

Synthese mit Mary

Modular Architecture for Research on speech sYnthesis

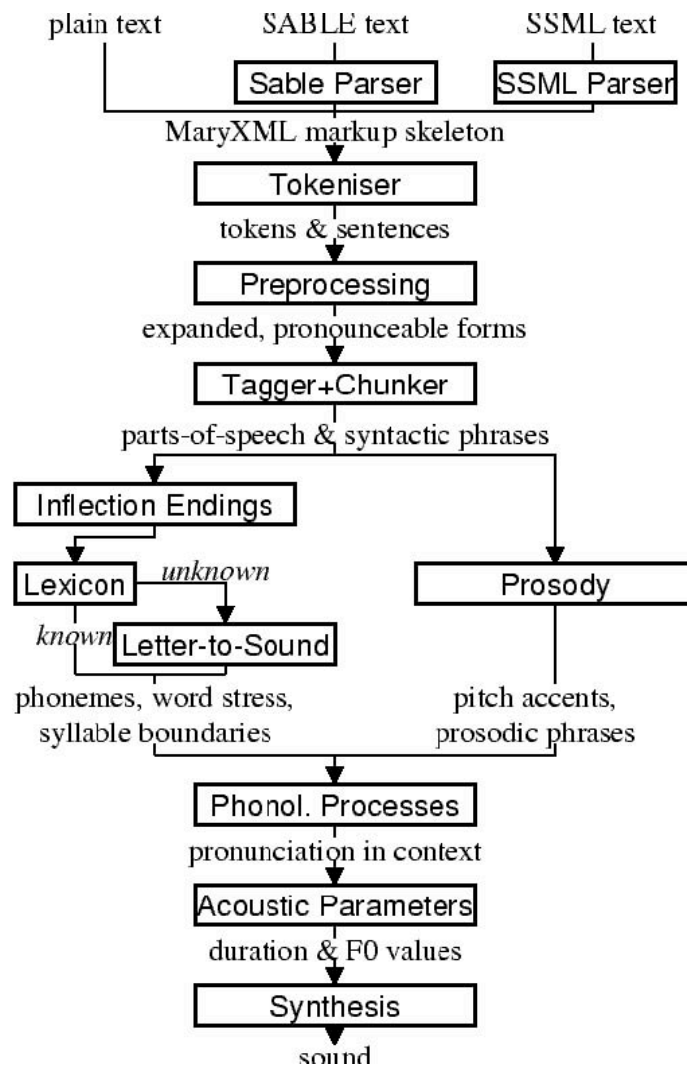
Entwickelt im DFKI und dem Phonetik-Institut, Universität Saarbrücken.

Siehe: <http://mary.dfki.de>

Tokeniser. Alle Module bis inkl. Tokeniser befassen sich damit, Wörter zu identifizieren, mit Kommas, Punkte, Fragezeichen usw. umzugehen. D.h. alle identifizierten Einheiten werden voneinander abgegrenzt.

Preprocessing: um Zahlen (1000 -> ein Tausend) und Abkürzungen (USA -> U S A) in Wörter umzusetzen.

Tagger & Chunker: um die syntaktischen Konstituenten vom Text aufzubauen – da die syntaktische Struktur vor allem die prosodische Struktur beeinflussen kann.



Inflection Endings

Wenn die Eingabe 'Mein 2. Angebot' war, wird der Text im Preprocessing in 'mein zwei Angebot' umgesetzt, und dann in 'mein zweites Angebot' durch dieses Modul.

Lexicon, letter to sound. Bestimmung der phonemischen Form entweder durch Abfrage aus dem Lexikon oder (wenn das Wort nicht im Lexikon vorhanden ist) durch letter-to-sound (Buchstaben-in-Phonem). Im zweiten Fall muss auch die Wortbetonung durch Regeln berechnet werden.

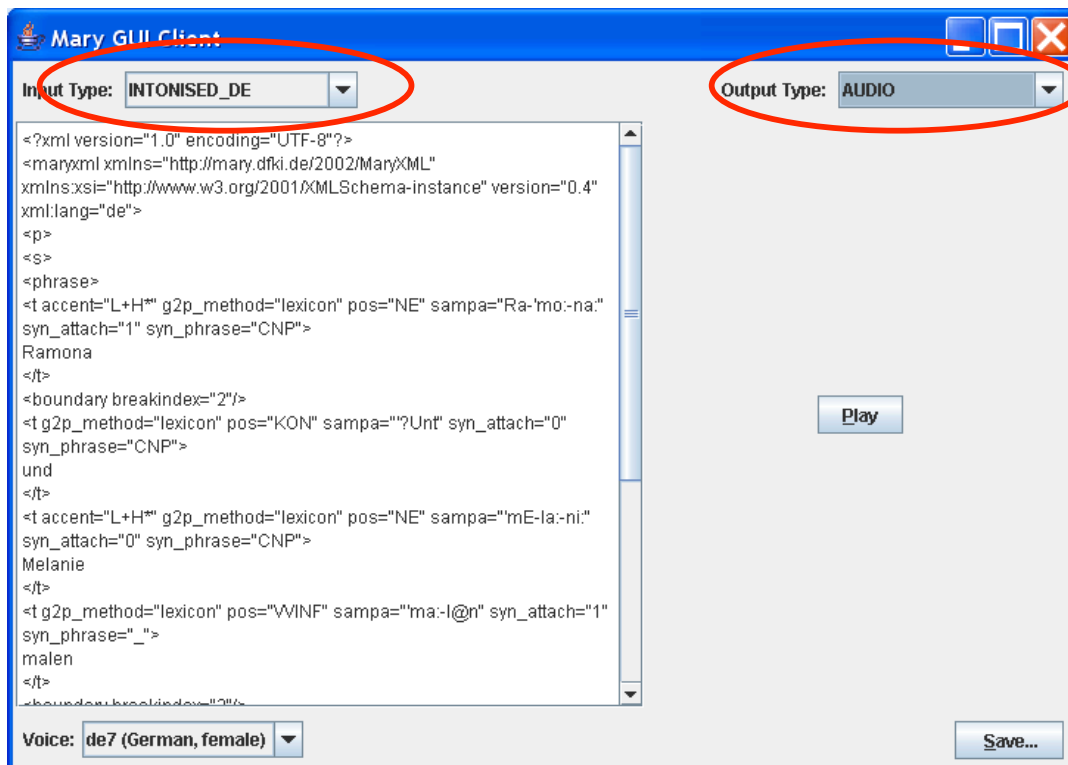
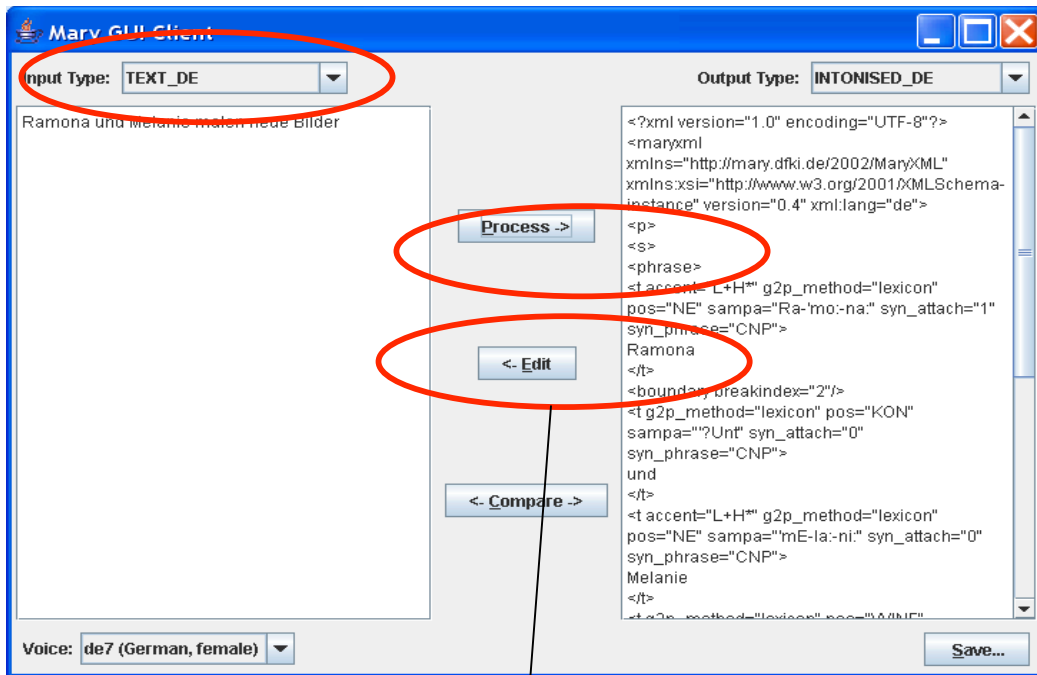
Prosody. Dem Text wird eine GTOBI-Etikettierung versehen, also Tonakzente, Grenztöne, und 'Break Indices' : 3 = Intermediärenphrase, 4 = äußerungsinterne Intonationsphrase, 5 & 6 = äußerungsfinale Intonationsphrase.

Die meisten Funktionswörter werden akzentuiert, jede Intermediärphrase enthält mindestens ein akzentuiertes Wort, ein Satz wird grob analysiert, ob es sich um eine Aussage, Frage, ja-nein Frage, handelt.

Phonological processes Einige Regeln um Elisionen, Tilgungen, Wortassimilationen also Vorgänge der Spontansprache anzuwenden.

Acoustic parameters, synthesis. Umsetzung in akustische Parameter und Synthese.

Parametereingabe für die Synthese der Prosodie



Wörter: allgemeine Information

Wortanfang und -ende von Melanie

Phoneme

Wortbetonung

Tonakzent

```
<t accent="L+H*" g2p_method="lexicon" pos="NE" sampa="mE-la:-ni:"  
syn_attach="0" syn_phrase="CNP">  
Melanie  
</t>
```

Phrasengrenzen

3 = Intermediärphrasengrenze, 4 = Intonationsphrasengrenze, 5 = Äußerungsfinal (4 und 5 sind fast identisch).

zB um eine Intermediärphrasengrenze nach 'und' zu setzen:

```
<t g2p_method="lexicon" pos="KON" sampa=""?Unt" syn_attach="0"  
syn_phrase="CNP">  
und  
</t>  
<boundary breakindex="3"/>
```

Phrasen und Grenztöne

(Ramona)H-	<boundary breakindex="3" tone="H-"/>
(Ramona)L-	<boundary breakindex="3" tone="L-"/>
[(Ramona)]L-L%	<boundary breakindex="4" tone="L-%"/>
[(Ramona)]H-L%	<boundary breakindex="4" tone="H-%"/>
[(Ramona)]L-H%	<boundary breakindex="4" tone="L-H%"/>
[(Ramona)]H-H%	<boundary breakindex="4" tone="H-^H%"/>

```
<t accent="L+H*" g2p_method="lexicon" pos="NE" sampa="Ra-'mo:-na:"  
syn_attach="1" syn_phrase="CNP">  
Ramona  
</t>  
<boundary breakindex="4" tone="H-%"/>
```

Tonakzente

zB wir wollen

[...malen neue Bilder] **H-H%**

L*

statt dem vorhandenen:

[...malen neue Bilder] **L-L%**

L+H* H*

</t>

<t accent="L*" g2p_method="lexicon" pos="VVINF" sampa="ma:-l@n"
syn_attach="1" syn_phrase="_">

malen

</t>

<boundary breakindex="2"/>

<t g2p_method="lexicon" pos="ADJA" sampa="nOY-@" syn_attach="1"
syn_phrase="NP">

neue

</t>

<t g2p_method="lexicon" pos="NN" sampa="bIl-d6" syn_attach="0"
syn_phrase="NP">

Bilder

</t>

<boundary breakindex="5" tone="H-^H%"/>

</phrase>

</s>

</p>

</maryxml>

Globale Änderungen

Die Spannweite um 200% in dem geänderten 'malen neue Bilder'.

```
</t>
<prosody range="+200%">
<t accent="L*" g2p_method="lexicon" pos="VVINF" sampa="ma:-l@n"
syn_attach="1" syn_phrase="_">
malen
</t>
<boundary breakindex="2"/>
<t g2p_method="lexicon" pos="ADJA" sampa="nOY-@" syn_attach="1"
syn_phrase="NP">
neue
</t>
<t g2p_method="lexicon" pos="NN" sampa="bIl-d6" syn_attach="0"
syn_phrase="NP">
Bilder
</t>
<boundary breakindex="5" tone="H-^H%"/>
</prosody>
</phrase>
</s>
</p>
</maryxml>
```

Andere <prosody > keywords

Sprechgeschwindigkeit um 50% erhöhen Durchschnittliche f0 um 50% senken, die f0-Spannweite um 30% senken in 'Ramona und Melanie'

```
<prosody rate = "+50%" pitch="-50%" range = "-30%">  
<t accent="L+H*" g2p_method="lexicon" pos="NE" sampa="Ra-'mo:-na:"  
syn_attach="1" syn_phrase="CNP">  
Ramona  
</t>  
<boundary breakindex="2"/>  
<t g2p_method="lexicon" pos="KON" sampa=""?Unt" syn_attach="0"  
syn_phrase="CNP">  
und  
</t>  
<t accent="L+H*" g2p_method="lexicon" pos="NE" sampa=""mE-la:-ni:"  
syn_attach="0" syn_phrase="CNP">  
Melanie  
</t>  
</prosody>
```

Aufgabe

Synthetisieren Sie den Satz auf S. 59 dieser Powerpoint Datei [ppt](#)

Nur hier und dort kann man noch ahnen wie schön sie war

Geben Sie zuerst den Text in 'Mary' ein. Notieren Sie die prosodische Etikettierung, die durch Mary erzeugt wird.

Versuchen Sie die Prosodie zu ändern, um den Satz mit der Prosodie und Melodie wie auf S. 59 der [ppt](#) Datei zu resynthetisieren.