Specifikacija spletnega sintetizatorja govora WebSpeak

WebSpeak  je spletna storitev, ki omogoča preprosto vgradnjo sintetiziranega govora v širok nabor odjemalcev.

Odjemalec, ki je bodisi spletna stran bodisi aplikacija, pošlje želeno besedilo na strežnik HTTP. Strežnik pretvori besedilo v govor, ki ga v obliki zvočnega zapisa pošlje nazaj odjemalcu v predvajanje.

Ta dokument opisuje nabor funkcij WebSpeak, njihove parametre in odgovor, ki ga vračajo. Namenjen je bralcem z osnovnim poznavanjem protokola HTTP.

## Spremembe

1.0 Začetna verzija

1.1 Dodatni parametri omogočajo zdaj tudi metodo POST.

 Uporabniške poverilnice Basic (glava Authorization) niso več obvezne.

 Dodana funkcija speak2.

2.0 Preverjanje odjemalca s ključem API

 Ukinjeno upravljanje z vnaprej pripravljenimi posnetki

2.1 Tudi funkcija info zahteva ključ API

# Kazalo

[Spremembe 1](#_Toc43105557)

[Kazalo 2](#_Toc43105558)

[Splošne funkcije 3](#_Toc43105559)

[Funkcija *info* 3](#_Toc43105560)

[Funkcije za sintetični govor 4](#_Toc43105561)

[Funkcija *speak* 4](#_Toc43105562)

[Funkcija *speak2* 4](#_Toc43105563)

[Funkcija *retrieve* 5](#_Toc43105564)

# Splošne funkcije

## Funkcija info

Namenjena je splošni poizvedbi o verziji govornega strežnika in glasovih, ki so na strežniku nameščeni. Odjemalec praviloma najprej kliče to funkcijo, da lahko po potrebi v svojem uporabniškem vmesniku prikaže spisek glasov, ki so uporabniku na voljo.

### Zahteva

* Parametri:
	+ f=info
	+ k=<ključ API>
* Metoda: GET
* Pomembne glave zahteve:
	+ Accept-Language: jezik, v katerem želi odjemalec opis napake

### Odgovor

* Status: 200 OK ali druga koda napake
* Vrsta: application/json, oziroma text/plain; charset=UTF-8 v primeru napake
* Vsebina:

{

"version" : <verzija strežnika, niz>,

"voices" : [

{

"id" : <ID glasu 1, niz>,

"name" : <ime glasu 1, niz>,

"lang" : <jezik, niz, ISO 639-1>,

"gender" : <spol, "m"|"f">,

"age" : <starost, "child"|"teen"|"adult"|"senior">

},

{

"id" : <ID glasu 2, niz>,

"name" : <ime glasu 2, niz>,

"lang" : <jezik, niz, ISO 639-1>,

"gender" : <spol, "m"|"f">,

"age" : <starost, "child"|"teen"|"adult"|"senior">

}, ... ]

}

# Funkcije za sintetični govor

## Funkcija speak

Namenjena je pretvorbi besedila v govor. Če je tehnično izvedljivo, začne strežnik takoj, ko je mogoče, pošiljati zvočni posnetek po koščkih (Chunked transfer HTTP/1.1).

Za daljša besedila priporočamo uporabo metode POST.

### Zahteva

* Parametri:
	+ f=speak (vedno GET)
	+ t=<besedilo, ki ga želi odjemalec sintetizirati v zapisu UTF-8>
	+ v=<ID glasu, navedemo enega izmed glasov, ki jih vrne funkcija info>
	+ o=<želen zvočni zapis odgovora: mp3 = MP3, ogg = OGG Vorbis>
	+ k=<ključ API>
* Metoda: POST ali GET
* Pomembne glave zahteve:
	+ Content-Type: application/x-www-form-urlencoded za zahteve vrste POST
	+ Accept-Language: jezik, v katerem želi odjemalec opis napake
	+ Origin: URL spletne strani, če je ključ API vezan na domeno. Vsebovati mora samo <protokol>://<ime strežnika>[:<vrata>] (npr. https://ebralec.si)

### Odgovor

* Status: 200 OK ali druga koda napake
* Vrsta: audio/mpeg ali audio/ogg, odvisno od parametra o, oziroma text/plain; charset=UTF-8 v primeru napake
* Vsebina: binarni podatki, ki predstavljajo zvočni posnetek v želenem zapisu ali opis napake v jeziku, ki ga je uporabnik navedel v glavi zahteve Accept-Language

## Funkcija speak2

Namenjena je pretvorbi besedila v govor. Funkcija pripravi sintetiziran posnetek v svojem medpomnilniku in vrne njegov ID. To funkcijo uporabimo, kadar so besedila predolga za metodo GET, hkrati pa želimo posnetek takoj predvajati in predvajalnik ne omogoča metode POST in zna le prenesti posnetek z uporabo metode GET (npr. HTML5 <audio>).

Posnetki pripravljeni s to funkcijo se hranijo na strežniku samo omejen čas. Strežnik redno čisti posnetke. Za takojšnje predvajanje ali trajno hrambo posnetkov uporabimo funkcijo speak, ali po uporabi funkcije speak2 uporabimo še funkcijo retrieve.

Funkcijo speak2 sicer lahko uporabimo v svojem uredniškem sistemu, da sprožimo izdelavo posnetka novice vnaprej in s tem zmanjšamo zakasnitev pri dostavi posnetka prvemu poslušalcu po objavi novice. Vendar jo moramo vedno uporabiti tudi na spletni strani tik pred funkcijo retrieve. To zagotovi, da bo strežnik po potrebi ponovno izdelal posnetek, če je bil med tem slučajno že počiščen iz medpomnilnika.

### Zahteva

* Parametri:
	+ f=speak2 (vedno GET)
	+ t=<besedilo, ki ga želi odjemalec sintetizirati v zapisu UTF-8>
	+ v=<ID glasu, navedemo enega izmed glasov, ki jih vrne funkcija info>
	+ k=<ključ API>
* Metoda: POST ali GET
* Pomembne glave zahteve:
	+ Content-Type: application/x-www-form-urlencoded za zahteve vrste POST
	+ Accept-Language: jezik, v katerem želi odjemalec opis napake
	+ Origin: URL spletne strani, če je ključ API vezan na domeno. Vsebovati mora samo <protokol>://<ime strežnika>[:<vrata>] (npr. https://ebralec.si)

### Odgovor

* Status: 200 OK ali druga koda napake
* Vrsta: application/json, oziroma text/plain; charset=UTF-8 v primeru napake
* Vsebina: niz z ID-jem posnetka za funkcijo retrieve ali opis napake v jeziku, ki ga je uporabnik navedel v glavi zahteve Accept-Language

## Funkcija retrieve

Namenjena je predvajanju sintetiziranih posnetkov pripravljenih s funkcijo speak2.

### Zahteva

* Parametri:
	+ f=retrieve (vedno GET)
	+ i=<ID posnetka, ki ga vrne funkcija speak2>
	+ o=<želen zvočni zapis odgovora: mp3 = MP3, ogg = OGG Vorbis>
	+ k=<ključ API>
* Metoda: POST ali GET
* Pomembne glave zahteve:
	+ Content-Type: application/x-www-form-urlencoded za zahteve vrste POST
	+ Accept-Language: jezik, v katerem želi odjemalec opis napake
	+ Origin: URL spletne strani, če je ključ API vezan na domeno. Vsebovati mora samo <protokol>://<ime strežnika>[:<vrata>] (npr. https://ebralec.si)

### Odgovor

* Status: 200 OK ali druga koda napake
* Vrsta: audio/mpeg ali audio/ogg, odvisno od parametra o, oziroma text/plain; charset=UTF-8 v primeru napake
* Vsebina: posnetek ali opis napake v jeziku, ki ga je uporabnik navedel v glavi zahteve Accept-Language