

Emofilt

A Prosody Filter to Simulate Emotional Speech

Felix Burkhardt

felix.burkhardt@t-systems.com

... **T** ... Systems

Get It ;-)

<http://emofilt.sourceforge.net/>

Emofilt

Mbrola

Based On MBROLA

- Free for non-commercial-use
Speech synthesis engine
- diphone based
- Available in 34 languages

Emofilt

Multilingual

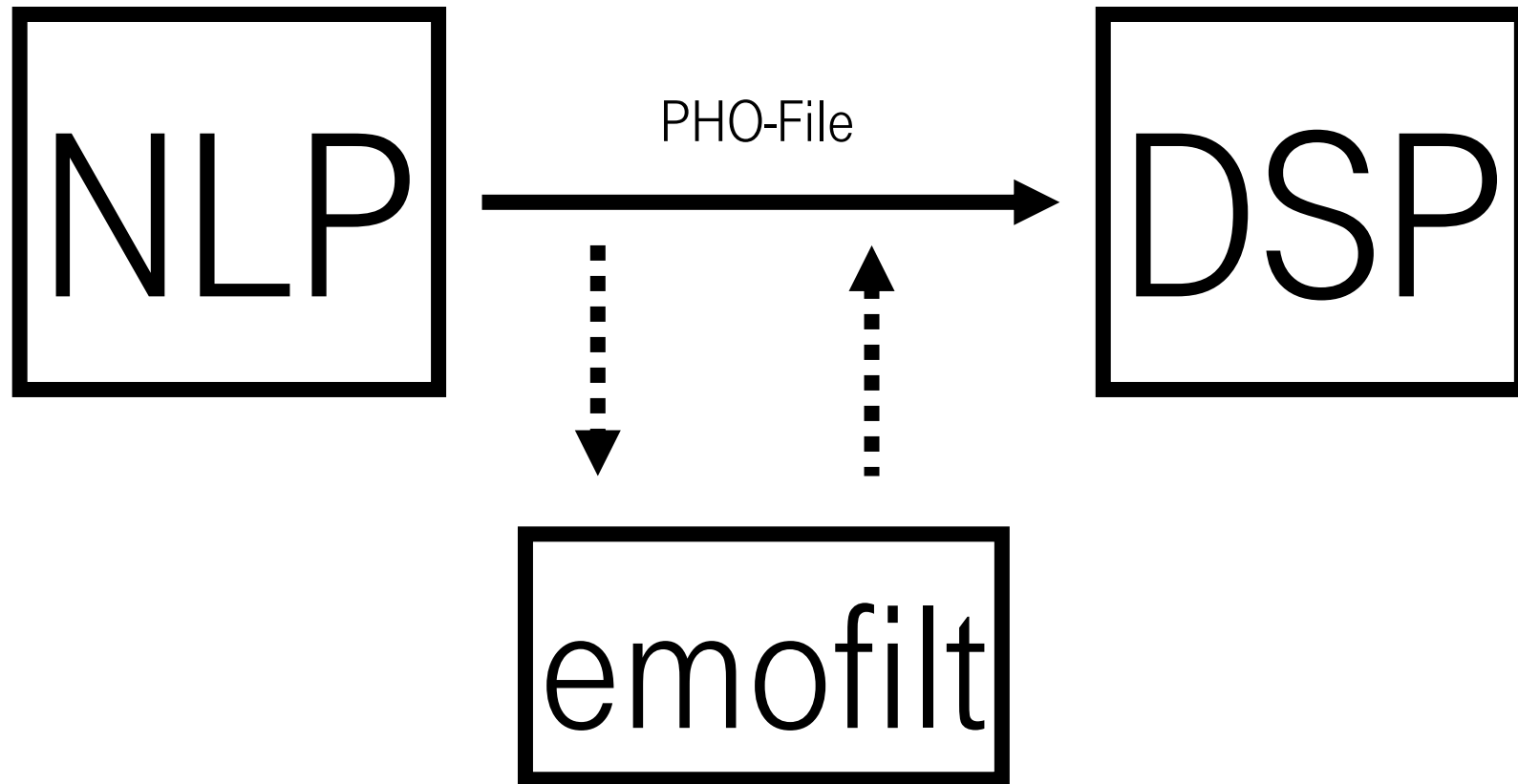
Mbrola Available in 34 languages

Language / voice dependent parts
stored in external XML file



Emofilt

Filtering Prosody



Emofilt

PHO-File Example

```
_ 35 54 160  
a 57 7 160 42 160 77 160  
n 54 13 160 50 160 87 160  
d 63 2 162 33 164  
e: 59 17 166 51 169 85 172  
n 64 17 175 48 178 80 181  
v 78 9 184 35 186 60 188 86 190  
O 63 14 191 46 191 78 191  
x 69 9 191 38 191 67 190 96 190
```

Emofilt

Emotion-File Example

```
<emotion name="boredom">  
  <pitch>  
    <changeMean rate="80"/>  
    <changeRange rate="50"/>  
    <changeVariability rate="80"/>  
  </pitch>  
  <duration>  
    <changeSpeechRate rate="120"/>  
    <changeDurSylsFocus rate="140"/>  
  </duration>  
  <phonation>  
    <changeEffort effort="soft"/>  
  </phonation>  
  <articulation>  
    <vowelTarget target="undershoot"/>  
  </articulation>  
</emotion>
```

Emofilt

Language-File Example

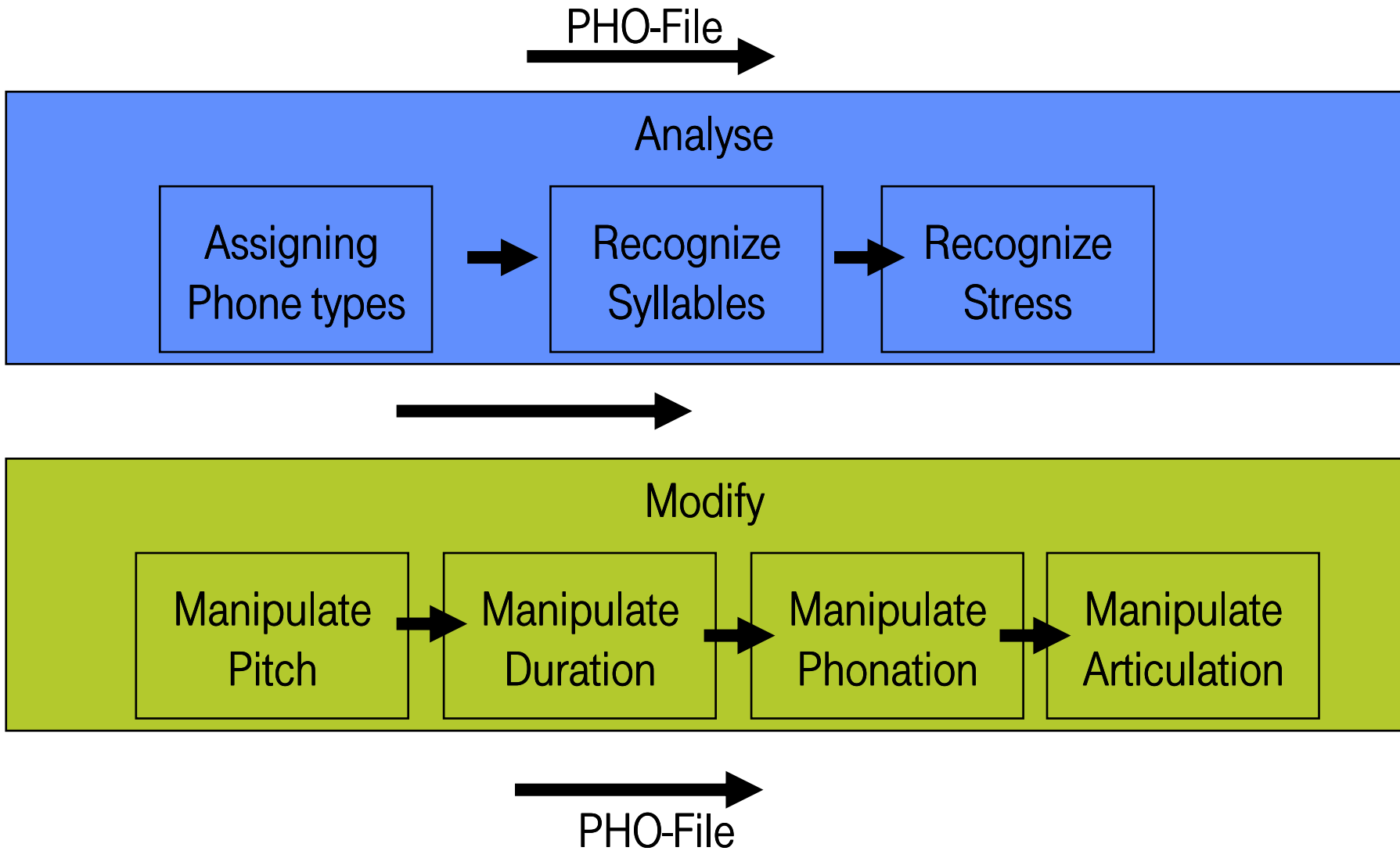
```
<language name="ex1">
  <description></description>
  <replacements>
    <replacement central="6" decentral="a:"/>
  </replacements>
  <phonemes>
    <phoneme name="a:" manner="long_vowel" voiced="true" />
    <phoneme name="a" manner="short_vowel" voiced="true" />
    <phoneme name="l" manner="approximant" voiced="true" />
    <phoneme name="n" manner="nasal" voiced="true" />
    <phoneme name="z" manner="fricative_voiced" voiced="true" />
    <phoneme name="S" manner="fricative_voiceless" voiced="false" />
    <phoneme name="pf" manner="fricative_voiceless" voiced="true" />
    <phoneme name="b" manner="stop_voiced" voiced="true" />
    <phoneme name="p" manner="stop_voiceless" voiced="false" />
    <phoneme name="_" manner="silence" voiced="false" />
  </phonemes>
</language>
```

Emofilt

Manipulations

- Pitch
 - Level
 - Range
 - Contour: Phrase Syllable
- Duration
 - For each sound class, for stressed/unstressed syllables
- Phonation
 - Jitter Simulation
 - German Voices DE6 and DE7 with three voice qualities
- Articulation
 - Simulation of Vowel target over/undershoot: Substitution of de-/centralized vowels

Emofilt Process Chain



Emofilt Developer

file emotion language utterance

phoneme borders
 syllable borders
 phoneme labels
 frequency scale
 time scale

andeny Ox n Embl n LOEt st lra: xhaU z @ @ a: Fo Unt ha: b @ a: g r E sb @ u: x t _

0s 1s 2s 3s 4s

pitch

level: rate 200
 range: rate 50
 variability: rate 90
 phrase contour: rate 0, cont. type rise
 phrase contour from stress: rate 100
 level unstress: rate 100
 level word-stress: rate 100
 level focus-stress: rate 100
 last syl contour: rate 0, cont. type rise
 contour unstress: rate 0, cont. type rise
 contour word-stress: rate 0, cont. type rise
 contour focus-stress: rate 20, cont. type fall
 wave model: rate 1 100, rate 2 100

duration

speech rate: rate 120
 focusstressed syllables: rate 100
 wordstressed syllables: rate 100
 unstressed syllables: rate 100
 v fric: rate 100
 vl fric: rate 100
 nas: rate 100
 aprx: rate 100
 vowel: rate 100
 vl stop: rate 100
 v stop: rate 100
 break: rate 100

phonation

jitter: rate 10
 vocal effort: effort normal
 unstressed: effort normal
 wordstressed: effort normal
 focusstressed: effort normal

articulation

vowel target: target normal

emotion: despair, voice: de6