

# Syntéza řeči pro nevidomé

Seminář z asistivních technologií  
podzim 2014

Adéla Štromajerová  
Klára Kuřová

# Syntéza řeči

- pomoc handicapovaným
- další aplikace
  
- proces „umělého“ vytváření řeči
- syntetizér řeči
- systém TTS
  
- cíl: vytvářet řeč v takové podobě a kvalitě, aby nebyla rozpoznatelná od řeči člověka

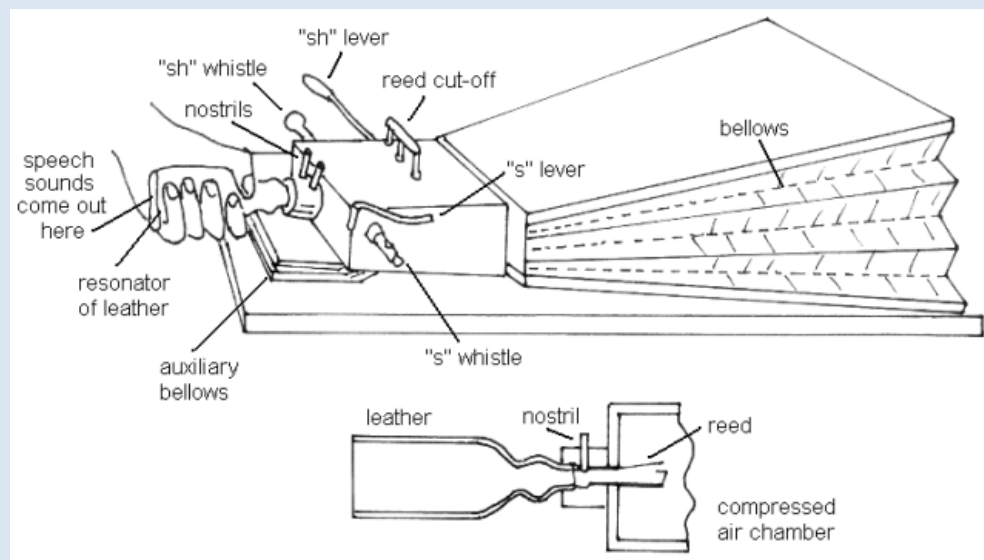
# Schéma syntetizéru řeči

- **Text → Zpracování přirozeného jazyka →  
Fonetická a prozodická informace →  
Syntetizér řeči → Řeč**
- syntetizér – jádro každého TTS
- zpracování přirozeného jazyka
  - analýza textu (předzpracování, normalizace, interpretace značkováného textu, lingvistická analýza)
  - fonetická transkripce (slovník vs. pravidla)
  - generování prozodie (intonace, rytmus, rychlost, hlasitost, přízvuk apod.)

# Historie

- Wolfgang von Kempelen
  - „mluvící stroj“
- Alexandr Graham Bell
  - pro potřeby hluchoněmých
- 1939 VODER – 1. syntetizér
- 50. – 80. léta – formantová syntéza (PAT, MITalk, DECTalk)
- 90. léta – boom konkatenáční syntézy
- 2000 – korpusově založená konkatenáční syntéza

# Kempelenův mluvicí stroj



# VODER – 1. syntetizér



# Syntéza ve frekvenční oblasti (formantová)

- simulace procesu vytváření řeči člověkem
- systém frekvenčních generátorů – hlasivky (parametry buzení)
- filtry a zesilovače – rezonance v dutinách (frekvence)
- syntéza podle parametrů modelu – pravidla
- dříve hojně využívaná, dnes už pouze v některých oblastech
- **výhody:** menší paměťové nároky
- **nevýhody:** robotický hlas, nedostatečně přesný model, výpočetně náročný vývoj systému

# Syntéza v časové oblasti (konkatenací)

- spojování předmluvených řečových segmentů (konkatenace)
- segmenty větší vs. menší
  - modelování intonace, přízvuku apod.
  - paměťové nároky
- obvykle difóny
- **výhody**: žádná složitá pravidla, reálný řečový signál, rychlejší a jednodušší návrh systému
- **nevýhody**: možná špatná kvalita, vyšší paměťové nároky
- v současnosti nejpoužívanější



# Artikulační syntéza

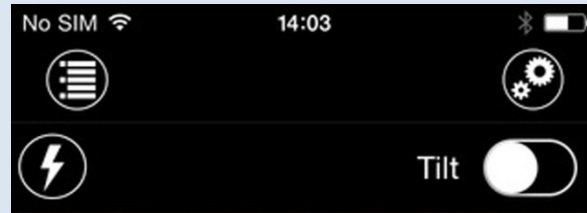
- kvalitní, nicméně nejnáročnější na implementaci
- artikulační parametry – př. retní průduch, retní výstupek, pozice špičky jazyka, šířka jazyka
- složité získávání dat pro model řečového aparátu
- prakticky pouze pro výzkumné účely

# Korpusová syntéza

- „unit selection synthesis“ – syntéza výběrem jednotek
- delší jednotky – difony až celé věty
- různé instance pro každou jednotku v korpusu
  - modelování intonace, přízvuku apod.
- **nevýhody:** obsáhlá databáze jednotek, možná nízká kvalita řeči (ne vše ve správném kontextu), výpočetně náročné

# Zařízení pro nevidomé

- **screen reader**
- elektronický zápisník
- mobil pro nevidomé (ONCE)
- čtečka knih (Beletrik)
- čtení titulků (KMPlayer)
- rozpoznání textu na fotografii (KNFB Reader)
- kamera pro odečítání textů z prostředí




Our challenge Sensotec was founded in 1986 as a company active in the development of aids for blind children, their parents and teachers in mainstream education. We have since grown into a business with 35 employees, operating from 3 sites. Over the



# Zahraníční syntetizátory

- Apple PlainTalk
- AT & T Natural Voices
- Cepstral
- eSpeak (MBROLA)
- Festival
- IVONA TTS
- Microsoft TTS

# Zahraníční screen readers

- Orca
- JAWS
- Acapela TTS Voices
- Google TalkBack

# České systémy

- Epos
- Demosthenes
- HLAS (Jiří Mojžíšek)
- Free(b)soft
- Speechech
- Galop

# Praktické ukázky



**KONEC**

Děkujeme za pozornost.