Parsing Grammatical Tone Using FLEX

Jonas Lau

SOAS, University of London
WHAT I WILL TALK ABOUT TODAY

PROBLEM:
I want to use FLEx, but I am working on a language featuring grammatical tone

QUESTION:
How can inflectional categories expressed by tone be glossed automatically?
WHAT I WILL TALK ABOUT TODAY

Ikaan
Tone
FLEEx
IKAAN

INTRODUCTION

Minority language spoken in Southern Nigeria
Around 10,000 speakers
Benue-Congo branch of Niger-Congo phylum
PRIMARY DATA
IKAAN

PRIMARY DATA

Deposit 0035 - Ikaan and related dialects of Ukaan

Sophie Salffner
**INTRODUCTION**

**Grammatical tone**

The function of tone is not limited to distinguishing two words, but can also distinguish two grammatical categories. (Hyman 2001: 1372)

**Tonal melodies**

By defining tonal melodies for specific word classes, multiple correlating sequences of surface tones can be explained and grouped together. (Gussenhoven 2004: 30)
• Verbal melodies encode TAM
• Melodies of the negated paradigm differ from the ones of the affirmative paradigm

(1) dʒòbár ìjóg
dʒɔ-LH-bár ì-jo = g
1SG-NEG.HAB-peel I4-yam = NEG
‘I do not peel yam.’ (neg.020)
(2) NFUT FUT CONT HAB

Affirmative d₃E LH d₃a F:H d₃E HD:HH d₃O F:H
LHL F:HH HD:LHHH
LHLL F:HHH HD:LHHHH

NEG Class 1 d₃E R:DH d₃O R:H d₃O LH
R:DHH R:HL LHL
R:DHHH R:HLL LHHH

NEG Class 2 d₃E R:H d₃O R:DH
R:HL d₃O LHH
R:HLL d₃O LHHH
<table>
<thead>
<tr>
<th>(2)</th>
<th>NFUT</th>
<th>FUT</th>
<th>CONT</th>
<th>HAB</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Affirmative</strong></td>
<td>d₃E LH</td>
<td>d₃a F:H</td>
<td>d₃E HD:HH</td>
<td>d₃O F:H</td>
</tr>
<tr>
<td></td>
<td>LHL</td>
<td>F:HH</td>
<td>HD:LHHH</td>
<td>F:HH</td>
</tr>
<tr>
<td></td>
<td>LHLL</td>
<td>F:HHH</td>
<td>HD:LHHHH</td>
<td>F:HHH</td>
</tr>
<tr>
<td><strong>NEG Class 1</strong></td>
<td>d₃E R:DH</td>
<td>d₃O R:H</td>
<td>R:DH</td>
<td>d₃O LH</td>
</tr>
<tr>
<td></td>
<td>R:DHH</td>
<td>R:HL</td>
<td>HD:HH</td>
<td>LHL</td>
</tr>
<tr>
<td></td>
<td>R:DHHH</td>
<td>R:HLL</td>
<td>LHH</td>
<td>LHLL</td>
</tr>
<tr>
<td><strong>NEG Class 2</strong></td>
<td>d₃E R:H</td>
<td>d₃O R:DH</td>
<td>R:DH</td>
<td>d₃O LH</td>
</tr>
<tr>
<td></td>
<td>R:HL</td>
<td>R:DHHH</td>
<td>R:DHHH</td>
<td>LHH</td>
</tr>
<tr>
<td></td>
<td>R:HLL</td>
<td></td>
<td></td>
<td>LHLL</td>
</tr>
</tbody>
</table>

**Ikaan**

**Tone**

**FLEX**
McGill (2007):
• Proposes a method for Toolbox to parse grammatical tone.
• Tones are parsed as infixes and can be read together as one morpheme.
• Not applicable to FLEEx

(3) \tx  \udukwa
\mb  \u- dukwa- H L
\ge  3SG- go  RLS
\ft  ‘He went’
**FLEx**

**XAmple**
- Default parser
- Linear parsing: left to right

**HermitCrab**
- “Item and process approach“ (Black 2014)
- Outside to inside
- Allows Phonological Rules and Affix Process Rules (APRs)
APRs model processes to form affixes

Advantages:

- Material of the stem can be reused (reduplication)
- Different affixation strategies can be combined
- Tonal melodies can be modelled as a combination of infixes carrying one gloss
(5)

RHL₁ v : NFUT NFUT.NEG

<table>
<thead>
<tr>
<th>Allomorphs</th>
<th>Ikaa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affix Allomorph (Process: Ikaa)</td>
<td></td>
</tr>
<tr>
<td>Morph Type</td>
<td></td>
</tr>
<tr>
<td>Affix Process Rule</td>
<td></td>
</tr>
<tr>
<td>prefix</td>
<td></td>
</tr>
<tr>
<td>Input</td>
<td>[C] [V] [C] [V] [C] [V] X</td>
</tr>
<tr>
<td>Index</td>
<td>1 2 3 4 5 6 7</td>
</tr>
<tr>
<td>Result</td>
<td>R:HLL</td>
</tr>
</tbody>
</table>

| Affix Allomorph (Process: Ikaa) | 
| Morph Type | 
| Affix Process Rule | 
| prefix | 
| Input | [C] [V] [C] [V] X | 
| Index | 1 2 3 4 5 | 
| Result | R:HL | 

| Affix Allomorph (Process: Ikaa) | 
| Morph Type | 
| Affix Process Rule | 
| prefix | 
| Input | [C] [V] X | 
| Index | 1 2 3 | 
| Result | R:H |
(6)  dʒěːʃénò́g
     đʒ e `ʃ e nd o g
(6) dʒěːʃénòg

dʒe ː 

R:HL

dʒE-

1SG
(6)  

\[
\begin{align*}
\text{džeʃénòg} & \\
\text{dʒe} & : \text{ʃe} \text{ŋ o} & \text{g} \\
\text{dʒE-} & -\text{g} \\
\text{1SG} & \text{NEG}
\end{align*}
\]
(6)  dʒeːʃénòg

dʒeːʃénòg

dʒE-            R:HL-              -g

1SG           NFUT.NEG           NEG
(6) dʒěːʃénòg

dʒe ː̌ ʃenog

dʒE-RːHL-ʃeno-g

1SG NFUT.NEG play NEG
(6)  

\[ dʒěːʃénòg \]

\[
\begin{align*}
\text{dʒe} & \quad \ddash \quad \ddash \quad \ddash \\
\text{dʒE-} & \quad \text{R:HL-} \\
1\text{SG} & \quad \text{NFUT.NEG} \\
\end{align*}
\]

\text{I did not play'}
HOW TO

- Plan representation of tone in your transcriptions/orthography
  - Form paradigms
- Group several paradigms together by forming templates to regulate cooccurrences
- Create APRs with allomorphs for words of different syllable structures
- Pay attention to the order of the rules (longer affixes come first!)
- Test with regular paradigms
CONCLUSION

Best Practice Guidelines
Four Best Practice Guidelines

• Try to keep a BALANCE between descriptive accuracy and workarounds
• Pay attention to the COMPLEXITY of the language structure, but try to minimize COMPLEXITY when modelling the structure for a parser
• Try to avoid UNDERSPECIFICATION
• Pay attention to the ORDER of rules
Literature


