

# CereVoice SSML Programmier-Anleitung

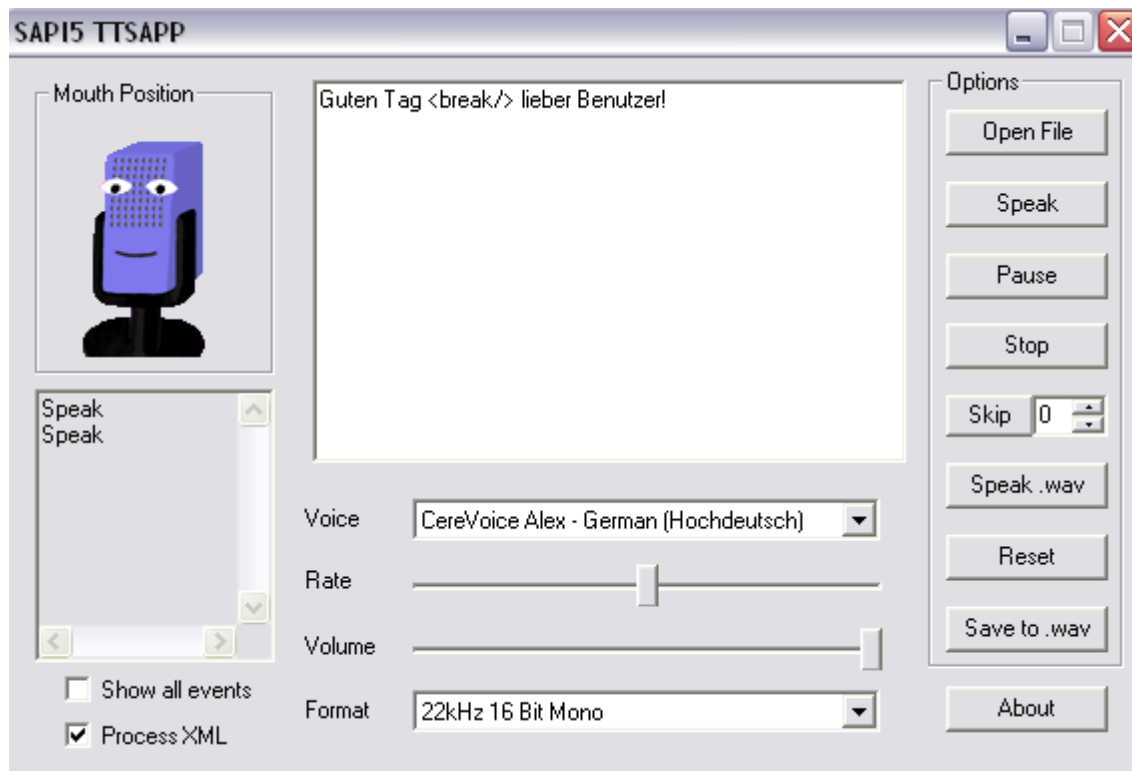
## 1. Speech Synthesis Markup Language SSML

Dieses Dokument erläutert im Kapitel 2 die von CereVoice unterstützte Markup-Sprache SSML und ihre Syntax. Speech Synthesis Markup Language (SSML) gehört als Subset von VoiceXML (VXML) zu XML-basierten Markup-Sprachen, standardisiert durch das World Wide Web Consortium (W3C).

Mittels SSML-Tags können Text-Passagen markiert werden, um die Parameter der Sprachsynthese zu verändern und somit das ausgegebene Sprachsignal zu alternieren, wie z. B. die Betonung, Sprechgeschwindigkeit oder Lautstärke. Die im Detail beschriebene Anwendung von SSML-Tags sowie einige Beispiele können der W3C-Homepage <http://www.w3.org/TR/speech-synthesis/> entnommen werden.

Kapitel 3 erklärt die von CereProc zusätzlich implementierte TTS-Funktionalität - CereVoice-Tags.

Um SSML- und CereVoice-Tags verarbeiten zu können, muss in der Standard-Schnittstelle SAPI von Microsoft die Option "Process XML" aktiviert werden (auf der Oberfläche links unten).



## 2. SSML-Tags in CereVoice

1. **break** erzeugt eine Pause von variabler Länge an angegebener Position mittels relativer Werte **x-weak, weak, medium, strong, x-strong** sowie numerischer Werte angegeben in Millisekunden (**ms**) oder Sekunden (**s**)

Hallo `<break/>` und guten Tag!

Hallo `<break strength="strong"/>` und guten Tag!

Hallo `<break time="250ms"/>` und guten Tag!

Hallo `<break time="3s"/>` und guten Tag!

2. **emphasis** betont den markierten Text mittels relativer Werte **reduced, moderate, strong, none**

Das ist ein `<emphasis level="strong">` großes `</emphasis>` Auto.

3. **prosody** alterniert die Sprachmelodie, indem die Werte der Attribute *volume*, *rate* und *pitch* verändert werden

**prosody volume** erhöht oder reduziert die Lautstärke mittels folgender Werte: **silent, x-soft, soft, medium, loud, x-loud, default** sowie numerischer Werte **0.1 - 100.0**

Sprechen Sie `<prosody volume="x-soft">` leise `</prosody>!`

Sprechen Sie `<prosody volume="2.0">` lauter `</prosody>!`

**prosody rate** erhöht oder reduziert die Sprechgeschwindigkeit mittels folgender Werte: **x-slow, slow, medium, fast, x-fast, default**

Er spricht `<prosody rate="x-slow">` langsam `</prosody>.`

**prosody pitch** erhöht oder reduziert die Tonhöhe mittels folgender Werte: **x-low, low, medium, high, x-high, default**

Heute bin ich `<prosody pitch="low">` müde `</prosody>.`

Heute bin ich `<prosody pitch="medium">` so glücklich `</prosody>.`

4. **sub** ersetzt den ursprünglichen Text mit dem im *alias*-Attribut angegebenen Text

`<sub alias="World Wide Web Consortium">` W3C `</sub>`

5. **p** und **s** markieren einen Absatz (p - paragraph) oder einen Satz (s - sentence)

```
<p>  
  <s> Das ist der erste Satz dieses Absatzes. </s>  
  <s> Das ist der zweite Satz dieses Absatzes. </s>  
</p>
```

### 3. CereVoice-Tags

1. **variant** ermöglicht alternative Aussprache mittels numerischer Werte 1 - 30

```
Guten Tag, lieber <usel variant="3"> Benutzer! </usel>  
Guten Tag, lieber Ben<usel variant="2">utzer! </usel>
```

Falls ein Satz, Wort oder ein Laut nicht richtig ausgesprochen werden, können mittels dieses Tags neue alternative Aussprachen - Varianten gewählt werden.

2. **phonemes** ermöglicht das Hinzufügen von Lautschrift-Zeichen, um die angegebene Aussprache vor der Standard-Aussprache zu bevorzugen

```
<lex phonemes="b aeh1 n t"> Band </lex>
```

Aufgrund dieser Markierung wird das Wort "Band" wie "b ae n t" in "eine Musik-Band" ausgesprochen und nicht wie "b a n t" in "das Band".

CereVoice unterstützt nur Lautschrift-Zeichen (Phoneme) aus dem CereVoice Phone-Set, zusammengefasst im Kapitel 4.

## 4. CereVoice Phone-Set Deutsch

Folgende Übersichts-Tabellen umfassen Lautschrift-Zeichen (Phoneme), die in CereVoice zum Wiedergeben der Aussprache von Lauten im Deutschen verwendet werden können.

Vokale		
Phonem	Beispiel-Wort	Lautschrift / Aussprache
@	machen	m a1 x @ n
a	Anfang	q a1 n f a ng
ah	Zahn	ts ah1 n
ae	hätte, nett bette	h ae1 t @ , n ae1 t b ae1 t @
aeh	spät	sh p aeh1 t
eh	Leder	l eh1 d er
e	Rekombi	r e1 k o m b ih
i	still bitte	sh t i1 l b i1 t @
ih	Niere biete Remmidemmi	n ih1 r @ b ih1 t @ r ae1 m ih d ae m ih
o	Motte	m o1 t @
oh	Dose Prozent	d oh1 z @ p r oh ts ae1 n t
oe	könnte	k oe1 n t @
oeh	Schön Öko	sh oeh1 n oeh1 k o
u	bunt	b u1 n t
uh	Mut genuin	m uh1 t g e n uh ih1 n
ue	hübsch	h ue1 p sh
ueh	Hüte Hypo	h ueh1 t @ h ueh1 p o
rv	der	d ae1 rv
er	Vater	f ah1 t er

Konsonanten		
Phonem	Beispiel-Wort	Lautschrift / Aussprache
p	Papa	p a p ah1
b	Ball	b a1 l
t	Tal	t ah1 l
d	Damm	d a1 m
g	Auge	q au1 g @
k	Kunst	k u1 n s t
f	Affe	q a1 f @
v	warum	v ah1 r u m
s	Tasse	t a1 s @
z	Rose	r oh1 z @
zh	Genie	zh e n ih1
sh	Schiff	sh i1 f
th	think	th i1 ng k
dh	that	dh ae1 t
q Glottal	Ei	q ai1
l	bald	b a1 l t
rl	traffic	t rl ae1 f i k
h	Hafen	h ah1 f @ n
ch	sicher	z i1 ch er
x	Buch	b uh1 x
w	water	w o1 t er
j	Jade	j ah1 d @
r	Radio	r ah1 d ih oh

Diphthonge		
Phonem	Beispiel-Wort	Lautschrift / Aussprache
ai	Reise	r ai <sup>1</sup> z @
oi	Leute	l oi <sup>1</sup> t @
au	Raute	r au <sup>1</sup> t @
ei	Steak	s t ei <sup>1</sup> k

Nasale		
Phonem	Beispiel-Wort	Lautschrift / Aussprache
m	Mal	m ah <sup>1</sup> l
n	Knall	k n a <sup>1</sup> l
ng	Wange	v a <sup>1</sup> ng @
en	Teint	t en <sup>1</sup>
an	Chance	sh an <sup>1</sup> s @
on	Pardon	p a rv d on <sup>1</sup>
oen	Parfum	p a rv f oen <sup>1</sup>

Affrikaten		
Phonem	Beispiel-Wort	Lautschrift / Aussprache
dzh	Jeans	dzh ih <sup>1</sup> n s
tsh	deutsch	d oi <sup>1</sup> tsh
ts	Zopf	ts o <sup>1</sup> pf
pf	Topf	t o <sup>1</sup> pf

Analog zum Kapitel 3.2. müssen diese Phoneme innerhalb des phonemes-Attributs im lex-Tag durch Leerzeichen getrennt werden.

Vokale am Wortanfang, in Komposita an den Wortgrenzen sowie in manchen Wörtern an Morphemgrenzen sind grundsätzlich mit vorangestelltem Glottal-Laut "q" (engl. glottal stop) zu transkribieren, wie z. B. "**q a n f a n g**" im Wort "**Anfang**", "**b u n d @ s q a m t**" im Wort "**Bundesamt**" oder "**f a e rv q ai s t**" im Wort "**vereist**" (vgl. Übersichtstabellen).

Um die Aussprache-Qualität zu erhöhen, kann die Hauptbetonung innerhalb der Lautschrift mittels der Zahl 1 markiert werden, z. B. "**q a<sup>1</sup> n f a n g**". Der betonte Vokal (Phonem) und die Zahl 1 dürfen nicht durch Leerzeichen getrennt sein - sie müssen zusammengeschrieben werden.