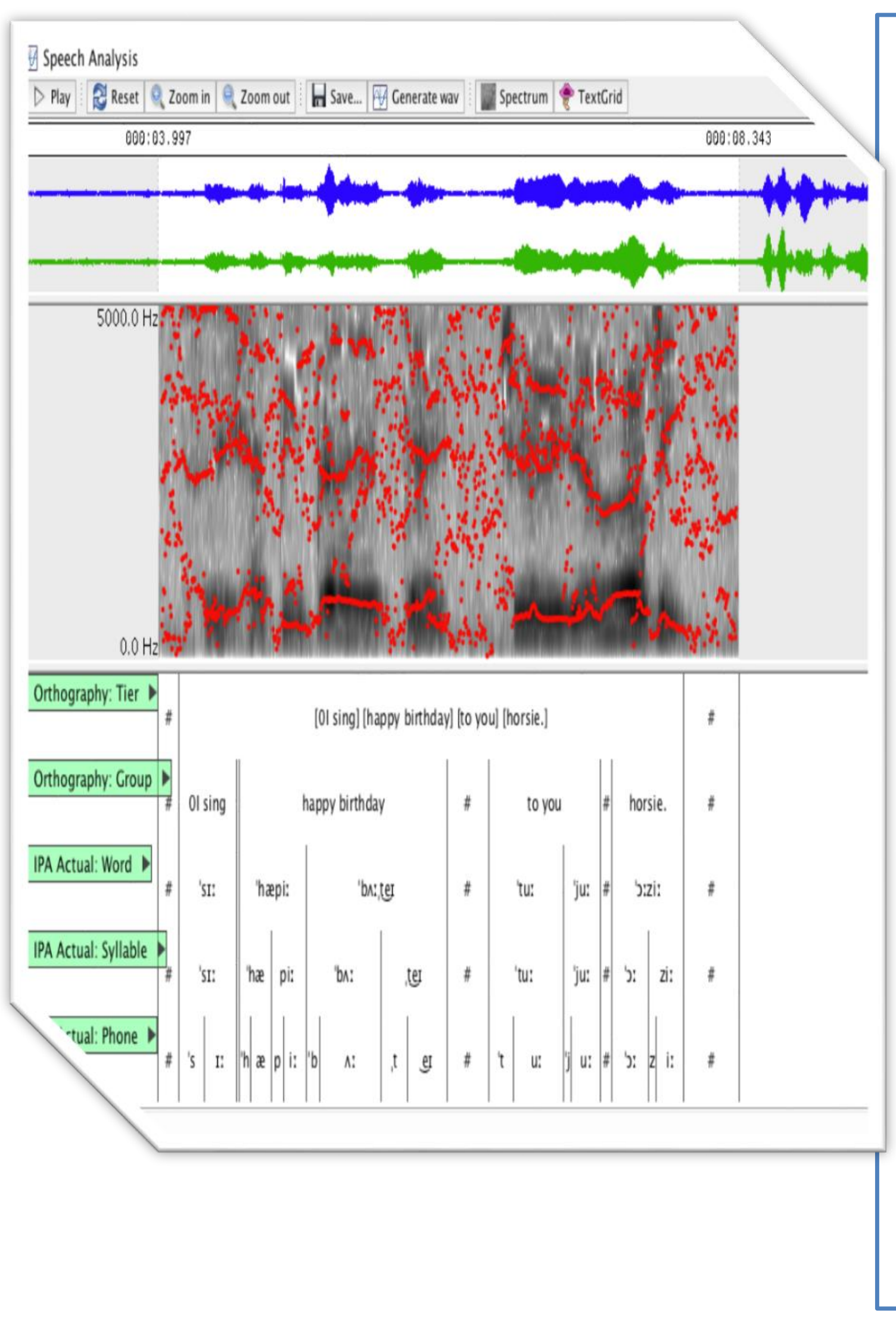
SCAnnAL – An Automatic Speech Corpus Annotator for African Speech Corpora

Moses Ekpenyong[1] – Eno-Abasi Urua[2] – Aniefon Akpan[2]

[1]Department of Computer Science; [2]Department of Linguistics and Nigerian Languages

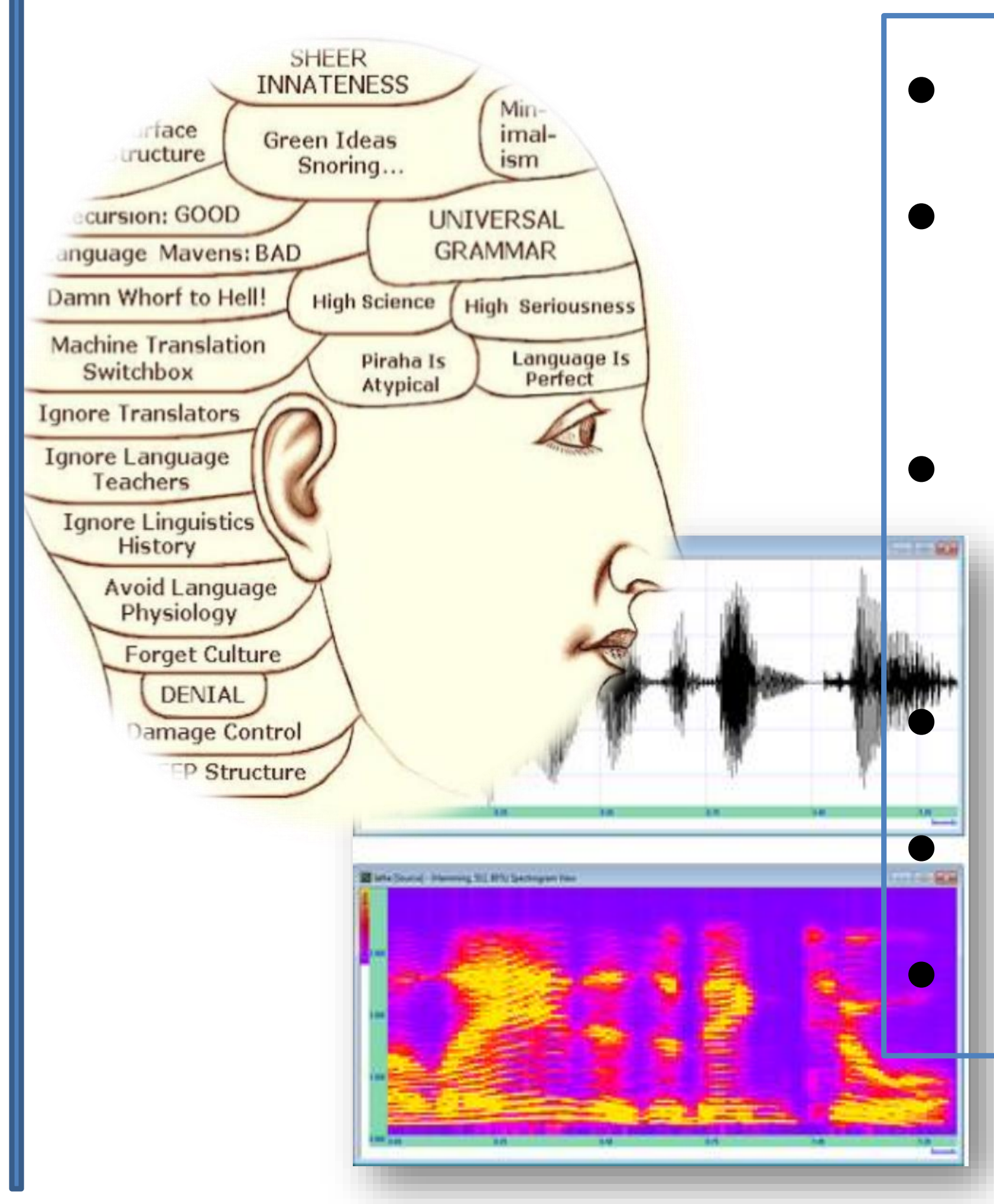
(mosesekpenyong@{uniuyo.edu.ng, gmail.com}; enoabasi@gmail.com; daniels.mmefon@gmail.com)

What is The Issue?



- Thousands of annotated speech corpora exist worldwide
- The demand for richly annotated corpora is fast growing
- The process accompanying corpora annotation has slowed research progress for African languages
- Current annotation Toolkits do not satisfy the challenges African speech systems present.

Audience:



- Linguists
- Fieldwork/language experts
- Computational linguists
- Historians
- Speech Engineers
- Speech Technologists

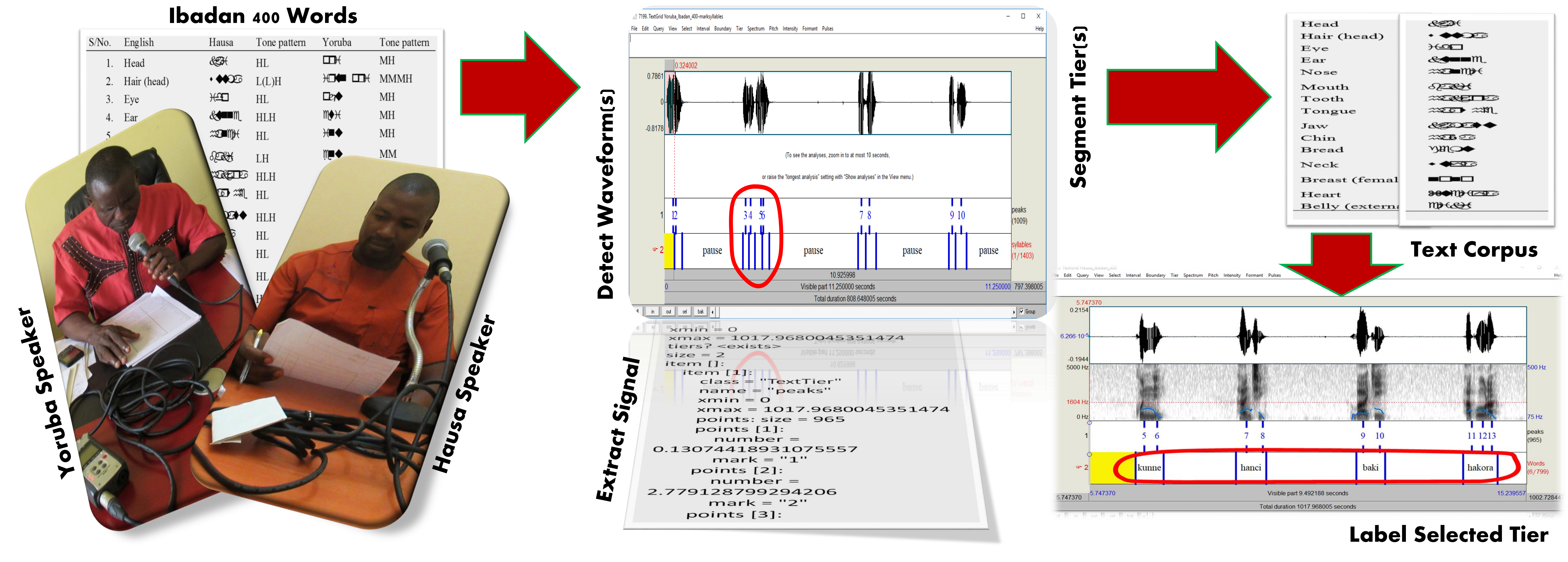
Our Goals:

- Examine current annotation Toolkits and identify their limitations
- Study the peculiarities of African tone languages
- Automate the annotation process using Signal Processing and NLP
- Adapt automated process to African languages
- Evaluate the annotator for precision

Our approach:

- Using Signal Processing, detect the respective speech waveforms
- Segment specified tier(s)
- Accept corresponding transcriptions
- Perform NLP to pre-annotate the transcriptions (e.g., syllabification, ...)
- Render transcriptions (phonemes; words; syllables; sentences; etc.), to respective segments using NLP
- Label segments

SCAnnAL Methodological Workflow:



Ibadan 400 Words

S/No.	English	Hausa	Tone pattern	Yoruba	Tone pattern
1.	Head	[illegible]	HL	[illegible]	MH
2.	Hair (head)	[illegible]	L(L)H	[illegible]	MMMH
3.	Eye	[illegible]	HL	[illegible]	MH
4.	Ear	[illegible]	HLH	[illegible]	MH
5.		[illegible]	HL	[illegible]	MH
		[illegible]	LH	[illegible]	MM
		[illegible]	HLH		
		[illegible]	HL		
		[illegible]	HLH		
			HL		
			HL		
			HL		

Yoruba Speaker

Hausa Speaker

Detect Waveform(s)

Segment Tier(s)

Head	[illegible]
Hair (head)	[illegible]
Eye	[illegible]
Ear	[illegible]
Nose	[illegible]
Mouth	[illegible]
Tooth	[illegible]
Tongue	[illegible]
Jaw	[illegible]
Chin	[illegible]
Bread	[illegible]
Neck	[illegible]
Breast (femal	[illegible]
Heart	[illegible]
Belly (externa	[illegible]

Text Corpus

Extract Signal

```
   xmin = 0
   xmax = 1017.9680045351474
   tiers? <exists>
  size = 2
  item []:
      item [1]:
          class = "TextTier"
          name = "peaks"
          xmin = 0
          xmax = 1017.9680045351474
          points: size = 965
          points [1]:
              number =
0.130744189310755557
              mark = "1"
          points [2]:
              number =
2.7791287992942O6
              mark = "2"
          points [3]:
```

Label Selected Tier