

FreieSoftwareOG –

Text2Speech -
ein pragmatischer Ansatz



Sprachsynthese - Begriffserklärung

Unter Sprachsynthese versteht man die künstliche Erzeugung der menschlichen Sprechstimme.

Ein Text-to-Speech-System (TTS oder Vorleseautomat) wandelt Fließtext in eine akustische Sprachausgabe.

Grundsätzlich lassen sich zwei Ansätze zur Erzeugung von Sprachsignalen unterscheiden.

Sprachsynthese - Begriffserklärung

Zum einen kann durch die sogenannte Signalmodellierung auf Sprachaufnahmen (Samples) zurückgegriffen werden.

Zum anderen kann das Signal aber auch durch die sogenannte physiologische (artikulatorische) Modellierung vollständig im Rechner erzeugt werden.

Sprachsynthese - Begriffserklärung

Während die ersten Systeme auf Formantsynthesen beruhten, basieren die zurzeit industriell eingesetzten Systeme vorwiegend auf Signalmodellierung.

Ein besonderes Problem für die Sprachsynthese ist die Erzeugung einer natürlichen Sprachmelodie.

Grundlagen – espeak vs. espeak-ng

espeak ist das alte, bei der Installation von espeak-ng wird automatisch ein Alias auf espeak erzeugt, sodaß man auch *espeak* „*Test*“ eingeben kann.

apt install espeak-ng

espeak – Installation

Installation von espeak unter GNU/Linux:

```
apt install espeak-ng
```



espeak – Kurzer Überblick

```
espeak-ng „Dies ist ein Test.“
```

```
espeak-ng --voices
```

espeak - Parameter

espeak-ng -v de+f3 „Dies ist ein weiterer Test.“

-v definiert die Stimme, f3 Geschlecht und Alter
Für detailliertere Informationen:

Andy	boris	f2	Gene	iven3	klatt	Lee	m3	m7	micel	rob	steph3	whisper
Annie	croak	f3	Gene2	Jacky	klatt2	linda	m4	Mario	'Mr serious'	robert	Storm	whisperf
AnxiousAndy	Denis	f4	iven	john	klatt3	m1	m5	max	norbert	steph	travis	zac
aunty	f1	f5	iven2	kaukovalta	klatt4	m2	m6	Michael	quincy	steph2	Tweaky	

espeak - Parameter

```
cd /usr/lib/x86_64-linux-gnu/espeak-ng-data  
cd voices/  
ls  
cd \!v/  
ls
```

espeak – Konfigurationsbeispiele

Es ist natürlich auch möglich, längere Texte direkt aus einer Datei lesen zu lassen:

```
espeak-ng -f lesetext.txt
```

Leider kann `espeak` nur `txt`-Dateien verarbeiten.

Mit dem kleinen Tool „`odt2txt`“ kann man das allerdings umgehen:

```
odt2txt --stdout lesetext.odt | espeak-ng --stdin
```

espeak – Export nach Audiodatei

Die Sprachausgabe kann mit `espeak-ng` auch in eine Wave-Datei exportiert werden:

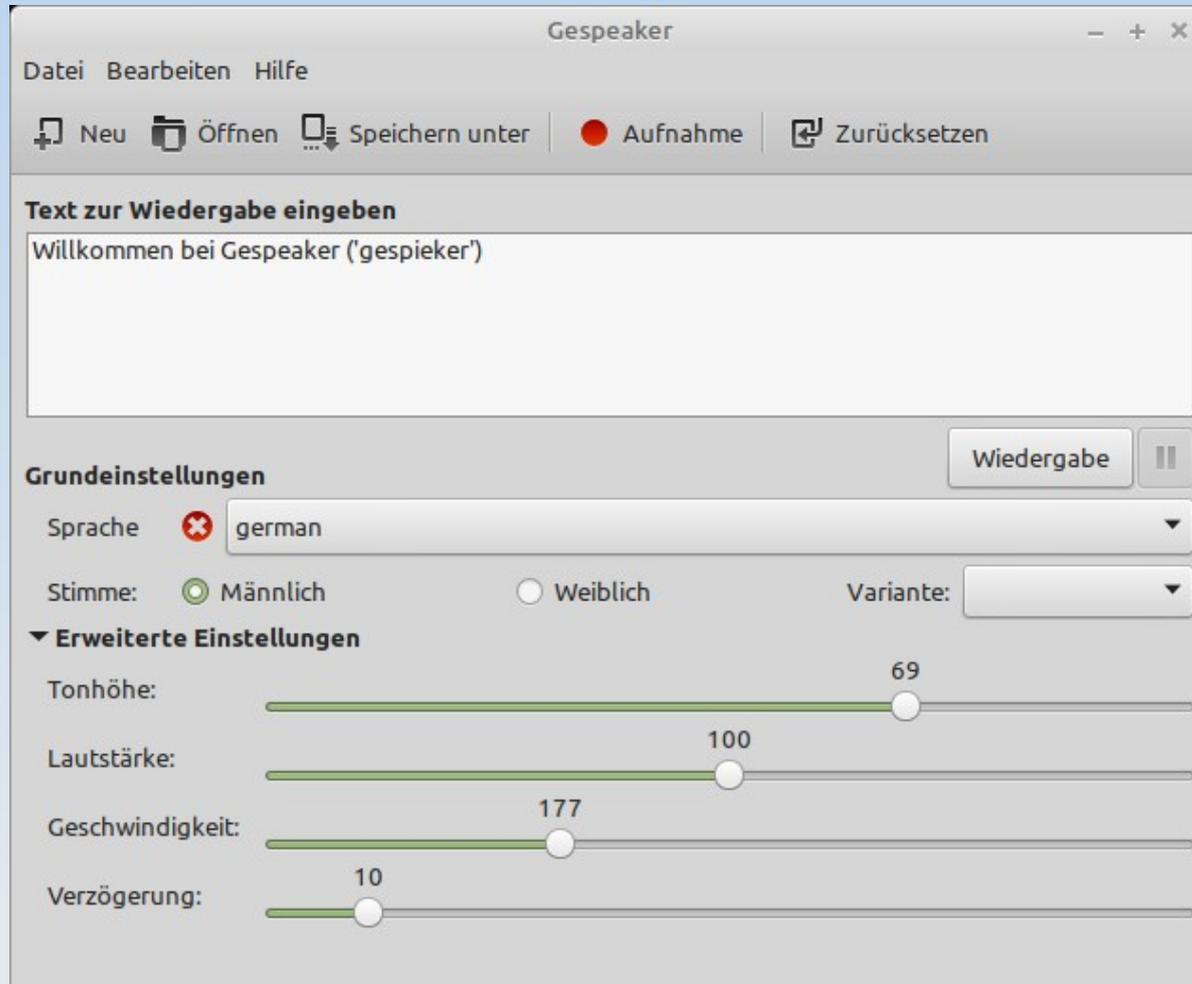
```
odt2txt --stdout langerlesetext.odt | espeak-ng -stdin -w LangeRede.wav
```

Witzigerweise dauert die Ausgabe in eine 10-Minuten Audiodatei nur wenige Sekunden.

espeak – Grafische Frontends

Wer es gerne klicki-bunti hat, kann sich mit „Gespeaker“ auch ein einfaches grafisches Frontend installieren.

espeak - Grafische Frontends



Exkurs - Spracherkennung

While this tutorial has been all about text-to-speech software, at the other end of that spectrum is voice recognition. In a world driven by Alexa, Cortana and the like, it shouldn't come as any surprise that voice recognition on Linux is also being actively pursued.

Unfortunately, there doesn't as yet exist a perfect solution to serve your every command.

Bitte beachten

Auf der Homepage findet sich immer das aktuelle Datum, sowie das Thema des nächsten Treffens!

Wer möchte, darf sich gerne auf die “Infomail”-Liste setzen lassen.

Weitergehende Informationen

<https://de.wikipedia.org/wiki/Sprachsynthese>

<https://wiki.ubuntuusers.de/espeak/>

<https://mycroft.ai/>

Weitere Informationen bekommen Sie hier:

<http://www.FreieSoftwareOG.org>

und

Kontakt@FreieSoftwareOG.org

oder kommen Sie doch einfach zu unserem regelmäßigen Treffen,
jeden 1. Mittwoch im Monat ab 20:00 Uhr.

(Treffpunkt laut Webseite)

