HORNMORPHO 2.5 Quick Reference

Michael Gasser
Indiana University, School of Informatics and Computing
gasser@cs.indiana.edu
11 July, 2012

Installation

1. Uncompress the file that you downloaded. This will yield a directory (folder) called `HornMorpho-2.5`, which contains all of the files that you need to run HORNMORPHO.
2. Go to the `HornMorpho-2.5` directory (folder), and enter the following, making sure that you are running Python 3.0, 3.1, or 3.2.

```
python setup.py install
```

Use

STARTING THE PROGRAM

Start up the Python interpreter, again making sure that you are running at least Python 3.0, and type the following to load the program.

```
import l3
```

FUNCTIONS

Options for each function are shown with their default values.

anal(*language*, *word*)
Options: roman=False, root=True, gram=True, citation=True, raw=False, nbest=100 [Amharic only]

Performs morphological analysis of the word. For ambiguous words returns the first `nbest` analyses. For Amharic only, analyses are ordered by their estimated frequency.

```
>>> l3.anal('ti', 'ናብ')
word: ናብ
```

```
>>> l3.anal('ti', 'ፔፕሲ')
?word: ፔፕሲ
```

```
>>> l3.anal('am', 'የማያስፈልጋትስ')
word: የማያስፈልጋትስ
```

```
POS: verb, root: <fl_g>, citation: አስፈለገ
 subject: 3, sing, masc
 object: 3, sing, fem
 grammar: imperfective, causative, relative, definite, negative
 conjunctive suffix: s
>>> l3.anal('om', 'afeeramaniiru')
word: afeeramaniiru
POS: verb, root: <afeer>, citation: afeeramuu
 subject: 3, plur
 derivation:  passive
 TAM: perfect
>>> l3.anal('ti', 'ብዘጋጥመና')
word: ብዘጋጥመና
POS: verb, root: <gTm>, citation: ኣጋጠመ
 subject: 3, sing, masc
 object: 1, plur
 grammar: imperfective, reciprocal, transitive, relative
 preposition: bI
>>> l3.anal('am', 'አይደለችም')
word: አይደለችም
POS: copula, root: <ne>
 subj: 3, sing, fem
 negative
>>> l3.anal('ti', 'ዘየብለይ')
word: ዘየብለይ
POS: verb, root: <al_e>, citation: ኣሎ
 subject: 3, sing, masc
 object: 1, sing
 grammar: present, relative, negative
>>> l3.anal('om', 'dubbanne')
word: dubbanne
POS: verb, root: <dubbadh>, citation: dubbachuu
 TAM: past, negative
POS: verb, root: <dubbadh>, citation: dubbachuu
 subject: 1, plur
 TAM: past
>>> l3.anal('am', 'lezemedocacnm', roman=True)
word: lezemedocacnm
POS: noun, stem: zemed
 possessor: 1, plur
 grammar: plural
 preposition: le, conjunctive suffix: m
>>> l3.anal('am', 'ቢያስጨንቁአቸው', root=False, gram=False)
word: ቢያስጨንቁአቸው
POS: verb, citation: አስጨነቀ
```

```
>>> l3.anal('am', 'ለዘመዶቻችንም', raw=True)
[('zemed', [-acc, cnj='m', der=[-ass], -dis, +plr, pos='n',
poss=[+expl, +p1, -p2, +plr], pp='le', rl=[-acc, +p], v=None])]

>>> l3.anal('am', 'ይመጣሉ')
word: ይመጣሉ
POS: verb, root: <mT'>, citation: መጣ
 subject: 3, plur
 grammar: imperfective, aux:alle
POS: verb, root: <mTT>, citation: መጠጠ
 subject: 3, plur
 grammar: imperfective, aux:alle
POS: verb, root: <mT'>, citation: ተመጣ
 subject: 3, plur
 grammar: imperfective, aux:alle, passive

>>> l3.anal('am', 'ይመጣሉ', nbest=1)
word: ይመጣሉ
POS: verb, root: <mT'>, citation: መጣ
 subject: 3, plur
 grammar: imperfective, aux:alle
```

anal_file(*language*, *input_file*, *output_file*)
Options: root=True, gram=True, citation=True, raw=False

Runs anal on the words in a file.

```
>>> l3.anal_file('am', 'l3/languages/am/data/ag.txt',
    'l3/languages/am/data/ag_out.txt')
Analyzing words in l3/languages/am/data/ag.txt
Writing to l3/languages/am/data/ag_out.txt
```

seg(*language*, *word*) [Amharic and Oromo verbs and Oromo nouns only]
Options: roman=False, gram=True, raw=False

Performs morphological segmentation on the word. Morphemes are separated by '-'; stems/roots appear within '{}'.

```
>>> l3.seg('am', 'ሲያጭበረብሩን')
ሲያጭበረብሩን:
s(cnj1)-y(sb=3sm|3p)-{Cbrbr+a12e3e4_5}(imprf,trans)-u(sb=2p|3p)-n(ob=1p)

>>> l3.seg('om', 'afeeramaniiru', gram=True)
word: afeeramaniiru
POS: verb, segmentation: {afeer-am}-an-r-u
 subject: 3, plur
 derivation: passive
 TAM: perfect
```

seg_file(*language*, *input_file*, *output_file*)
Options: gram=True, raw=False

Runs seg on the words in a file.

```
>>> l3.seg_file('am', 'l3/languages/am/data/ag.txt',
    'l3/languages/am/data/ag_out.txt')
Segmenting words in l3/languages/am/data/ag.txt
Writing to l3/languages/am/data/ag_out.txt
```

phon(*language*, *word*) [Amharic only]
Options: gram=True

Converts an Amharic word written in Ge'ez characters to a romanized form that shows consonant gemination and the epenthetic vowel (represented by '*I*'). If multiple pronunciations are possible, they are ordered by estimated frequency.

```
>>> l3.phon('am', "ይመታሉ")
yImetal_u (132) yIm_et_al_u (61)
```

```
>>> l3.phon('am', "ይመታሉ", gram=True)
-- yImetal_u
POS: verb, root: <mt'>
 subject: 3, plur
 grammar: imperfective, aux:alle
-- yIm_et_al_u
POS: verb, root: <mt'>
 subject: 3, plur
 grammar: imperfective, aux:alle, passive
```

```
>>> l3.phon('am', 'እንድብር')
?IndIbIr (0)
```

phon_file(*language*, *input_file*, *output_file*) [Amharic only]
Options: gram=True, print_ortho=False, word_sep='\n', anal_sep=' '

Runs phon on the words in a file.

```
>>> l3.phon_file('am', 'l3/languages/am/data/ag.txt',
    'l3/languages/am/data/ag_phon.txt')
Analyzing words in l3/languages/am/data/ag.txt
Writing analysis to l3/languages/am/data/ag_phon.txt
```

```
>>> l3.phon_file('am', 'l3/languages/am/data/ag.txt',
    print_ortho=False, word_sep=':')
Analyzing words in l3/languages/am/data/ag.txt
yIh:meShaf:yezarE:01:amet:gedema:bedenbu:mIrmera:alfo:tat_Imo:beweT_a:g
izE:tal_aq_ tal_aq:cIg_Ir:feTrob_IN_:neb_er:.
```

gen(*language*, *root/stem*, [*grammatical_features*])
Options: roman=False, guess=False [Amharic, Tigrinya only]

Generates the surface form of a word given a root or stem and optional grammatical features. With no features specified, a default form is output.

```
>>> l3.gen('am', "mWl'")
ሞላ
```

```
>>> l3.gen('am', "mWl'", roman=True)
mola
```

```
>>> l3.gen('om', 'sirb')
sirbe

>>> l3.gen('ti', "gWyy")
ጎየየ

>>> l3.gen('am', "mWl'", '[sb=[+p2,+fem],ob=[+plr,+l]]')
ሞላሽላቸው

>>> l3.gen('am', "mengst", '[+plr,+def]')
መንግስታቱ

>>> l3.gen('am', 'sdb', '[pos=n,v=agt,vc=cs,as=rc]')
አሳዳቢ

>>> l3.gen('am', 'brkt', '[pos=n,v=ins,pp=ke,cnj=m,+def]')
ከመበርከቻውም

>>> l3.gen('am', 'ne', '[+neg, sb=[+p1,+plr]]')
አይደለንም

>>> l3.gen('am', 'kongo', '[pp=be]')
This word can't be generated!

>>> l3.gen('am', 'kongo', '[pp=be]', guess=True)
በኮንጎ

>>> l3.gen('am', 'wddr', '[+gen, poss=[+p1,+plr]]')
This word can't be generated!

>>> l3.gen('am', 'wdd_r', '[+gen, poss=[+p1,+plr]]')
የውድድራችን

>>> l3.gen('om', 'sirb', '[sb=[+fem],tm=prf]')
sirbiteerti

>>> l3.gen('om', 'barbaad', '[+inf,cnj=f]')
barbaaduuf

>>> l3.gen('om', 'sob', '[der=[+autoben],sb=[+p2],+neg,tm=prs]')
sobattu

>>> l3.gen('ti', 'HSb', '[sb=[+p2,+fem],ob=[+plr]]')
ሐጸብክዮም

>>> l3.gen('ti', 'n|qTqT', '[vc=ps,tm=imf,sb=[+p1,+plr]]')
ንንቅጥቀጥ

>>> l3.gen('ti','gdf','[tm=j_i,+neg,sb=[+p2],ob=[+plr],vc=ps,as=rc]')
ኣይትጋደፎም
```