Acoustic and Perceptual Analyses of Sesotho Tone

Lehlohonolo Mohasi

Supervised by: Prof. Thomas Niesler
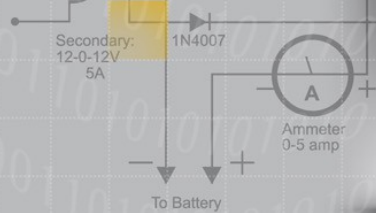**Prof. Hansjoerg Mixdorff**

UNIVERSITEIT
STELLENBOSCH
UNIVERSITY

Overview

* Research Objective
* The Sesotho Language
* The Fujisaki Model
* Part 1 – Acoustic Analysis
  - Aim
  - Methodology
  - Observations
  - Results
  - Summary & Conclusions
* Part 2 – Perceptual Analysis
  - Aim
  - Methodology
  - Results
  - Summary & Conclusions

Research objective

* **Speech synthesis in Sesotho**
  - **Intelligibility & naturalness**

* **Need to understand tone**
  - **Interrelationship between tone and general intonation in Sesotho is poorly understood**
  - **Technologically not addressed**

* **Tone not marked in orthography**
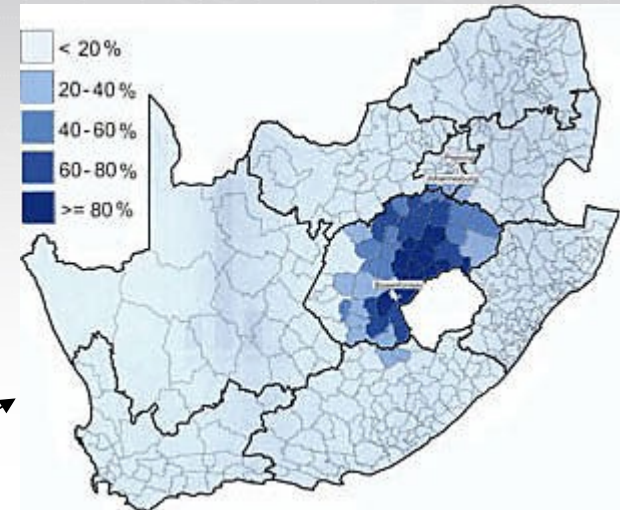
* **Very little published**

The Sesotho Language

* **Tonal Bantu language official in Lesotho (LS) and South Africa (SA)**

* **Different orthographies, same pronunciation**

* **Exotic phonemes include clicks, e.g.:**

  – Moqebelo (Saturday)

  – senqanqana (a frog)

  – qhaqhang-qhang (a fight)

The Sesotho Language

**Home language to ~ 7.9% of the SA population and ~ 99% of Lesotho population**

Area of SA where Sesotho language is dominant.
Langauges of South Africa, South African language copyright © South African tourism

The Sesotho Language

* Sesotho is a language with 2 contrasting tones: high (H) and low (L)

* To quantify tonal alignment in Sesotho, the Fujisaki tool was used

  – F0 contours parametrized

  – Initially developed for German language

  – Applied to other Asian tonal languages
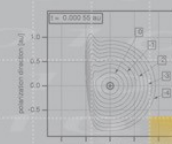
* Syllabic tones shown by tone/accent commands
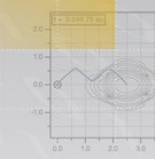
The Fujisaki model

* Decomposes a given log F0 contour into:
  - a base frequency (*Fb*)
  - a phrase component, capturing slower changes in the F0 contour as associated with intonation phrases
  - an accent component that reflects faster changes of F0 associated with accents and boundary tones

* The phrase and accent components can be interpreted as
  - smooth responses of the model to impulse-wise phrase commands
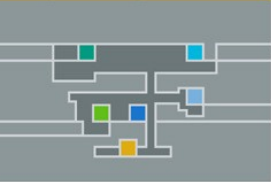 - box-shaped accent commands
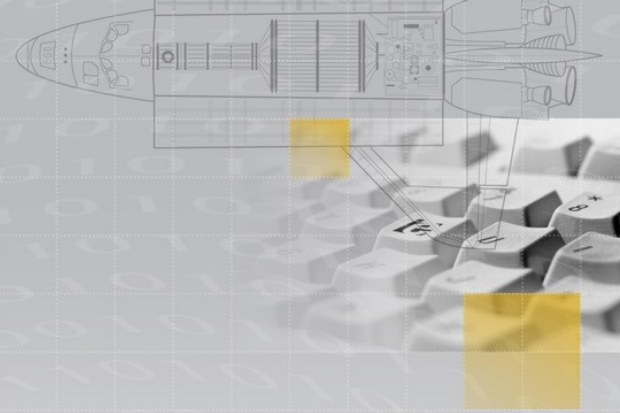
The Fujisaki model
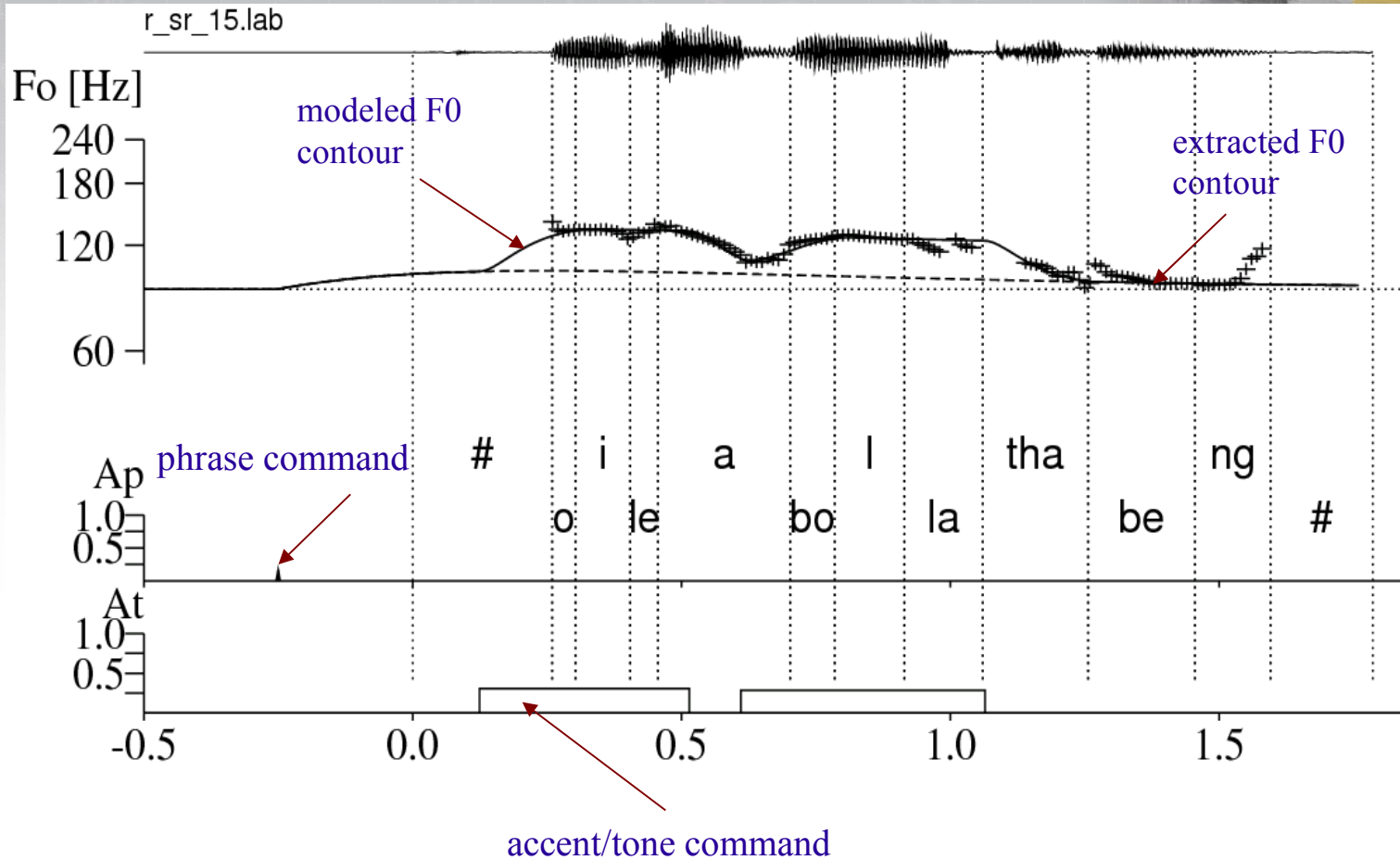
Part 1

Acoustic Analysis

Aim

* Carry out acoustic investigation into the realization of tone in Sesotho using the Fujisaki tool.

* All tone languages examined so far show tone commands of both positive and negative polarity.

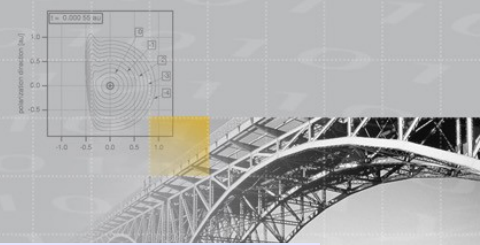* Examine whether this is also the case for Sesotho.

Creating a corpus

- Tonal minimal pairs + questions

- Own examples + from literature

- Annotation at word and syllable levels

- Syllable tones also marked

- 2 strategies employed: reading and repeating

Corpus extract

Tone pattern	Sample sentence	English meaning
HH	O ile a e **hlola**.	He conquered it.
LL	O ile a e **hlola**.	He created it.
HH	Motsoala o rata ho **seba**.	My cousin likes to gossip.
LL	Motsoala o rata ho **seba**.	My cousin likes doing mischief.
HH	Oa **tena**.	She is putting on clothes.
LL	Oa **tena**.	She is annoying.
HLH	O ile a **bolla** thabeng.	He was circumcised in the mountains.
LLL	O ile a **bolla** thabeng.	He (his body) decayed in the mountains.

Observations during corpus compilation

Auditory analysis: Tone recognition error rate

–   **Reading 23.6%**

–   **Repeating 7.4%**

Fujisaki analysis

–   **Revealed minimal pairs had both vowel difference and tonal difference**
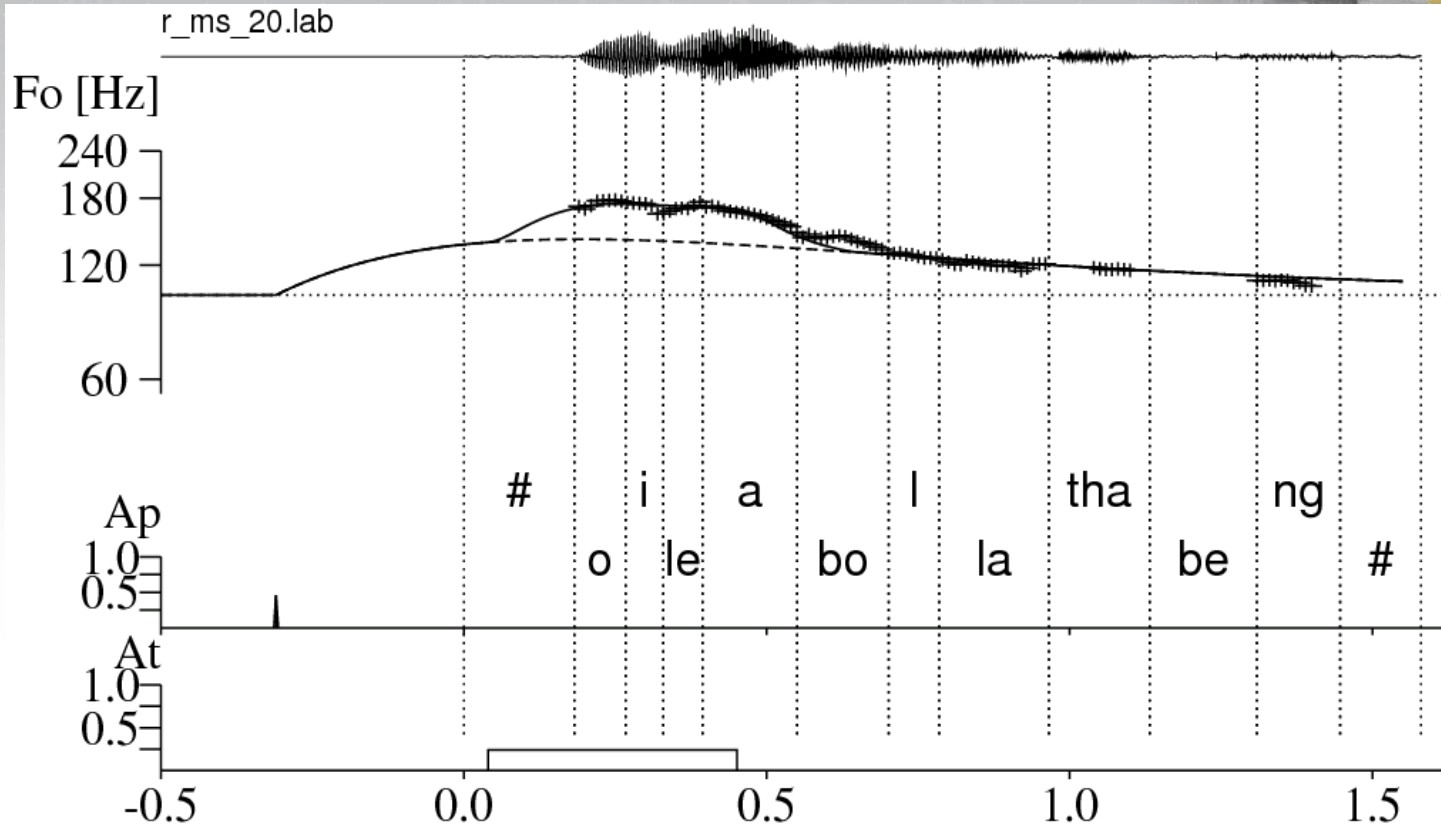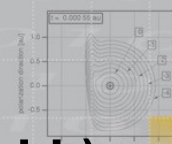
–   **Literature not consistent and sometimes wrong**

Examples of inconsistencies with literature

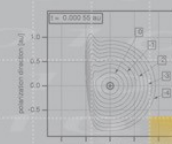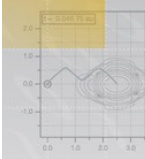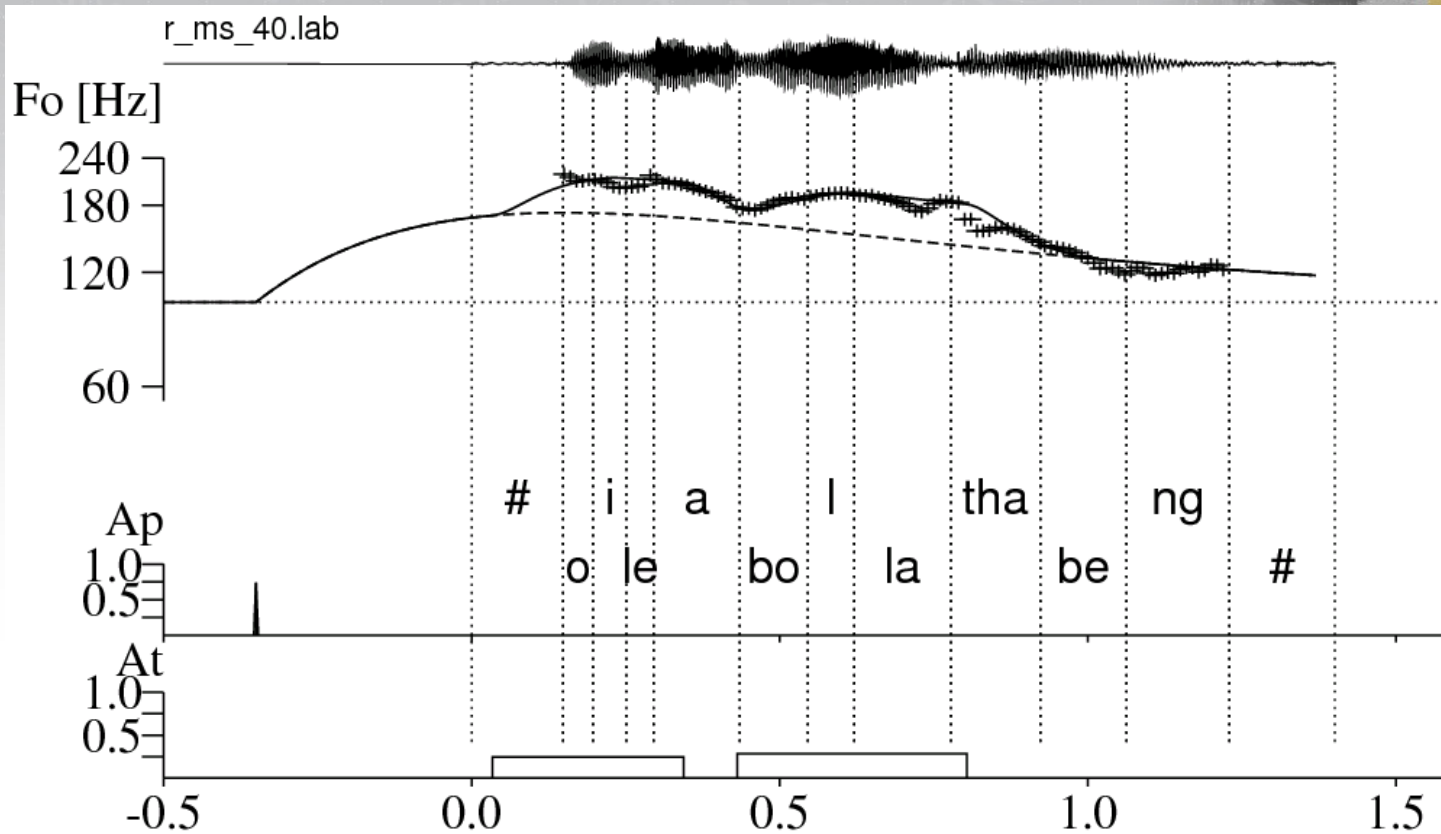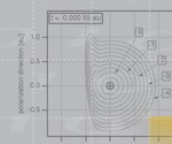| word	position	translation	hyp. tone	obs. tone	vowel
hlola	final	create	LL	HL	[O]
		conquer	HH	HL	[o]
seba	final	gossip	HH	HL	[e]
		do mischief	LL	LL	[e]
pota	medial	come over	LL	HH	[O]
		talk crap	HH	HH	[o]
tena	final	get dressed	HH	HL	[e]
		annoy	LL	LL	[e]

Fujisaki analysis: bolla - decay (LLL)

r_ms_20

Fujisaki analysis: bolla – circumcise (HLH)
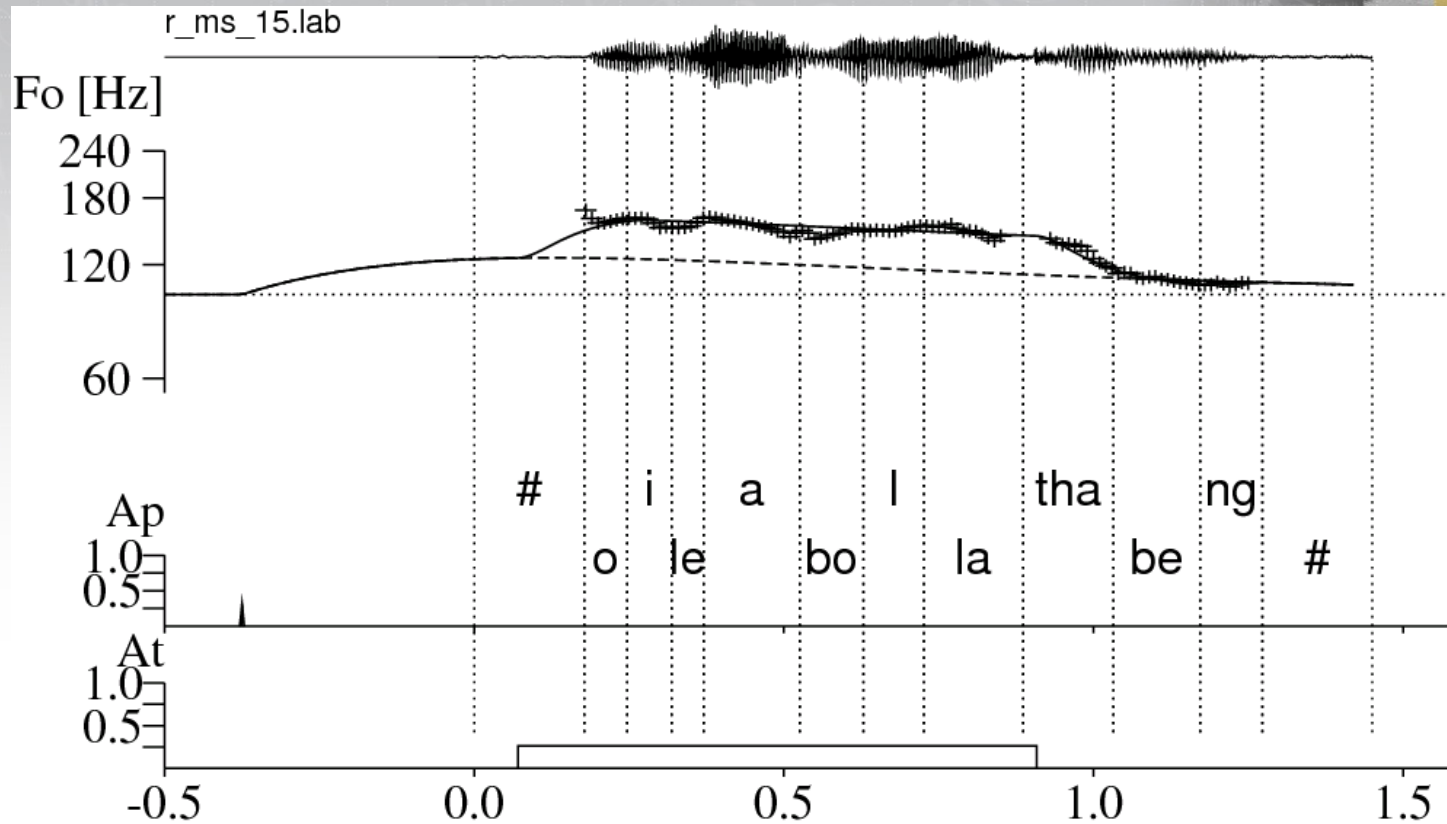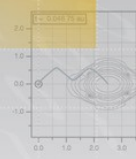
Fujisaki analysis: bolla – circumcise (HLH)   (a question)

r_ms_40   r_ms_15

Summary and Conclusions (Part 1)

*   The material examined so far only presents a first snapshot on tonal realizations in Sesotho.

*   The tonal organization of Sesotho is different from that found in Asian tone languages where every syllable is assigned a specific tonal target.

*   Tones specified in existing literature do not always correspond to observations.

*   Sometimes tone differences are associated with vowel differences.
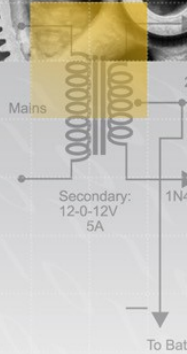
Summary and Conclusions (Part 1)

*   It was found that in Sesotho only high tone syllables are associated with tone commands and other syllables are either transient or their F0 follows the phrasal contour.
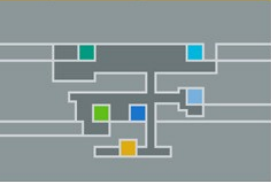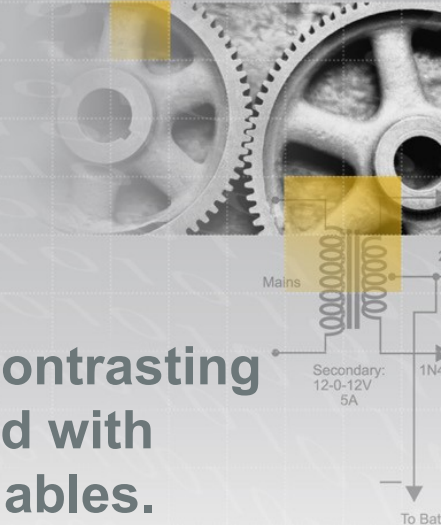
Part 2

Perceptual Analysis
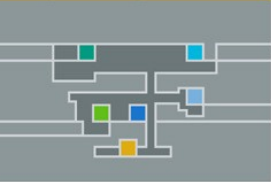
In Part 1, it was observed that some of the contrasting words in the minimal pairs examined differed with respect to the vowel quality of their first syllables.

Some even showed only vowel differences, and no tonal differences.

Aim

* **Investigate to what extent tonal perception interacts with vowel perception to facilitate word identification**

* **Investigate to what extent pitch, speech rate and the length of the penultimate syllable affect the perception of statements vs. questions**

* **Use prosodically modified stimuli for perceptual tests**
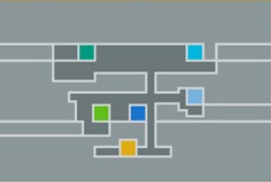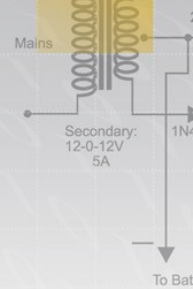
Methodology

* Used a subset of the corpus of minimal pairs from the acoustic analysis experiment.

* Examples chosen had tonal and vowel differences.

* Examples also included question/statement pairs with identical orthography.

* Stimuli modified using Praat and the FujiParaEditor plug-in and compiled into a perceptual experiment.
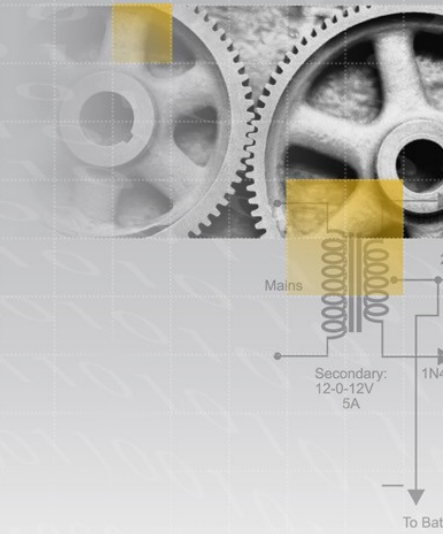
* Included repeats to test for consistency.
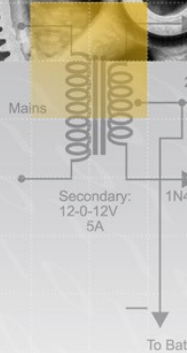
Perceptual tests

Performed at:

– **Stellenbosch University (SU)**

– **National University of Lesotho (NUL)**

|	SU	NUL	Total
Male	12	5	17
Female	3	8	11
Total	15	13	**28**

Perceptual tests

* **Results from the two groups at SU and NUL were evaluated separately.**

    – **Correlation of group result means was found to be 0.96, these were pooled for subsequent analysis.**

* **Correlation between results at first and second presentation was 0.90, hence their average was taken for the repeated stimuli.**
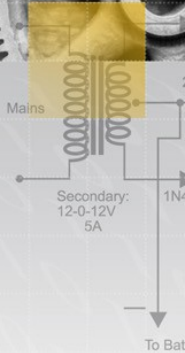
Word identification stimuli

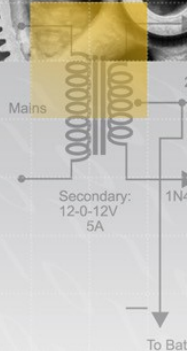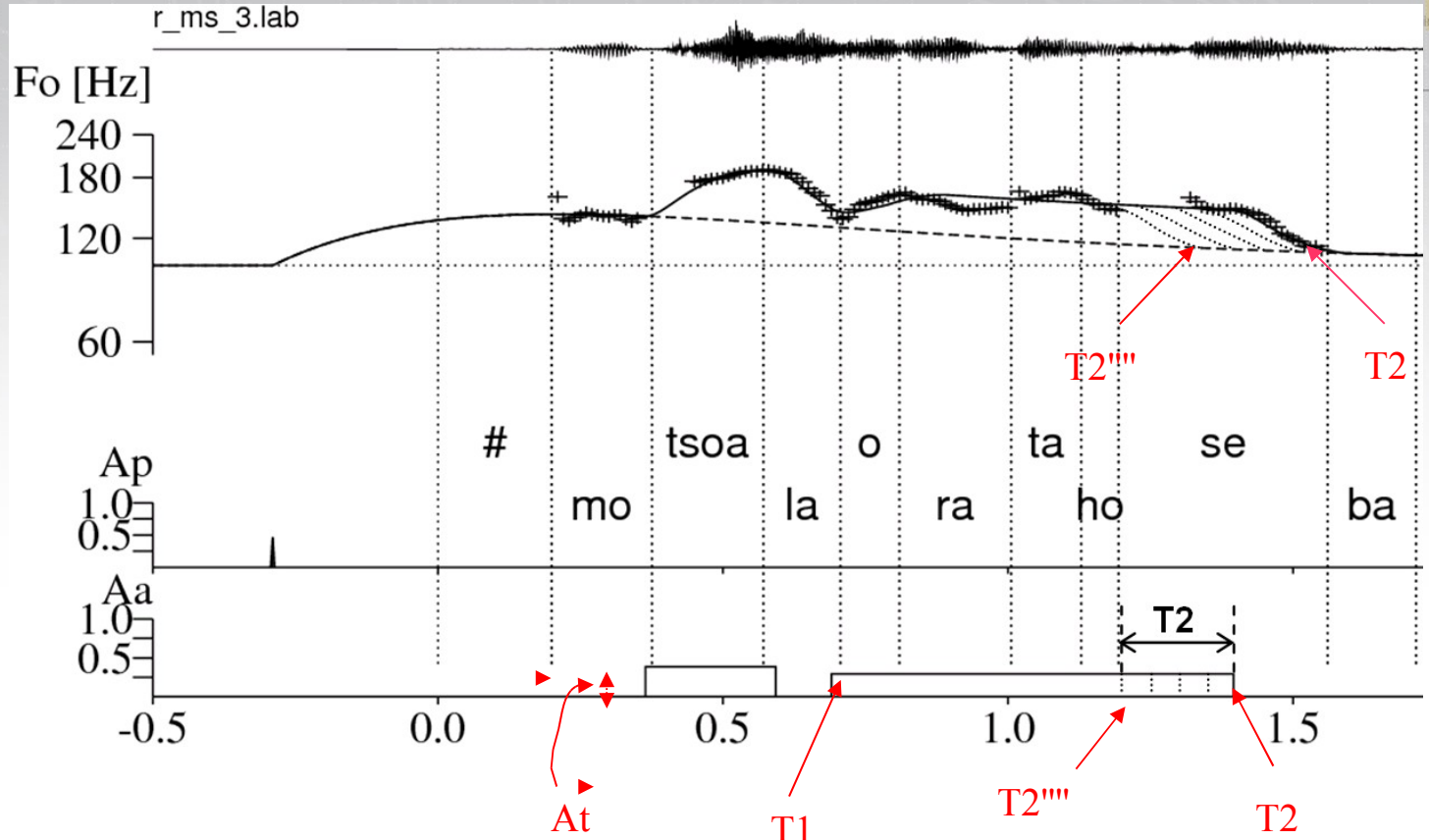| word	translation	vowel	tone
lehata	skull	[e]	HHH
	liar	[e]	LLL
seba	gossip	[e]	HL
	do mischief	[e]	LL
bolla	circumcised	[o]	HHH
	decayed	[O]	LLL
ts'ela	crossed	[e]	HL
	poured	[E]	HL
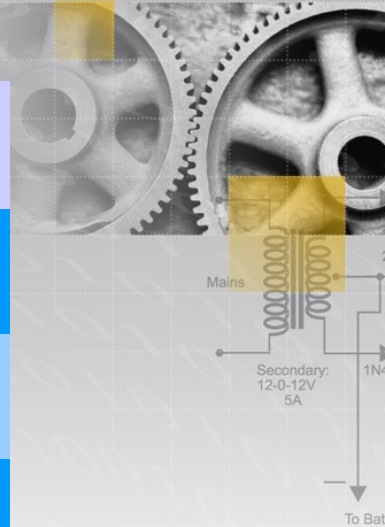
Word identification

Aims

2 points were of main interest:

* What is the minimum **duration of a tone command** associated with a high tone syllable?

* Does **reducing/increasing the tone command amplitude** on a high/low tone syllable lead to perception of the low/high tone partner – even without modifying the vowel quality?

Word identification
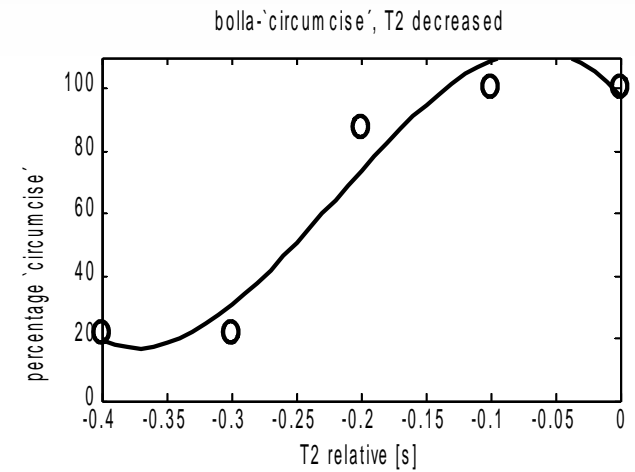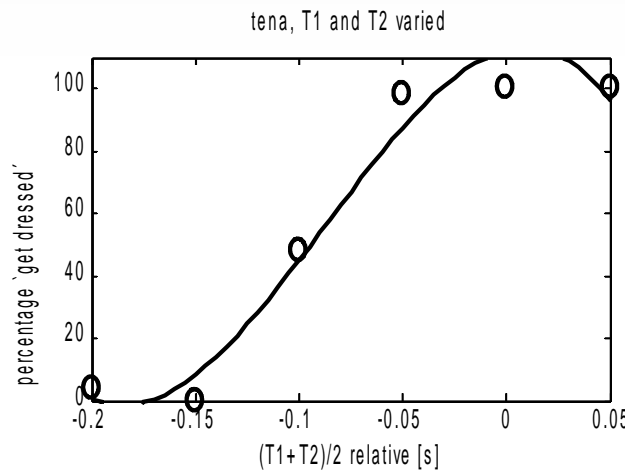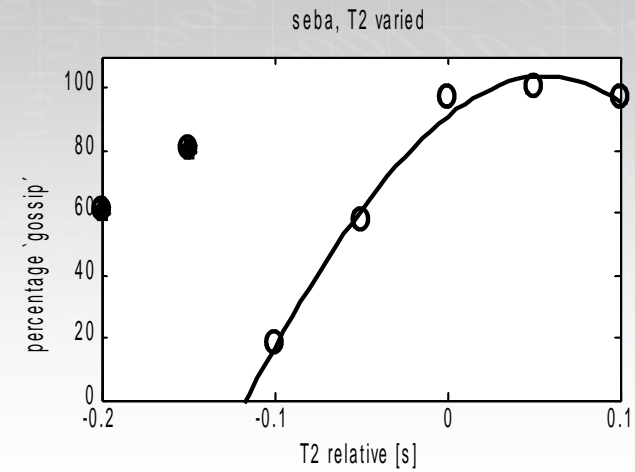
Example: seba - gossip; variation of T2

Word identification - Modifications evaluated

word	translation	vowel	tone	modification
lehata	skull	[e]	HHH	Increase of T1
	liar	[e]	LLL	T1 later
seba	gossip	[e]	HL	Variation of T2
	do mischief	[e]	LL	T2 earlier
bolla	was circumcised	[o]	HHH	Reduction of T2, reduction of At
	decayed	[O]	LLL	Increase of At

Word identification

Results
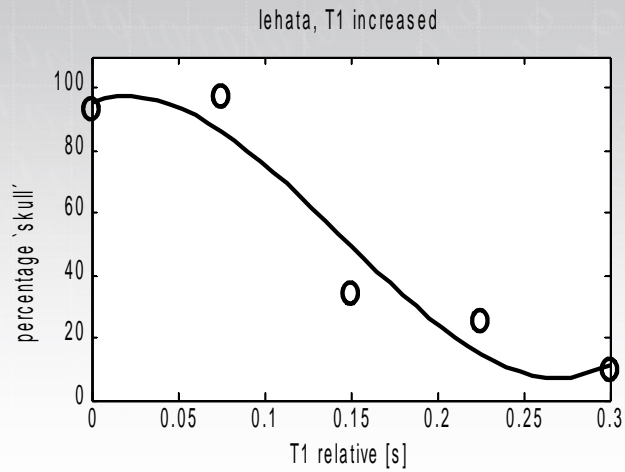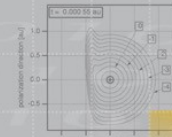
Question vs. Statement
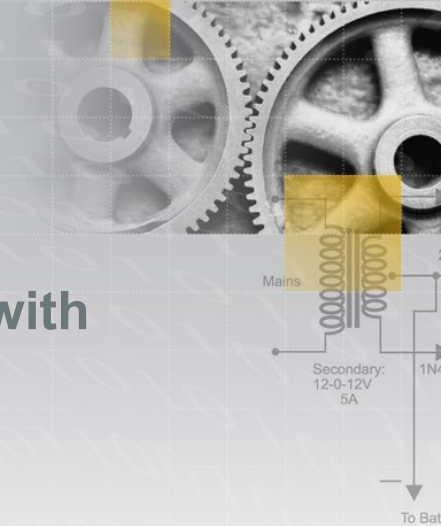
Aims

Two utterances of statements are modified with respect to the following parameters:

- The phrase command magnitude (Ap)

- Speech rate (Sp)

- Shortening of penultimate syllable (SP)

The perceptual experiment tested whether these modifications could change a statement to a question.

Question vs. Statement

Results

Effectiveness of each type of modification:

- phrase command magnitude (Ap): correlataion = -0.25

- speech rate (Sp): correlation = -0.36

- presence/absence of penultimate shortening (PS): correlation = -0.83

Correlations are between each modification and the successful perception of a question.
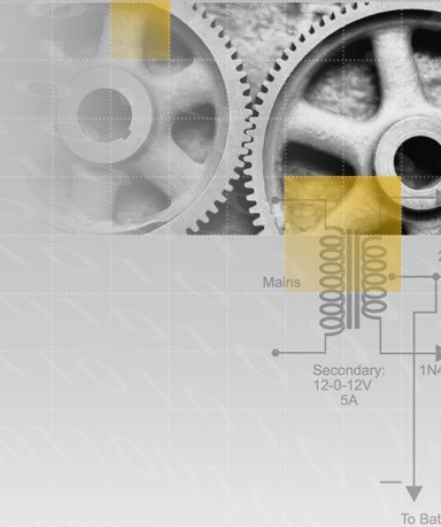
Questions vs. Statement

Results

Stimuli	Ap	Sp	PS	Outcome – Identified as ...
Unmodifed	unchanged	unchanged	unchanged	statement by 97.6%
Modified	increased to max.	unchanged	unchanged	statement by 84.8%
Modified	unchanged	increased	unchanged	statement by 86.9%
Modified	unchanged	unchanged	shortened	statement by 66.1%
Modified	highest	increased to max.	present	question by 89.3%

Summary & Conclusions

* With a few exceptions, results regarding word identification are in line with our expectations.

* Reduction in the amplitude as well as in the duration of the tone command for high tone stimuli convert them into their low tone counterparts.
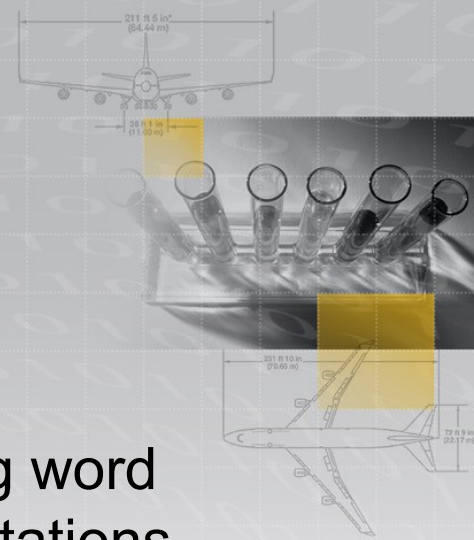
Summary & Conclusions

*   Increasing the amplitude of the tone command for a low tone word has the opposite effect.

*   F0 modifications even seem to override vowel differences between words, as was shown for 'bolla'.

Summary & Conclusions

*   Shortening of the penultimate syllable, increasing the speech rate, and increasing the phrase command magnitude all make it more probable that an utterance is judged to be a question.

*   The shortening of the penultimate had the strongest impact among the modifications considered.
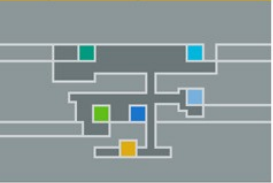
Thank You!!!

Questions??

Comments??

Suggestions??

lmohasi@sun.ac.za