and was defeated
Language and Diversity: How has Nigeria Faired

- Currently we have the Boko Haram – a hydra headed monster,
- paved a further slice to language unity, even in the North and is re-awakening Biafra
Language and Diversity: How has Nigeria Fared

- After the Civil War …?
  - Modernization did set in and led to Developments
  - Influence of the churches led to the abandonment of our local traditions
  - Drive for western education and the use of English affected the various languages
  - Government were more interested in power, and ignored the very essence of life
Language and Diversity: How has Nigeria Faired

- Are we heading to another Civil War?
  - I personally don’t think so
  - But some people think it is necessary
  - The politics of language once dominated us in a bid to evolve a new currency of language called WAZOBIA coined after the English word ‘come’ Wa-Yoruba, Zo-Hausa and Bia-Igbo
    - Suddenly there was an uprise by the minor languages
      - They felt marginalized
    - Then, every language community woke up to resist this dominance
    - This was how the initiative faded
    - perceiving another crises, every states of Nigeria is proposing as a matter of urgency the development of a widely spoken language
Why must we empower our Languages

- Because language is life and defines our very existence
- Some of the reasons why we want to empower the languages in Nigeria include, to:
  - Raise the status of the language
  - Save them from extinction
  - Ensure fairness to the speech community
  - Bring development s to the speech community
  - Preserve and sustain the language into the future
  - Encourage the diversity of the language (the ability to speak more than one language)
What Language Policies do we propose?

- Policies that would favour the use of indigenous languages (not just one official ones at the expense of others). This is to avert language endangerment.
  - In realization of the endangerment threat, the state Governments now put in place policies designed to protect and promote languages whose viability is threatened.
- Policies that would support Lexicography (e.g. we need new dictionaries focused at bridging the gap between the African languages and English)
- Policies that would promote the use of our indigenous languages in various domains (e.g. In the health domain, agricultural domain, education domain, political domain, etc.).
- Policies that would encourage and fund as a matter of urgency the archiving of important aspects of our language using state-of-the-art/standard tools
Challenges we identified

- Data sparseness (lack of sufficient corpus and domain specific vocabularies in the language)
- Are these policies politically-driven? Do we need to lobby to enact them?
- Where are our language regulators? – bodies that regulate standard languages, (The Yoruba Academy, Society of promoting Igbo language and culture).
  - Do we need more academies? Are they important? Will they help the indigenous languages?
- Do we have the manpower and interest to empower and save our dying languages from heritage collapse?
- Will we be justified to see our languages go just like that?
- What do we do?
Proposed solution

- We need to reflect on these things and aggressively halt this chain reaction, so we feel every one should
  - Be proud of his/her language, speak it, advocate it, interact with it, preach it, act it, pray with it, …
  - Use it in the home (parents should speak it to their children, hand it over to the next generation)
  - Interdisciplinary research collaborations
  - Establish Centres of Excellence for Language Technology Research with community-aware content
Proposed Solution

- We should,
  - Disseminate/Integrate language resources into state-of-the-art/ubiquitous tools (e.g. the cell phones)
  - Encourage students in the languages to write research Projects/Thesis, and examine them that will impact on the language communities, not just reviewing theoretical concepts
  - Encourage the Government and Industries to adopt a BT approach to development. E.g., we can make them see the need to fund community-driven projects
Ibibio (ISO: 639-3, Ethnologue IBB)

- Tone language
- Spoken by approx. 4,000,000 speakers
- In Akwa Ibom State
- Southeast coastal region of Nigeria
- West Africa
- From the (new) Benue Congo language family
Empowering Ibibio

- Primary and Secondary Education
  - Ibibio has for the past few years been examined by NECO and this April 2014, WAEC will begin the examination of the Ibibio language.
  - More books for the schools have now been published in all areas - grammar, literature, culture, history, etc. and also used by teachers and students.
  - We hope this will sustain and empower the Ibibio language.
  - On February 21 of every year when the world marks International Mother Language day, essay competitions, quizzes, poetry, drama and cultural events are presented and prizes given to groups that excel.
The Ibibio Language

- Syllable Structure
  - A language is built around its syllable structure
  - The Ibibio syllable structure is simple and comprises a simple onset - vowel or a nasal prefix and a rhyme-consonant (Urua, 2000).
  - The nucleus in the Ibibio syllable include the vowel and syllabic nasal
The Ibibio Language

- **Tense System**

  - The allomorphs generated by each tense are syntactically determined by certain grammatical category in a sentence.
    - mé (present tense) is used with the first and second person pronouns.
    - The use of the past tense marker - màá/ké (non-proximate) is syntactically conditioned.
    - ké is used in a negation sentence and màá in a declarative sentence. For instance, Nkékòppó – i did not hear, mmàákòp – I heard it.
    - The distinction between proximate and non-proximate in past and future tenses is neutralised in a negative sentence. The proximate future tense marker yáà/yáá occurs in a simple declarative sentence and ‘dì’ occurs in negative sentence. For instance, nyáàkà – i will go, ndìkàhà – i will not go.
The Ibibio Language

- **Word formation**
  - Most adverbs in Ibibio language are formed through reduplication. For instance
  - **Noun:**  `inọ` – thief;  `inọ - inọ` – like a thief
  - **Adjective:**  `Idiók` – bad;  `Idiók-Idíók` – very badly
  - Most abstract nouns Ibibio language are formed by adding affixes to a verb root or stems
    - Verb roots       Abstract noun
      - má – love       í-má – love
      - sửá – hate      ù-súá – hatred
  - The sentence structure of Ibibio language is SVO (Subject Verb Object)
The Ibibio Language

- Vowel and Tone System

**Table 1  Ibibio vowel system**

<table>
<thead>
<tr>
<th>Vowel height</th>
<th>Front</th>
<th>Central</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>i, ii</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Mid</td>
<td>e, ee</td>
<td>a, A</td>
<td>0, oo, ə, əə</td>
</tr>
<tr>
<td>Low</td>
<td></td>
<td>a, aa</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2  Ibibio consonant system**

<table>
<thead>
<tr>
<th>Manners of articulation</th>
<th>Stops</th>
<th>voiceless</th>
<th>kp, p, pp</th>
<th>t, tt</th>
<th>k, kk</th>
</tr>
</thead>
<tbody>
<tr>
<td>voiced</td>
<td></td>
<td>b</td>
<td></td>
<td>d</td>
<td></td>
</tr>
<tr>
<td>Non stops</td>
<td></td>
<td>nasal</td>
<td>m, mm</td>
<td>n, nn</td>
<td></td>
</tr>
<tr>
<td>fricative</td>
<td></td>
<td>f</td>
<td>s</td>
<td></td>
<td>h</td>
</tr>
<tr>
<td>tap</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>trill</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>approximant</td>
<td></td>
<td>w</td>
<td></td>
<td>r</td>
<td>j</td>
</tr>
</tbody>
</table>
The Ibibio Language

- Tone System
Empowerment through Lexicographic Materials: Nwed Usem Ibibio (Ibibio Dictionary)

- Was born a decade ago during a WALA project in 2004)
- A conundrum of Ibibio lexical units (words)
- Aimed at bringing to the fore the aims and objectives of the National Policy on Education (Section 4:19a) which stipulates that
  - “the mother tongue or the language of the immediate community should be the language of instruction in the first three years of Primary Schooling. This stipulation is however subject to the availability of teaching material in those languages-orthographies, primers, textbooks etc.”
Nwed Usem Ibibio (Ibibio Dictionary)

- Structure of the Lexical Entries:
  - Each lexical entry in the dictionary has five important information fields.
    - Orthographic information: This is presented in Essien’s Standard Ibibio orthographic characters
    - Phonetic/phonological information: presented as string of IPA symbols in a conventional IPA fonts.
    - Also supra-segmental information like tone is perfectly presented for all lexical entries in this dictionary (this is very rare to the best of my knowledge)
Nwed Usem Ibibio (Ibibio Dictionary)

- Structure of the Lexical Entries:
  - Syntactic Information: presented as Parts of speech (POS) information (e.g. tv - for transitive verbs (fat), n- for nouns, iv- for intransitive verbs, ideo - for ideophones, adv for adverbs etc)
  - Semantic information: represented in free text: phrasal definition is employed as opposed to sentence definition (e.g. in entry for abaakan on page xi; a kind of tree which often grows very large in the bush)
  - Collocation information: represented in example sentences which the lexical entries are used in collocation with other lexical units in the language.
<table>
<thead>
<tr>
<th>Nwed Usem Ibibio (Ibibio Dictionary)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>a, A</strong></td>
</tr>
</tbody>
</table>

| aa  | [áá] interj indicates surprise; áá, áké dámmá idáhá ákè |
| aa  | [áá] interj indicates surprise, áá, ntórò ké ádó |
| aafo | [ááfò] n [Eng half] half penny; ñkpò émi iyohókè ááfò. |
| aafu | [ááfù] n [Eng half] half penny; nkédép bésin ódò ké ááfù. |
| aak | [álák] interj please (emphatic, persuasive, very polite); álák, kúyát ésit. |
| aakpño | [áakpño] n named child (someone named after another, each is akpño to the other); áakpño mòfò? |
| aba edem | [ábà èdèm] n the back (anat) (edem - back, side) ñkpò áfít ányé ke ábà èdèm. |
| aba¹ | [ábà] n cavity, hollow, hole; ékpú aféghe ádák àbà ùsáñi. |
| aba² | [ábà] num forty; ékáéte ódò ówó ìsùá àbà. |
| aba³ | [ábà] adv again, anymore, still (used most commonly with a negative verb); kùnám úyòm àbà. |
| abaaak | [ábaak] adj young, not full grown (abaaak can be used with the name of any domestic animal); èdòní ódò ódò àbááak |
| abaaakañ | [ábaàkañ] n a kind of tree which often grows very large in the bush, its succulent leaves are used for goat fodder, african nutmeg (adams), wild nutmeg (dalziel), pycnanthus kombo (dalziel); ábààkañ mmò ódá ké ádàñá ikót ódò. |
What this book did not cover

- Since it was based on a pioneering work of Kaufman (1981)
- Full sentence definitions
- Domain-specific words (in the area of politics, medicine, etc.)
  - More word variations and examples
  - We intend to crowdsource data by transforming the e-copy version of the dictionary into a Web version
Empowerment through NLP/LT Tools: Ibibio Speech Resources

- Ibibio Speech Synthesis System
  - Mid 2012, the first Statistical Parametric speech synthesis system for Ibibio was born (Ekpenyong, Watts, Urua, King and Yamagishi, 2014)
  - The realization of this groundbreaking research began in a Speech Technology Research Scholarship that paralleled a WALA Project (Gibbon, Ahuoa, Gbery, Urua and Ekpenyong, 2004), about a decade ago
  - Our confidence was based on the assumption that nothing is impossible?
Empowerment through NLP/LT Tools: Ibibio Speech Resources

- **Ibibio Speech Synthesis System**
  - So we started the journey by adapting a Spanish front end (Festival synthesis System)
  - But, as we went deeper into the project we discovered a lot of missing links
    - System robustness
    - Replicability and adaptability of the speech resources
    - Reliability and reverse engineering
    - Supervised or unsupervised or hybrid?
    - Tone harmony
    - What next?
HMM context modeling

Fig. 2. The training procedure. Close copy synthesis:
#a-ke-di-dio0-NO dem#o-do#a-di-jIk
#1-02-04-0222-01-010#1-02#1-01-010

Speech systems performance

Naturalness

Intelligibility
Empowerment through NLP/LT Tools: Ibibio Speech Resources

- **Generic Tone Marking System**
  - We are exploring robust tone pattern modeling
  - in a data driven tone marking system