



Univerza v Ljubljani
Naravoslovnotehniška fakulteta
Oddelek za tekstilstvo
Grafična tehnologija

Seminarska naloga

PODJETJE AMEBIS IN PROJEKT GOVOREC

Predmet: Jezikovne tehnologije

Avtorja:
Andreja Židanik, Martin Volk

Ljubljana, Maj 2006

KAZALO

UVOD	- 3 -
1. 1 Presis	- 4 -
1. 1. 1 Vmesniki.....	- 5 -
1. 2 Besena	- 6 -
1. 2. 1 Modul za preverjanje besed.....	- 7 -
1. 2. 2 Modul za deljenje.	- 7 -
1. 2. 3 Vmesniki.....	- 7 -
1. 3 Prest.....	- 8 -
1. 4 Kolos	- 8 -
1. 4. 1 Zgradba sistema Kolos	- 9 -
2 SINTEZA GOVORA	- 11 -
2. 1 Zgradba sistema za sintezo govora	- 12 -
2. 2 Pretvorba teksta v enote sinteze.....	- 12 -
2. 3 Pretvorba teksta v metrične parametre.....	- 13 -
2. 4 Pretvorba sinteznih enot v govor	- 13 -
3 GOVORNI SIGNAL.....	- 14 -
4 PROGRAMI ZA SINTEZO GOVORA.....	- 17 -
4. 1 Avtomatska sinteza govora	- 17 -
4. 2 Postopki avtomatske sinteze govora	- 17 -
4. 3 Program za sintezo govora	- 19 -
5 PROJEKT GOVOREC	- 19 -
5. 1 Delovanje sistema.....	- 20 -
5. 2 Uporaba in navodila sistema Govorec.....	- 23 -
ZAKLJUČEK.....	- 26 -
LITERATURA IN VIRI	- 27 -

UVOD

Danes so jezikovne tehnologije temelj prihodnosti, ki bodo olajšale delo nam in računalnikom pri hitrejši oboje smerni komunikaciji. Prav tako so jezikovne tehnologije napredovale pri aplikacijah. z dodajanjem razpoznavanja govora v sisteme. Sistem je zasnovan tako, da se odzove na način z uporabo kombinacij naprej posnetih sposobnosti in umetno generiranih glasov, kar pa uporabniku z govorom omogoča lažje in hitrejše nakupovanje preko spleta, prevajanje, vnos besedila, itd...

Seminarska naloga vsebuje predstavljena področja iz prevajanja, opisa govornega signala in sinteze govora, ter s predstavitvijo realnih uporab tehnologij v aplikacijah.. V Sloveniji se z prevajanjem s pomočjo jezikovnih tehnologij ukvarja podjetje Amebis, ki je razvil dokaj širok spekter jezikovnih tehnologij. Poleg strojnega prevajanja iz angleščine v slovenščino in obratno se podjetje Amebis ukvarja tudi z projektom Besana in Kolos. Preko interneta se lahko znotraj projekta Kolos lahko pogovarjamo z Petrom Klepcem, ki odgovarja na vsa naša vprašanja. Napredek se kaže tudi v sistemih za sintezo govora, s pomočjo katerih stvari, ki so podane v tekstovni obliki spremenimo v govor. Ti nudijo predvsem pomoč slepim ali slabovidnim lahko pa tudi služijo za učenje jezika, itd... Osnovna naloga sinteze govora je pretvorba digitaliziranega besedila v govorjeno besedo. Projekt z imenom Govorec, ki omogoča slepim in slabovidnim lažje sporazumevanje s svetom je primer aplikacije..

Problematika pri jezikovnih tehnologijah je, da je jezik naravna stvar in se vedno spreminja in je potrebno zato potrebne korpuse sproti tudi popravljati.

1 PODJETJE AMEBIS

Podjetje Amebis se ukvarja pretežno z razvojem programske opreme na področju jezikovnih tehnologij in elektronskim založništvom. V Sloveniji so edino podjetje, ki je razvilo programsko opremo na področju strojnega prevajanja. Ukvarjajo se tudi z drugimi razvojnimi potmi na področju jezikovnih tehnologij in elektronskega založništva.

Najbolj pomembna jezikovna tehnologija, ki jo trenutno razvijajo je prav gotovo prevajalni sistem Presis.

1.1 Presis

Presis je prevajalni sistem, ki je namenjen prevajanju besedil iz slovenskega jezika v tuje jezike in obratno. Primeren je za različne skupine uporabnikov, zaradi vgrajenega pomnilnika prevodov. Presis omogoča hitro in učinkovito pomoč pri prevajanju, z dodatnimi nasveti pa pomaga pri oblikovanju končnega prevoda. Pri tem upošteva slovnična pravila slovenskega in angleškega jezika. Predvsem pa je še bolj učinkovit v kombinaciji z elektronskimi slovarji, ki jih je razvilo podjetje Amebis. Vendar pa še vedno ne smemo pričakovati, da bo Presis vsako besedilo vedno in v celoti pravilno prevedel, saj je to zaradi spreminjajočih se jezikov, pomenskih odtenkov ter različnih želja uporabnikov skoraj nemogoče. Trenutno Presis deluje samo na področju dveh jezikovnih parov, torej prevodov iz slovenščine v angleščino in iz angleščine v slovenščino.

Prevajalni sistem Presis ponavadi ne prevaja povsem pravilno. Prevod je namenjen približnemu razumevanju originalnega besedila ali osnovi za nadaljnje prevajanje. Vendar je kakovost prevoda odvisna tudi od dejavnikov, na katere lahko vplivamo sami.

Za boljše prevajanje se je potrebno držati določenih navodil:

- Omogočena JavaScript in piškotki (Cookies) sta nujna za pravilno delovanje teh strani.
- Stavke in lastna imena pišem z veliko začetnico (imena jezikov, dni v tednu in mesecev v slovenščini niso lastna imena).
- Ne uporabljamo SAMIH VELIKIH ČRK.
- Besedilo za prevajanje naj bo slovnično pravilno, saj to pomaga pri boljši analizi.
- Pisati moramo »kaša« ali »ka^sa« (če nimamo nameščene slovenske tipkovnice) in ne »kasa« »kasha« ali »kas"a«.

- Ne smemo pisati presledkov pred ločili (vejica, pika, podpičje, dvopičje, vprašaj, klicaj ipd.), ampak za ločili. Pred oklepajem in za zaklepajem naj bo presledek, razen če zaklepaju sledi ločilo.
- Vsa besedila, ki jih prevajamo na straneh, si shranjujemo kot pomoč pri nadaljnjem razvoju prevajalnih tehnologij.
- Predstavitvena verzija prevajalnega sistema Presis je omejena na prevajanje besedil z dolžino največ 200 znakov. V predstavitveni verziji ni mogoče dodajanje lastnih prevodov.

1. 1. 1 Vmesniki

Prevajalni sistem Presis omogoča delo s tremi uporabniškimi vmesniki. Vsi trije uporabljajo isto prevajalno jedro, zato se prevodi med uporabniškimi vmesniki ne razlikujejo. Različna je le njihova uporabnost in načini dela.

Poznamo torej tri vmesniki, to so: Presis Mini, Presis za Internet Explorer in Presis Pro.

- **Presis Mini** je hiter in enostaven uporabniški vmesnik, ki je lahko vedno na zaslonu. Lahko ga uporabljamo tudi kot slovar, pri čemer nam besed ni potrebno vpisovati v osnovni obliki, saj prevajalnik upošteva tudi slovnico.
Program lahko spremlja spremembe na odložišču, kar pomeni, da imamo prevod na voljo takoj, ko besedo (stavek ali celotno besedilo) postavimo na odložišče.
- **Presis za Internet Explorer** je dodatek pri uporabi prevajalnega sistema Presis in Dodatek za uporabo je uporaben med brskanjem po svetovnem spletu. Pritisk na gumb v programu Internet Explorer prikaže prevedeno vsebino trenutne strani. Pri tem na zaslonu ostanejo aktivne vse povezave, kar pomeni, da nas Presis ne ovira pri nadaljevanju brskanja.
- **Presis Pro** je najzmogljivejši uporabniški vmesnik, ki edini omogoča dodajanje besed in stavkov, povezavo s slovarji ASP32 in ima vgrajen črkovalnik. Omogoča hitro menjavanje različnih možnosti za posamezen prevod in omogoča shranjevanje prevedenega besedila.
Vendar je ta uporabniški vmesnik na voljo samo v poslovnem paketu. Besedilo, ki ga želimo prevesti, lahko enostavno vtipkamo v vnosno polje, lahko ga prenesemo s pomočjo odložišča, ali pa enostavno odpremo datoteko v formatu .txt ali .rtf. Pred prevajanjem lahko z vgrajenim črkovalnikom preverimo pravilnost zapisa in po potrebi popravimo napake, saj te lahko poslabšajo prevod. Prevajamo lahko na dva načina. Prvi način je prevajanje po stavkih, pri čemer naenkrat vidimo

le en stavek. Ko ga s pomočjo programa in njegovih nasvetov dokončno prevedemo, skočimo na prevajanje naslednjega stavka. Z ustreznim ukazom lahko sprožimo tudi avtomatsko prevajanje celotnega besedila. V tem primeru se po določenem času, ki je odvisen predvsem od hitrosti računalnika ter dolžine besedila, v prevodno okno izpiše prevod celotnega besedila. Tipična hitrost prevajanja je nekaj 10 strani besedila formata A4 v minuti. Za lažje delo si lahko celotno okno programa poljubno povečamo ali pomanjšamo, prav tako pa to velja tudi za velikost vnosnega in prevodnega polja ter velikost črk znotraj posameznega polja.

Presis Pro z barvami, ki si jih lahko prilagodimo tudi po svoji želji, označi štiri tipe prevedenih besed:

- 1) *besede z enim znanim prevodom,*
- 2) *besede z več možnimi prevodi,*
- 3) *neprevedena lastna imena in*
- 4) *neznane besede.*

S pomočjo barvnega označevanja lahko hitreje opazimo napake in jih zato lažje in hitreje ustrezno popravimo.

Prevod lahko poljubno spreminjamo, pri čemer nam program pomaga z nasveti. Pri besedah z več možnimi prevodi se ti nasveti pojavijo avtomatsko v trenutku, ko pridemo s kazalcem na določeno besedo ali pa jih priključimo s pritiskom na desni miškin gumb, ki odpre ustrezno okence/dialog. Z izbiro enega izmed ponujenih nasvetov program sam zamenja ustrezno besedo v besedilu, ki pa jo lahko še dodatno sami popravimo.

Če je na računalniku nameščen en ali več elektronskih slovarjev iz sistema ASP32, nam delo olajša tudi vgrajena povezava z njimi. Če smo s kazalcem na neki besedi v enem izmed obeh oken in pritisnemo ustrezen gumb, se nam odpre izbrani slovar in avtomatsko prikaže prevod besede v njem. Če program ne prevede določene besede oz. zanjo nima ustreznega prevoda, lahko prevod (v naš osebni slovar) dodamo sami. Prav tako lahko v pomnilnik prevodov shranimo tudi cele stavke, ki se pogosto pojavljajo v naših besedilih. Tako si lahko z nekaj vnosi, ki pokrivajo naše pogoste izraze, zelo izboljšamo kakovost prevoda.

1.2 Besena

Besana je črkovalnik, to je program, ki išče napačno napisane besede. Običajno uporablja slovar pravih besed in opozarja na besede, ki jih v tem slovarju ni. Besana je namenjena vgradnji v široko paleto Microsoftovih izdelkov. Uporabljamo ga lahko v Office 97, Office 2000, Office XP, Office 2003 in Outlook Express. Uporabljamo ga lahko v vseh programih, ne glede na jezikovno verzijo, ki jo

uporabljamo - torej lahko nadgradimo tudi črkovalnik, ki smo ga doslej uporabljali v slovenskih paketih. Nova verzija črkovalnika pozna več kot 1.2 milijona slovenskih besednih oblik (kar je podvojeno število glede na verzijo, ki je priložena v Office XP!) in je bolj oprta na novi slovenski pravopis.

Programski paket BesAna je namenjen vsem, ki uporabljajo računalnik za pisanje, saj veliko pripomore k slovnični pravilnosti besedila. BesAna ne more nadomestiti lektorja - ne zna najti vseh vrst napak, marsikatera napaka pa postane opazna šele na pomenski ravni, kar je za računalniško tehnologijo še nedosegljivo. Zaradi tega program lahko svetuje neustrezen popravek, vseeno pa tistemu, ki ureja besedilo, pomaga, da hitro odpravi očitne napake in se lahko bolj posveti drugim vidikom besedila.

BesAna omogoča delo z dvema uporabniškima vmesnikoma. Oba uporabniška vmesnika uporabljata isto pravopisno jedro, zato se poprave med uporabniškimi vmesniki ne razlikujejo. Različen je le način dela. Sestavna dela črkovalnika sta modul za preverjanje besed in modul za deljenje.

1. 2. 1 Modul za preverjanje besed

Osnovna naloga črkovalnika je iskanje pravopisnih in tipkarskih napak. Program pregleduje besedilo, dokler ne naleti na napako. Takrat odpre okno, v katerem ima uporabnik možnost besedo popraviti, pri tem pa mu program pomaga tudi z nasveti. Nasveti so po obliki podobne besede, ki jih ima črkovalnik v slovarju, ne glede na njihov pomen. Koristni so v primerih, ko uporabnik ne ve, katera oblika je pravilna. Tak primer so npr. besede "daljni", "dalnji", "daljnji", "nadaljni", "nadalnji" in "nadaljnji".

1. 2. 2 Modul za deljenje

Modul za deljenje omogoča lepo poravnavo besedila pri pisanju. Modul je v osnovi algoritmičen, vendar ima vgrajen tudi slovarski del, ki vsebuje izjeme, ki jih algoritem ne deli pravilno.

1. 2. 3 Vmesniki

Na voljo sta dva vmesnika.

BesAna Mini

je hiter in enostaven uporabniški vmesnik, ki je lahko vedno na zaslonu, ali pa skrit na vrstici opravil. Program lahko spremlja spremembe na odložišču, kar pomeni, da imamo popravo na voljo takoj, ko besedo (stavek ali celotno besedilo) postavimo na odložišče. BesAna Mini omogoča tudi tehniko povleci in spusti (angl. Drag & Drop), s katero dodajamo besedila v popravo.

BesAna za Word

Dodatek za uporabo programskega paketa BesAna je uporaben pri urejanju vseh dokumentov, ki jih lahko odpremo s programom Microsoft Word. Uporabniški vmesnik je narejen po Microsoftovih standardih in se skoraj neopazno vključi v delovanje programov Word in Outlook. BesAna za Word lahko preverja slovnico med pisanjem sprotno, ali na uporabnikovo zahtevo.

BesAna ne lektorira vedno pravilno. Kakovost preverjanja je odvisna tudi od dejavnikov, na katere lahko vplivate. Za boljše preverjanje se je potrebno držati spodnjih navodil.

Omogočena JavaScript in piškotki (Cookies) sta nujna za pravilno delovanje teh strani.

1.3 Prest

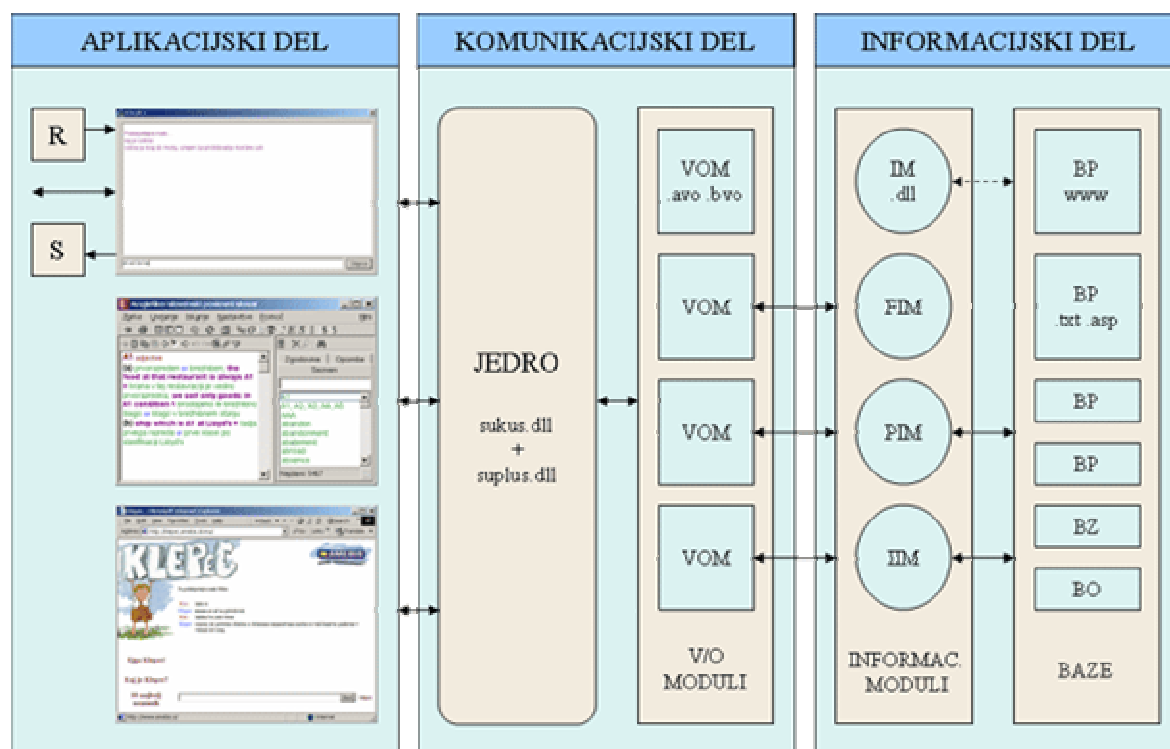
Prest je prvi program za optično branje dokumentov z vgrajenim slovenskim črkovalnikom. Namenjen je vsem, ki nočejo izgubljati časa za prepisovanje besedil v računalnik. Noben program ne more popolnoma natančno prebrati besedila, ki je zapisano kot slika, saj imamo pri tem težave celo ljudje. Zato si ljudje pomagamo tako, da iz sosednjih črk in besed "uganemo" črke, ki jih ne moremo prebrati. Tako deluje tudi program Prest. Prest zna pretvoriti vse normalne vrste črk in števil v obliko, ki je namenjena nadaljnjemu obdelovanju z urejevalnikom. Tudi program Prest ima že vgrajen enostaven urejevalnik. Vgrajen ima tudi črkovalnik (slovenskega jezika), zato že med delom s programom Prest dosežete veliko zanesljivost pretvorbe. Vgrajena možnost učenja posameznih znakov pa nam pomaga takrat, ko program ne pozna kakega tipa črk ali števil.

1.4 Kolos

Priti do takega uporabniškega vmesnika, ki bi nam včasih omogočil lažjo pot do informacij, seveda sploh ni tako enostavno. Glavni problem oziroma oviro predstavlja jezik, saj računalniki nikakor ne obvladujejo komunikacije v naravnem (slovenskem) jeziku tako kot bi si želeli ljudje. To seveda ne pomeni, da česa takega ni mogoče narediti.

Glavni cilj projekta Kolos je izgradnja informacijskega sistema Kolos, ki bo sposoben (inteligentno) komunicirati v naravnem jeziku, njegove module pa bodo lahko tudi drugi razvijalci in proizvajalci vgrajevali v svoje aplikacije.

1. 4. 1 Zgradba sistema Kolos



Slika: Zgradba sistema Kolos

Sistem Kolos lahko v grobem razdelimo na tri glavne dele:

- aplikacijski del - komunikacija med človekom in računalnikom teče preko uporabniških vmesnikov ali programov, ki so povezani oz. imajo vgrajene funkcije za tak način komuniciranja. Tak program je lahko namensko razvita aplikacija za pogovarjanje, nadgrajena in prilagojena že obstoječa aplikacija, spletna stran ali drugo. Programi, ki sestavljajo aplikacijski del projekta Kolos torej skrbijo za neposreden stik z uporabnikom, bodisi na pisni ali govorni način.

V aplikacijski del med drugim sodi tudi program za klepetanje preko spleta z imenom **Klepec**, ki služi kot univerzalni uporabniški vmesnik za komunikacijo v naravnem jeziku. Uporablja lahko vse ali pa samo nekatere V/O module s pripadajočimi informacijskimi podsistemi.

- komunikacijski del - ta del sistema predstavlja vezni člen med aplikacijskim in informacijskim delom. Ta del je ključnega pomena za uspešen potek dialoga, saj predstavlja most med človekom in njegovim jezikom na eni strani, ter računalnikom in njegovim "jezikom" na drugi strani.

Komunikacijski del sestavljajo:

Vgradni programski modul Sucus (.dll), ki ga lahko kličemo iz drugih programov.

Modul omogoča pretvorbo vhodnega vprašanja v naravnem jeziku v ustrezen odgovor. Pri nekaterih programih ja lahko izhod ustrezen logičen izraz, pri drugih kar ustrezen odgovor v naravnem jeziku. Sucus je programsko jedro komunikacijskega dela sistema Kolos, katerega glavna naloga je odgovarjati na zastavljena vprašanja oz. pretvarjati vhodna vprašanja v naravnem jeziku v ustrezne odgovore.

Odgovori so odvisni od modulov V/O, ki so na voljo.

Z moduli V/O (Vprašanje/Odgovor) Sucus odgovarja na določena vprašanja. Moduli V/O lahko vsebujejo osnovno znanje, ki je pri komunikaciji v naravnem jeziku potrebno za dialog z uporabnikom, lahko pa so le vmesnik, ki dodatnim informacijskim podsistemom doda ustrezen vhodni in izhodni format. Moduli V/O so v bistvu izvorna/izvajalna koda v lastno razvitem programskem jeziku z delovnim imenom K2, ki je posebej namenjen razvoju vmesnikov za dialog v naravnem (slovenskem) jeziku.

Seveda bi teoretično lahko take vmesnike razvijali tudi v drugih (klasičnih) programskih jezikih, vendar bi bilo to zaradi pregibnosti slovenskega jezika in prostega besednega reda besed v stavku izjemno težko in bi zahtevalo daleč preveč programskih vrstic kode že za reševanje zelo enostavnih problemov. Način programiranja v jeziku K2 se razlikuje od večine klasičnih programskih jezikov, saj je Sucus, ki interpretira oz. izvaja programe v jeziku K2 vzorčno voden sistem. To pomeni, da tok oz. proces izvajanja ukazov ne teče strogo usmerjeno, ampak je njegov tok v večini primerov nepredvidljiv. Sistem namreč pregleda vse ali večino vzorcev, jih primerja z vhodnim nizom in izvede tistega, ki je najbližji oz. ustreza še dodatnim pogojem.

Če želimo narediti vmesnik za komunikacijo v naravnem (slovenskem) jeziku, potrebujemo programsko jedro oz. dinamično knjižnico Sucus ali program, ki kliče Sucus, in vsaj en modul VOM (izvorni .AVO ali prevedeni/binarni .BVO), napisan v jeziku K2, v katerem so zapisana ustrezna navodila, kakšni naj bodo odgovori na določena vprašanja.

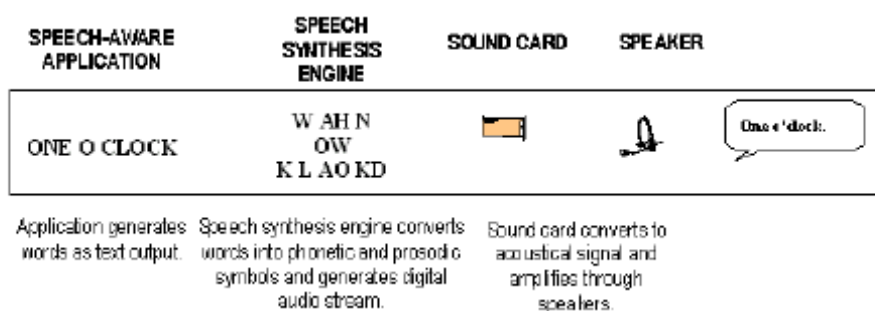
- informacijski del - informacije, ki jih računalnik preko dialoga prenaša uporabniku, zagotavlja informacijski del. Sestavljajo ga različne podatkovne zbirke in baze podatkov, funkcije oz. postopki za njihovo iskanje in vpisovanje, morebitni moduli umetne inteligence in podobno.

Informacijski del sestavljajo:

Informacijski moduli - moduli, ki module V/O oskrbujejo z informacijami, ki jih ti vsebujejo. Informacijski moduli lahko delujejo kot funkcije, ki za svoje delovanje ne potrebujejo dodatnih podatkovnih ali drugih baz, lahko pa si pri delovanju pomagajo tudi z njimi - baze - baze znanja, dejstev in podatkov v poljubnem formatu, s pomočjo katerih informacijski moduli posredujejo informacije modulom V/O.

2 SINTEZA GOVORA

Sinteza govora ali tekst-v-govor je proces pretvarjanja teksta v govor. To vključuje razbijanje besed v foneme; analiza za posebno rokovanje teksta kot so: številke, valut, sklanjatev in postavljanje ločil ter generiranje digitalne avdio oblike za predvajanje. Obstaja nekaj funkcij, ki jih sintetizator izvaja in so prikazane na spodnji sliki.



Slika : Delovni proces tekst-v-govor

Sintetizatorji ali tekst-v-govor glasovi izvajajo jezikovne sinteze, rokujejo z kompleksnostjo pretvorbe teksta in generirajo govorni jezik. Ti sistemi generirajo zvoke podobne tistim, ki jih govorijo ljudje in vključujejo različne filtre, ki simulirajo dolžino vratu, velikost ustne votline, obliko ustnic in pozicijo jezika. Kljub temu, da je razumljiv, zvok ki je narejen s to tehnologijo zveni manj človeško kot pa zvok narejen z digitalnim snemalnikom.

Kljub temu pa so tekst-v-govor aplikacije boljša alternativa v situacijah kjer je digitalno snemanje nepraktično.

Ponavadi se uporablja tekst-v-govor aplikacije v sledečih primerih:

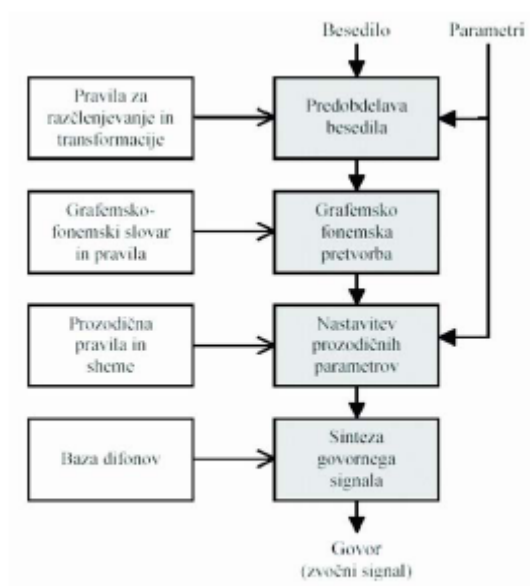
- Avdio snemanje je preveliko za hranjenje na disku ali predrago.
- Aplikacijski odgovori potrebujejo kratke fraze.
- Aplikacije ne morejo predvideti, če bo potrebna komunikacija ali odgovor.
- Uporabniki imajo raje avdio odgovor ali realističen.

2. 1 Zgradba sistema za sintezo govora

Postopek sinteze govora lahko razdelimo na tri glavna področja:

- Pretvorba teksta v enote sinteze
- Pretvorba teksta v metrične parametre
- Pretvorba sinteznih enot v govor

Shematičen prikaz pretvorbe teksta v govor:



Slika 3: Proces sinteze govora

2. 2 Pretvorba teksta v enote sinteze

Ker večina metod sinteze deluje na osnovi fonemov so enote sinteze fonemi, lahko pa bi uporabljali kakšne druge enote, vendar se taka izbira pokaže kot najboljša.

Kljub temu, da so ponavadi vhod ASCII znaki, je pretvorba večstopenjska operacija katere optimum še ni odkrit.

Prva stopnja pri pretvorbi je normalizacija. Z normalizacijo interpretiramo odstavke, ločila, zaporedja števk se razvijejo v ustrezne številke – glavne ali vrstilne. Okrajšave in akronimi se zapišejo v polni obliki. Zaporedje besed se nato pretvori v zaporedje enot sinteze (fonemov). Pri pretvarjanju besed lahko najprej iščemo v slovarju pogostih izgovorjav, ki ga je predhodno potrebno zgraditi za vsak jezik posebej. Če besede ne najdemo v slovarju ali ga nimamo, ji moramo izgovorjavo določiti sami. Drugi del pretvorbe predstavlja napovedovanje mesta naglasa. Mesto naglasa je zlog, na katerem ima beseda jakostno in tonsko izrazitost.

Prav tako ima lahko posamezna beseda več mest naglasa. Zadnja faza je pretvorba besed v fonetični zapis. Za to pretvorbo lahko uporabljamo produkcijska pravila za posamezne glasove, ki jih moramo predhodno sami določiti in so odvisna od jezika.

2. 3 Pretvorba teksta v metrične parametre

Določanje metričnih parametrov spet lahko razdelimo na tri dele.

Prvi del je določanje intonacijskih mej.

Za idealno določanje mej je potrebno poznavanje celotne sintaktične analize stavka. Pogosto pomanjkanje intonacijskih mej povzroči, da je glas bolj grob. Preveliko število intonacijskih mej pa povzročajo zmedo in motnje. Najenostavnejša metoda je postavljanje mej, kjer le te določajo ločila. Malo bolj izpopolnjena rešitev je rešitev v kateri imamo shranjeno listo funkcijskih besed pri katerih nastopajo meje. Poleg teh dveh enostavnih metod obstaja še veliko bolj izpopolnjenih metod.

Drugi del je določanje dolžin posameznih segmentov.

Dolžina segmentov se ponavadi določa z množico pravil. Pravila so sestavljena iz velikega števila faktorjev, ki bi lahko vplivali na dolžino segmenta. Nekateri izmed teh faktorjev so: frekvenca besede, sintaktična kategorija, fonetičen kontekst in še mnogi drugi.

Tretji del je določanje frekvenčnih skic.

Osnovna frekvenca človeškega govora je določena kot kombinacija velikega števila faktorjev. Sistemi za sintezo govora ne upoštevajo vsebine teksta, ki ga sintetizirajo, ker bi morali poznati vsebino teksta. Prav tako ne upoštevajo nekaterih drugih lastnosti človeškega glasu, zato je govor bolj zglajen kot človeški.

2. 4 Pretvorba sinteznih enot v govor

Ko imamo določene vse parametre - stavčno intonacijo, naglašene zloge besed, znakovni zapis pretvorjen v zaporedje glasov izgovorjave moramo glasove še pravilno realizirati.

To lahko naredimo na dva načina:

Z generiranjem glasov s pomočjo simulacije človeškega govornega trakta

- Artikularni: fizični model je zasnovan na podlagi opisa fiziologije in akustike govornega trakta
- Formantni sintetizatorji: prenosna funkcija je opisana s formantnimi frekvencami in formantnimi amplitudami - umetno rekonstrukcijo formantnih značilnosti dosežemo z vzbujanjem niza resonatorjev in antiresonatorjev (periodično vzbujenje za zveneče glasove in vzbujaje s šumom za nezveneče glasove)

Za dobro generiranje potrebujemo dober matematični model govornega trakta, poraba pomnilnika pa je majhna. Slabosti pa sta kompliciran matematični model pri artikularnih sintetizatorjih in težavno nastavljanje formantnih parametrov pri formantnih sintetizatorjih.

Z združevanjem predhodno posnetih enot

Predhodno posnete enote so vzete iz naravnega govora - na primer: alofoni, difoni, trifoni večji deli pogosteje uporabljenih besed, celih besed ali kombinacijo le-teh. Potrebujemo torej bazo osnovnih enot, s katerimi bomo tvorili govor. Pri manjših osnovnih enotah je poraba pomnilnika manjša, vendar pa so pri večjih enotah rezultati boljši. Zaradi spektralnih nezveznosti, ki nastanejo pri lepljenju osnovnih enot, uporabljamo postopke, ki te nezveznosti zgladijo.

Najbolj razširjeni postopki, ki se uporabljajo pri obdelavi signala, so:

- večpulzno linearno napredovnanje (LPC)
- obdelovanje signalov sinhrono z osnovno periodo (PSOLA in WSOLA)
- kombiniranje harmonskih in šumnih modelov (HNS, MBR-PSOLA)

3 GOVORNI SIGNAL

Pri razpoznavanju govora se ukvarjamo s problemom pretvorbe signala, ki ga proizvaja govorec s svojimi govorili, v zaporedje simbolov - govornih enot (fonemov, zlogov, besed), ki jih je govorec s tem signalom posredoval.

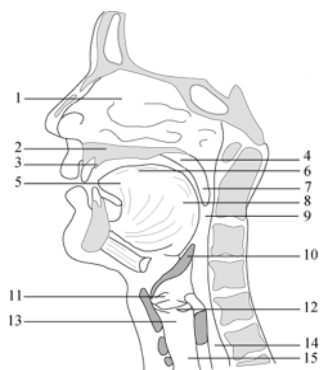
Tako lahko razumemo razpoznavanje govora kot transformacijo akustičnega signala v zaporedje simbolov, ki jih vnaprej predpišemo. Pri teoriji razpoznavanja vzorcev to prevedemo v iskanje najbolj verjetnega zaporedja simbolov, ki ga opiše zaporedje danih akustičnih dogodkov.

Akustični dogodki so ponavadi opisani z vektorji značilk, ki jih pridobivamo iz govornega signala. Pridobivanje akustičnih značilk iz govornih signalov poteka na podlagi različnih časovno–frekvenčnih predstavitev.

Govorni signal je težaven za časovno–frekvenčno analizo. Vsebuje tako izrazite ne stacionarnosti, kratke, hipne spremembe (npr. odpore pri zapornikih), kot tudi daljše stacionarne dele (npr. samoglasniki). Zato predstavlja velik izziv časovno–frekvenčnim predstavitvam.

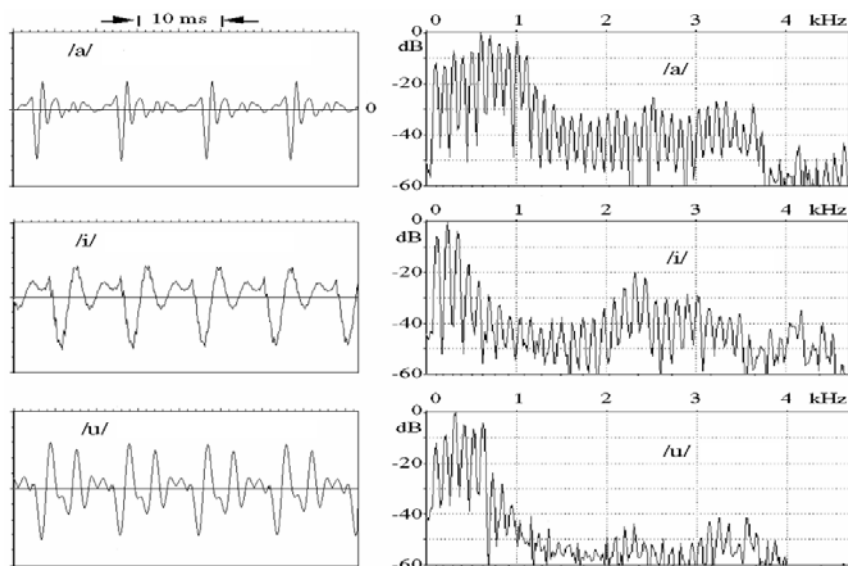
Dejstva, ki jih je potrebno upoštevati pri govornem signalu:

- govor ni zaporedje neodvisnih fonemov
- govor vsebuje več informacije kot besedilo: čustva, odnos govorca, razpoloženje
- pomen besedila (gori na gori gori)
- glavni vir energije glasu so pljuča (izdišni zrak)
- pot zraka: sapnik, grlo, grlna votlina, goltna votlina, ustna, nosna votlina.
- glasilki, 2 gibalna hrustanca. odpiranje in zapiranje zračne poti, nihanje (F0)
- razmaknjeni glasilki: H (aha). zaprti in nihata: grlni K.
- ostale glasove tvorimo z oblikovanjem grlne cevi nad grlom (razmik izgovornih ploskev).



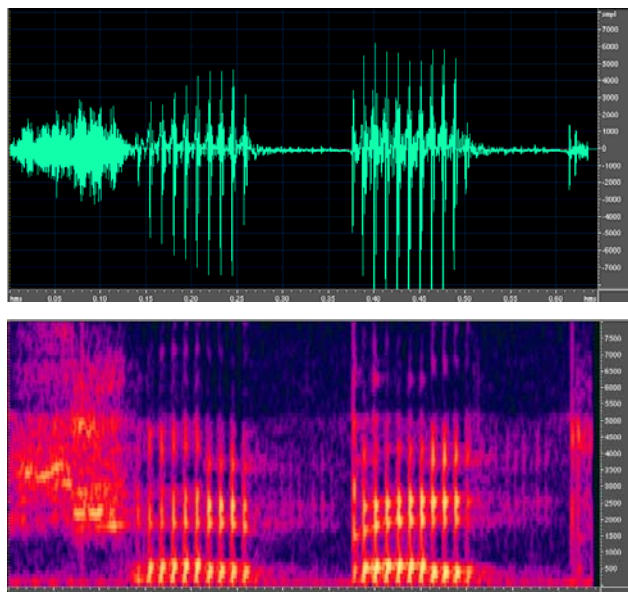
Slika : Govorni organi pri človeku

Fonetika je zvočna podoba jezika od glasu, naglasa v besedi do celotnega besedila. Osnovna razdelitev glasov je na zvoneče in nezvoneče. Zvoneče delimo na osnovno frekvenco (F0) in višje harmonske frekvence, glas pa tvorijo glasilke. Pri nezvonečih glasovih so glasilke razmaknjene zato ni F0, ni višjih frekvenc, so tišji, poleg tega se glas se tvori nekje na poti do ust.

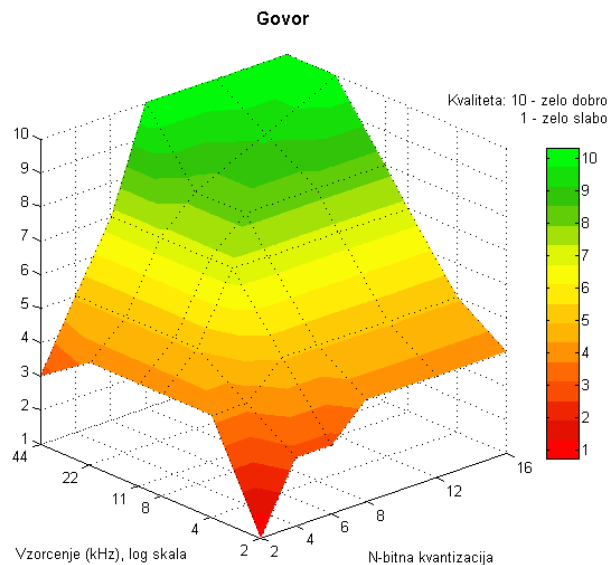


Slika : Časovna in sprekrnalna predstavitev samoglasnikov a, i in u

Samoglasniki imajo več energije (svetle barve) pri nižjih frekvencah, soglasniki pri višjih. Nezveneči soglasniki imajo zelo malo energije (temne barve). Glasovni format besede šepet v obliki wav preko interneta sepet.wav.



Slika : Časovna in spektralna predstavitev besede »šepet«



Slika : Grafični prikaz kvalitete govora

4 PROGRAMI ZA SINTEZO GOVORA

4. 1 Avtomatska sinteza govora

Gre za posredovanje informacij, ki jih s pomočjo signala stroj posreduje uporabniku – človeku.

Prednosti uporabe tovrstne sinteze pri komunikaciji med človekom in strojem so:

- Uporabnik zlahka razume govorjena sporočila, ki mu jih posreduje stroj brez predznanja in zahtevanja večje koncentracije, saj je govor najpogostejše in najučinkovitejše sredstvo za sporazumevanje med ljudmi.
- Uporabnik lahko sprejme informacijo, tudi ko opravlja druge aktivnosti npr. hodi, dela na drugih objektih.
- Posredovano sporočilo ne potrebuje medija za svoj zapis papirja ali zaslona, kar je lahko pri posredovanju daljših sporočil tudi slabost.

4. 2 Postopki avtomatske sinteze govora

Postopke avtomatske sinteze danes razdelimo v tri skupine:

Sinteza s pomočjo kodiranja govornega signala

Za tvorjenje govornih sporočil uporabljamo vnaprej posnet govorni signal, ki je lahko zaradi zmanjšanja pomnilniških zahtev kodiran. Kvaliteta sestavljenega sporočila je odvisna od ujemanja značilk govora med posameznimi sestavljenimi besedami ali frazami. Med akustične značilke prištevamo ovojnico,

amplitudo, osnovno frekvenco in hitrost izgovora. Če shranimo daljše fraze, stavke dosežemo večjo naravnost in razumljivost sintetiziranega govora. Če pa posnamemo manjše enote, kot so zlogi in fonemi, lahko tvorimo večje število različnih besed in stavkov, vendar je kvaliteta slabša.

V praksi se shranjujejo besede in fraze, nato se pri generiranju sporočila uporabi ustrezne besede v kombinaciji s frazami. Pri tovrstni sintezi pa je največji problem nenaravnost govora, saj se izgovorjen stavek precej razlikuje od zaporedja besed, ki so izgovorjena izolirano. To rešujemo tako, da posamezno besedo posnamemo z različnimi poteki osnovne harmonske frekvence (naraščajoča, konstantna in padajoča). Vendar sintetiziran govor še vedno ni popolnoma zadovoljiv, saj je potrebno rešiti tudi problem naglasa v stavku, ritem in intonacijo, ki so odvisni od sintaktičnih in semantičnih dejavnikov.

Problem, ki se še pojavi pri tovrstni sintezi je razširitev besed. Če želimo dodati novo besedo, to ne moremo storiti brez dodatnega poprejšnjega snemanja. Torej moramo imeti zmeraj na voljo ustrezen prostor z enakim akustičnim okoljem in referenčnega govorca, ki bo izgovoril novi material. Če starega govorca nimamo na voljo, mora novi govorec izgovoriti, celotni slovar, čeprav smo želeli dodati le eno besedo.

Sinteza zasnovana na postopku analiza-sinteza

Sinteza zasnovana na postopku analiza-sinteza je izboljšana različica prejšnje, saj obsega manjše število podatkov, ki so potrebni za sintezo govora. Tu se prvotno lotimo analize izgovorjave besed, posameznih fraz in določimo zaporedje značilk, ki podajajo časovno spreminjanje značilnosti analiziranega govora. Pri sintezi nato sestavljamo značilke posameznih osnovnih enot - fonemov,, besed, fraz. S pomočjo sestavljenih zaporedij značilk sintetizator generira govorni signal.

Na ta način tvorimo želeno govorno sporočilo. Ker moramo za posamezno osnovno enoto shraniti le zaporedja značilk, je obseg podatkov, ki jih potrebujemo za sintezo, mnogo manjši kot pri postopku kodiranja govora. Prednost je tudi ta, da je enostavneje spreminjati značilnosti sintetiziranega govora (višina tona, hitrost govora), saj lahko to dosežemo s spreminjanjem vrednosti parametrov (vektorjev značilk). Kot zaporedja značilk lahko uporabimo LPC in PARCOR koeficiente, ki jih določimo s pomočjo linearne predikcije. Sisteme, ki izvajajo postopke določanja parametrov govora (vektorjev značilk) imenujemo vokoderji in z njihovo pomočjo lahko v veliki meri zmanjšamo pomnilniške kapacitete.

Sinteza s pomočjo pravil. Tvorimo govorni signal s pomočjo fonetičnih in lingvističnih pravil.

4. 3 Program za sintezo govora

Sinteza govora pri projektu Govorec

Tudi na našem tržišču se lahko pohvalimo s sistemom za pretvorbo besedila v govor, ki se imenuje Govorec in je nastal v okviru Skupine za inteligentne sisteme na Institutu Jožef Štefan. Glavni namen sistema je olajšati branje slepim in slabovidnim osebam pri delu z računalnikom. Za delovanje potrebujete vsaj 8 Mb prostora na disku. Program deluje pod vsemi inačicami operacijskega sistema Windows od 95 naprej. Računalnik pa mora imeti tudi primerno zvočno kartico.

Real speak

RealSpeak je program, ki spremeni tekst v zelo kvaliteten moški ali ženski govor. Njegova dobra lastnost je, da omogoča uporabo v avtomobilskih navigacijskih sistemih, branje zaslonov za slepe in slabovidne ali pa poveča klicne usluge. RealSpeak je zato dobra odločitev ker je prijazen, izrazit, visoko kakovosten jezikovni vmesnik.

- REALSPEAK TELECOM je vodilni program, ki spremeni tekst v čist človeški govor v več kot 20 jezikov in 30 glasov v različnih naglasih in stilih govora.
- REALSPEAK SOLO 4.0 je rešitev, ki je narejen posebej, da poveča kakovost vgrajenih pogovornih aplikacij. Z njim se zagotovi izjemen visoko kakovosten jezik.
- REALSPEAK WORD uporablja izjemen pristop k tekst-v-govor tehnologiji, da doseže največjo kvaliteto govora iz slovarja besed in jezikovnih posebnosti ter omogoča učencem jezika, da slišijo kako morajo biti besede izgovorjene. Zasnovan je na arhitekturi, ki ne potrebuje nobenega snemanja in zasede 3MB prostora za 60,000 besed.
- REALSPEAK MOBILE je družina tekst-v-govor izdelkov, katerih cilj so mobilni aparati, ki vsebujejo SMS in elektronsko sporočanje.

5 PROJEKT GOVOREC

Govorec je sistem za sintezo govora, ki so ga razvili v Skupini za inteligentne sisteme na Institutu Jožef Stefan. Razvit je bil z namenom, da pomaga vsem slepim in slabovidnim ljudem v Sloveniji. Osnovni namen je olajšanje branja slepim in slabovidnim osebam pri delu z računalnikom. Sistem je brezplačen in ga lahko uporablja vsakdo. Sistem je kompatibilen z Okni, kar omogoča uporabo Govorca v katerikoli aplikaciji pri Oknih, če je le-ta Speech API kompatibilna. Govorca je mogoče integrirati v program za

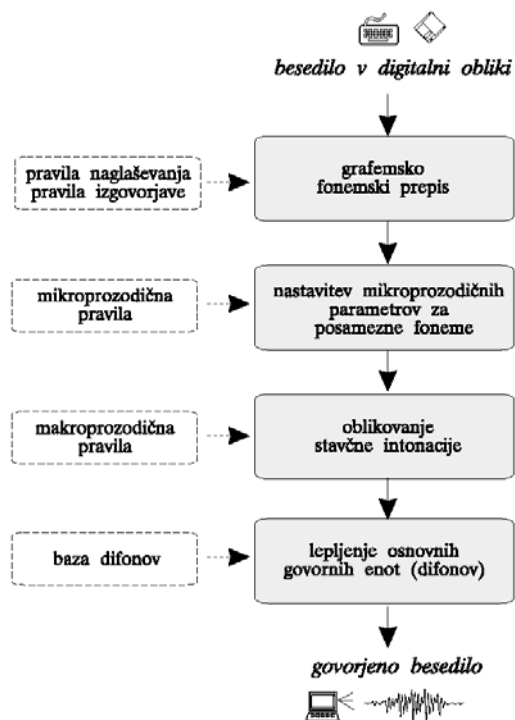
branje Jaws, prav tako pa deluje tudi s programom Narrator, ki je del operacijskega sistema Windows 2000.

Sistem sestavljajo trije neodvisni moduli: modul za grafemsko fonemsko pretvorbo, modul za nastavitev prozoičnih parametrov in modul za združevanja osnovnih govornih enot (difonov). Vhod v sintetizator predstavlja poljubno besedilo. Besedilo se sprva na podlagi pravil in fonemskega slovarja, ki vsebuje 500,000 besed, pretvori v ustrezen fonemski zapis. Sledi določanje osnovne frekvence in trajanja za posamezne foneme, nato se oblikuje še celotna stavčna intonacija. Sam postopek sinteze govornega signala temelji na lepljenju osnovnih govornih enot s pomočjo algoritma TD-PSOLA, ki zagotavlja visoko naravnost sintetiziranega govornega signala. Za osnovne govorne enote so bili izbrani difoni. Govorni signal se nato predvaja na zvočni kartici (SoundBlaster), oziroma se generira datoteka v WAV formatu. Dosedanje meritve kažejo, da je sistem razumljiv nad 95%, vendar delež še naraste, ko se poslušalec navadi na glas sintetizatorja.

5. 1 Delovanje sistema

Sistem sestavljajo trije neodvisni moduli: modul za grafemsko fonemsko pretvorbo, modul za nastavitev prozoičnih parametrov in modul za združevanja osnovnih govornih enot (difonov). Vhod v sintetizator predstavlja poljubno besedilo. Besedilo se sprva na podlagi pravil in fonemskega slovarja, ki vsebuje 500.000 besed, pretvori v ustrezen fonemski zapis. Sledi določanje osnovne frekvence in trajanja za posamezne foneme, nato se oblikuje še celotna stavčna intonacija. Sam postopek sinteze govornega signala temelji na lepljenju osnovnih govornih enot s pomočjo algoritma TD-PSOLA, ki zagotavlja visoko naravnost sintetiziranega govornega signala. Za osnovne govorne enote smo izbrali difone. Govorni signal se nato predvaja na zvočni kartici (SoundBlaster), oziroma se generira datoteka v WAV formatu. Sistem je tudi Microsoft Speech API 4.0 kompatibilen, kar omogoča uporabo Govorca v katerikoli aplikaciji v Oknih, če je le-ta Speech API združljiva. Dosedanje meritve kažejo, da je sistem razumljiv nad 95%, vendar delež še naraste, ko se poslušalec navadi na glas sintetizatorja.

Sinteza govora



Slika : 3 koraki

Grafemsko-fonemska pretvorba

- pretvorba odvisna od jezika (Slovenščina: sorazmerno enostavna, Angleščina: zapletena).
- množica pravil z izjemami.
- fonemski slovarji.
- zamenjava nedovoljenih znakov: q==ku, y==i, w==v, x==ks.
- pretvorba števnikov in datumov.
- okrajšave (RAM, IBM).
- simboli: \$==dolarjev, @=at, +=plus.
- določevanje mesta in vrste naglasa
- za slovenščino je značilno prosto naglaševanje (naglas na kateremkoli zlogu).

Prozodija

- parametri odvisni od: spola in starosti govorca, čustev (jeza, sreča, žalost), pomena stavka (vprašanje, vzklík).
- nastavljanje trajanja in osnovne frekvence fonemov.
- določitev mest in trajanja premorov.

- naglaševanje besed.
- stavčna intonacija.

Trajanje glasu

- odvisno od: naglašenosti zloga, dolžine besede, položaja glasu, poudarjenosti besede, hitrosti govora.
- postopno popravljanje trajanja.
- fonem, beseda, podaljšanje naglašenih zlogov, stavki.
- določitev mesta in trajanja premorov: vezniki (in, pa, ter).
- naglaševanje besed, poudarki.

Osnovna frekvenca

- intonacijske krivulje. odvisne od vrste besede in stavka.
- globalne in lokalne komponente.
- govor ni zaporedje neodvisnih fonemov (zveznost).
- stavčna in besedna intonacija.
- Sinteza govornega signala
- sinteza signala na podlagi fonetičnega zapisa besed.

Artikularotna sinteza

- direktno posnemanje človeških govornih organov.
- potencialno najboljša.
- najzahtevnejša tako implementacijsko kot računsko.

Formantna sinteza

- opis s formantnimi frekvencami in amplitudami.
- vzpodbujanje niza filtrov z virom.
- periodično vzbujanje -zvaneči glasovi.
- sum - nezvaneči glasovi.
- običajni parametri F_0 (frekvenca), A_0 (amplituda), F_n , Q_n (formantne frekvence), V_L , V_H (zvaneče komponente amplitude) F_L , F_H (nezvaneče komponente amplitude).

Sinteza z združevanjem

- združevanje osnovnih govornih enot.
- največja naravnost sintetiziranega govora.
- omejenost na enega govorca, en glas.
- velika baza.
- izbira osnovne enote (fonem, difon, trifon, zlog, beseda).
- krajše enote: težja sinteza, označevanje, manjša naravnost, večja kontrola, manjša baza.
- daljše enote: lažja sinteza, lažje označevanje, večja naravnost, manjša kontrola, večja baza.
- težave: nezveznosti na mestih lepljenja, pomnilniške zahteve, označevanje.

Izboljšave

- nova baza.
- boljša pravila (učenje na slovarju).
- boljša prozodija.
- podarjeno Zvezi slepih in slabovidnih Slovenije.
- <http://ai.ijs.si/govorec/>.

5. 2 Uporaba in navodila sistema Govorec

Sistem Govorec vsebuje dva programa: SiTalk in SiRead. Vsak ima svoje okence. SiTalk se uporablja za splošno delo z besedili, SiRead pa stalno teče v ozadju in ob pritisku določenih tipk prebere označeno besedilo.

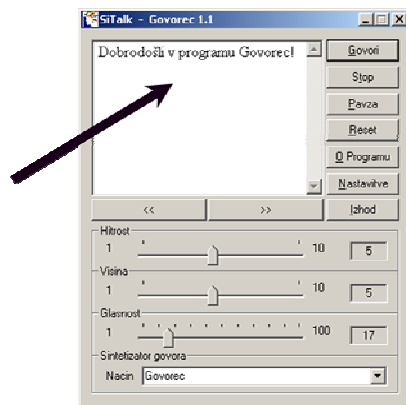
• SiTalk - Govorec 1.1

SiTalk je splošen program za branje besedil. Primeren je za branje sproti vpisanega besedila. Besedilo najprej natipkamo v okencu programa, nato pa ga program prebere. Program omogoča nastavitve hitrosti branja, frekvence govornega signala in glasnosti. Z drsniki Hitrost, Višina in Glasnost nastavljamo hitrost branja, frekvenco glasu in glasnost. Pri vrednostih Hitrost in Višina pomeni manjša vrednost počasnejše branje oz. nižjo frekvenco (nižji glas).

Program omogoča uporabo poljubnega sintetizatorja govora, ki je Microsoft Speech API 4.0 združljiv. Na dnu okna v področju Sintetizator Govora (Mode) je možna izbira sintetizatorja govora. Število izbir je odvisno od števila sintetizatorjev govora nameščenih na računalniku. Privzet je slovenski sintetizator govora Govorec.

Primer uporabe programa SiTalk:

Zagnati je potrebno program SiTalk (sitalk.exe). V prostor, označen s puščico, kliknemo z miško in natipkamo poljubno besedilo. Besedilo lahko tudi prenesemo v okence kot običajno v Windowsih.



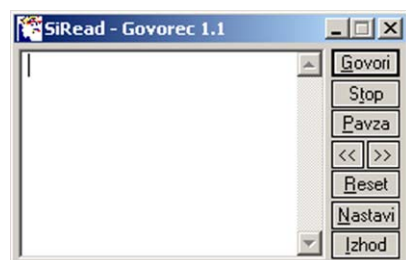
Kliknemo gumb Govori ali pritisnemo kombinacijo tipk Alt+g in program SiTalk prične z branjem besedila. Med branjem uporabimo gumbe Pavza in Stop oziroma ustrezne kombinacije tipk za začasno zaustavitev ali prekinitvev branja besedila.

- SiRead - Govorec 1.1

SiRead je program za branje besedila, ki se trenutno nahaja v odložišču (clipboard). Program je vedno nad vsemi ostalimi okni, pa tudi ostale funkcije so posebej prilagojene slepim in slabovidnim, tako da z nekaj tipkami začnemo brati izbrano besedilo v poljubnem programu. Za uporabo tega programa ni potrebna uporaba miške.

Primer uporabe programa SiRead

Zagnati je potrebno iread.exe.

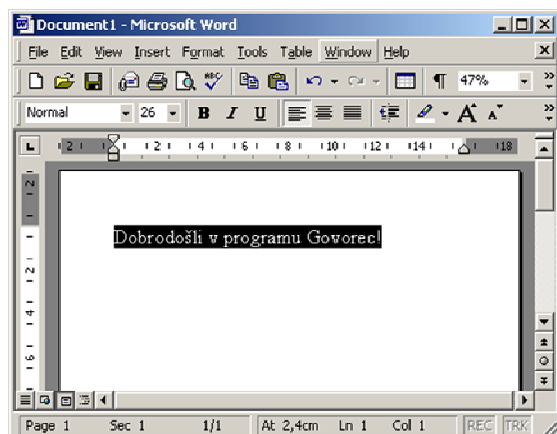


Odpremo program za delo z besedili (WinWord, Notepad itd.). Lahko odpremo tudi program za delo z internetom (Internet Explorer, Netscape) in se postavimo na neko stran.

Označimo besedilo, ki ga želimo prebrati. To lahko storimo tako, da držimo pritisnjeno tipko Shift in s smernimi tipkami označimo besedilo, lahko pa pritisnemo Ctrl + a, v tem primeru bomo označili celotno besedilo.

Pritisnemo Ctrl + C (Copy). S to kombinacijo tipk skopiramo označeno besedilo v odložišče.

Pritisnemo Ctrl + Shift + R in besedilo, ki smo ga označili, se prebere. Branje lahko prekinemo s kombinacijo tipk Ctrl + Shift + T.



ZAKLJUČEK

Jezikovne tehnologije so področja, ki imajo zagotovljeno visoko mesto v prihodnosti in sam tehnološki razvoj od sistemov za sinteze govora, razpoznavanja in strojno prevajanje. Te tehnologije so postale del današnje družbe si ne more več predstavljati živeti brez njih, ker omogočajo hitrejšo komunikacijo, izobraževanje, združevanje narodov.

Veliko je še neraziskanih področij, kljub velikemu napredku z gigantskimi dosežki. Dve vodilni podjetji Systran in Nuance se trudita in težita po razvoju novih in vse bolj naprednih tehnologij, da bi zagotovila svojim uporabnikom orodja in okolja, ki bi zadovoljila njihovim potrebam. V Sloveniji pa si za vse razvite jezikovne tehnologije v slovenskem jeziku, zasluži častno mesto podjetje Amebis in skupina, ki je naredila program Govorec.

LITERATURA IN VIRI

ROMIH, M.: Jezikovne tehnologije. Življenje in tehnika, januar 2003

ERJAVEC, T.: Označevanje korpusov. Jezik in slovstvo, let. 48, št.3/4

STABEJ, M.: Jezikovne tehnologije in jezikovno načrtovanje. Jezik in slovstvo, let. 48, št.3/4, maj/avg 2003

ERJAVEC, T.: Jezikovne tehnologije za slovenski jezik. Mednarodna konferenca

KAČIČ, Z.: Komunikacija človek – stroj, Maribor : Fakulteta za elektrotehniko, računalništvo in informatiko, Inštitut za elektrotehniko, 1995

SISTEM PRESIS: <<http://presis.amebis.si/prevajanje/>> , 26. 4. 2006

PROJEKT TC-STAR: <<http://www.tc-star.org/>>, 1. 5. 2006

SAY I CAN SPEECH RECOGNITION: <<http://www.sayican.com/>>, 26. 4. 2006

SISTEM GOVOREC: <<http://ai.ijs.si/govorec/>>, 1. 5. 2006

READ PLEASE: <<http://www.readplease.com/>>, 26. 4. 2006

WIKIPEDIA: <www.wikipedia.org>, 26. 4. 2006

The Festival Speech Synthesis System: <<http://www.cstr.ed.ac.uk/projects/festival/>>, 24. 4. 2006

The MBROLA Project: <<http://www.tcts.fpms.ac.be/synthesis/>>, 26. 4. 2006

ATRC: <<http://www.utoronto.ca/atrc/reference/tech/voicerecog.html>>, 24. 4. 2006

MICROSOFT: <<http://www.microsoft.com/speech/evaluation/techover>>, 24. 4. 2006

SINTEZA GOVORA:<http://laps.fri.uni-lj.si/dps_arhiv/seminarske/indihar/sinteza.htm>, 1. 5. 2006

AMEBIS:<<http://www.amebis.si/>>, 1. 5. 2006

