

SSML

SSML ist eine Markup-Sprache die vom World Wide Web Consortium (W3C) entwickelt wurde und als Subset von VoiceXML (VXML) zu den XML-basierten Markup-Sprachen gehört.

Mit SSML können Text-Passagen um Informationen wie Betonung, Lautstärke, Tonhöhe oder zum Beispiel auch Aussprachedauer erweitert werden die dann von der Sprachsynthese interpretiert werden. Eine detaillierte Beschreibung aller standardisierten SSML-Tags können auf der W3C-Website <http://www.w3.org/TR/speech-synthesis/> nachgeschlagen werden.

Die von Aristech angebotenen CereVoices unterstützen außerdem weitere nicht von der W3C standardisierte Tags.

Standard Tags

`<break/>`

Der break-Tag ist der einzige selbstschließende Tag und bietet die Möglichkeit pausen bestimmter Länge zu setzen.

Der Tag kann das Attribut `time` enthalten (empfohlen), indem die Pausenlänge in Sekunden oder Millisekunden beschrieben wird. Die Pausenlänge kann außerdem mit den Werten **x-weak**, **weak**, **medium**, **strong** oder **x-strong** und dem Attribut `strength` beschrieben werden. Der Tag kann allerdings auch ohne Attribute benutzt werden, was jedoch zur Folge hat, dass die Interpretation der Pausenlänge, wie auch bei dem `strength` Attribut der TTS überlassen wird und somit eventuell über unterschiedliche TTS-Hersteller unterschiedlich lange Pausen erzeugt werden.

Beispiele

```
Ist das dein Ernst?<break time="1s"/> Ich rede mit dir!
```

```
Ist das dein Ernst?<break strength="x-strong"/> Ich rede mit dir!
```

```
Ist das dein Ernst?<break/> Ich rede mit dir!
```

Standard Tags

`<emphasis/>`

Der emphasis Tag legt fest wie stark ein Wort oder Abschnitt betont werden soll. Er wird durch sein level Attribut weiter spezifiziert welches die Werte **reduced**, **moderate**, **strong** oder **none** enthalten kann.

Beispiele

```
Das ist eine <emphasis level="strong">große</emphasis>  
Herausforderung!
```

`<prosody/>`

Der prosody Tag kann sowohl die Lautstärke, die Geschwindigkeit als auch die Dauer und Tonhöhe eines Textabschnitts anpassen. Hierfür bietet er die Attribute **volume**, **rate**, **duration**, **pitch** und **contour**.

`<prosody volume/>`

Das volume Attribute kann floating point Werte enthalten, welche angeben, die wievielfache Lautstärke für den markierten Textabschnitt verwendet werden soll oder auch die Werte **silent**, **x-soft**, **soft**, **medium**, **loud**, **x-loud** und **default**.

Beispiele

```
<prosody volume="0.3">Bitte leise! Das ist eine Bibliothek.</prosody>  
<prosody volume="2.0">Kannst du bitte lauter Sprechen?</prosody>  
<prosody volume="x-soft">Bitte nur flüstern.</prosody>
```

Standard Tags

`<prosody/>`

`<prosody rate/>`

Das `rate` Attribut kann die Sprechergeschwindigkeit mit den Werten **x-slow**, **slow**, **medium**, **fast**, **x-fast** und **default** regulieren.

Beispiele

```
<prosody rate="x-fast">Schnell ich muss los.</prosody>
<prosody rate="x-slow">Ich bin total müde.</prosody>
```

`<prosody duration/>`

Mit dem `duration` Attribut kann die Sprecherzeit des markierten Abschnitts auf eine genaue Zeit festgelegt werden. Hierfür können wie bei dem `break`-Tag Zeiteinheiten in Sekunden oder Millisekunden festgelegt werden.

Beispiele

```
<prosody duration="4s">Wow!</prosody> Was war das?
```

Standard Tags

`<prosody/>`

`<prosody pitch/>`

Das pitch Attribut legt die Tonhöhe des markierten Abschnitts fest. Hierzu können wie beim volume-Attribut entweder floatingpoint Werte oder die Werte **x-low**, **low**, **medium**, **high**, **x-high** oder **default** verwendet werden.

Beispiele

```
<prosody pitch="1.5">Ich mag Helium Luftballons.</prosody>
<prosody pitch="0.7">Ich nicht.</prosody>
Ich bin <prosody pitch="high">so Glückliche.</prosody>
```

`<sub/>`

Mit dem sub-Tag können zum Beispiel Abkürzungen mit deren ausgeschriebenen Text mit Hilfe des alias Attributes ersetzt werden.

Beispiele

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

CereVoice Tags

<usel/>

Mit dem CereVoice usel-Tag kann durch das variant Attribut eine Alternative Aussprache für ein Wort festgelegt werden. Variant kann Werte von 0-30 annehmen, wobei 0 der Standardaussprache entspricht.

Beispiele

```
Wie wird <usel variant="3">Zwergzebu</prosody> ausgesprochen?
```

<phoneme/>

Der phoneme-Tag ermöglicht es in dem Attribut ph mit Hilfe des CereVoice Phone-Sets die Aussprache eines Wortes genau zu modellieren.

Beispiele

```
<phoneme phonemes="dzh ih1 n s">Jeans</phoneme>
```

<lex/> (nicht empfohlen)

Der lex-Tag ermöglicht es in dem Attribut phonemes mit Hilfe des CereVoice Phone-Sets die Aussprache eines Wortes genau zu modellieren.

Wir empfehlen den phoneme-Tag anstatt des lex-Tags zu verwenden.

Beispiele

```
<lex phonemes="dzh ih1 n s">Jeans</lex>
```