

Razvoj multimedijskega večjezičnega sistema za učenje in vadbo pravilnega izgovora govorno in slušno prizadetih otrok

Robert Veronik*, Zdravko Kačič*, Klara Vicsi**, Ferenc Csatari**, Peter Roach***,

Anne-Marie Öster⁺, Peter Barczikay⁺⁺

*Fakulteta za elektrotehniko, računalništvo in informatiko, Univerza v Mariboru

Smetanova 17, 2000 Maribor, Slovenija

**Technical University of Budapest, Hungary

***University of Reading, United Kingdom

⁺Kungl. Tekniska Högskolan, Sweden

⁺⁺RCS – Robot Control Software, Hungary

robert.veronik@uni-mb.si, kacic@uni-mb.si

Povzetek

Predstavljen je razvoj multimedijskega večjezičnega sistema za učenje in vadbo pravilnega izgovora govorno in slušno prizadetih otrok. Razvoj sistema poteka v okviru raziskovalnega projekta SPECO programa INCO -Copernicus. Osnovni namen projekta je razvoj sistema za učenje in vadbo, ki bo pomagal slušno in govorno prizadetim otrokom, da bodo sposobni nadzorovati in izboljšati lastno izgovorjavo. Sistem pomaga otrokom tako, da lahko ti istočasno primerjajo slikovno predlogo normalne referenčne izgovorjave z lastno. Program je zgrajen iz dveh modulov. Glavni del programa predstavlja splošni jezikovno neodvisni sistem za učenje govora, drugi modul je urejevalnik baze. Sistem bo v osnovi razvit za štiri evropske jezike. Osnovni del programa je razdeljen na dva dela. Prvi je namenjen učenju izgovorjave pripornikov, drugi pa za učenje izgovorjave vokalov.

Abstract

The paper describes a "SPECO - Multimedia Multilingual Teaching and Training System for Speech Handicapped Children", which is developed in the framework of the INCO-Copernicus research project SPECO. The aim of the project is to develop an audio-visual training and teaching system, which will help hearing and speech handicapped children to control and improve their speech pronunciation. The system helps children to improve their speech production in the way that they can simultaneously compare the visual pattern of the normal acoustic speech signal with their own one. The software system is built from two modules. The main module is a general language independent analysis system and the other is database editor. The system will be developed for four languages. The general language independent analysis system consists of two parts. The first is for teaching fricatives pronunciation and second for vowels.

1. Predstavitev sistema

Problem gluhotе je mnogokrat obravnavan kot manjša telesna pomanjkljivost, še posebej s strani ljudi, ki se poklicno ne srečujejo s tem problemom. Toda naglušnost predstavlja drugo najvišjo stopnjo invalidnosti. To pa je zaradi tega, ker je lahko normalni razvoj, še posebej otroka, zelo otežen in brez ustrezne terapije obstaja nevarnost duševne zaostalosti. Z ustreznim zdravljenjem in učenjem lahko to pomanjkljivost odpravimo. Pri slušno prizadetih se poleg slušne prizadetosti še mnogokrat pojavi nemost, ker zaradi prizadetega sluha in pomanjkanja povratne informacije, oseba ne zna pravilno uporabljati govornih organov. To navadno ni zaradi tega, ker bi bili njihovi govorni organi poškodovani ali pa niso zmožni izgovorjave glasov, ampak zaradi tega, ker svojih govornih organov sploh ne znajo uporabiti.

Govorna komunikacija je podobna verigi. Obstaja več dogodkov, ki povezujejo govorca in poslušalca. Možgani kontrolirajo govorne organe govorca. Premikanje govornih organov pomeni tvorjenje določenih zvokov. Ti prenesajo informacijo preko zraka do poslušalčevih ušes oziroma posredno do njegovih možganov. Če je eden izmed členov te verige poškodovan ali manjka, je komunikacija otežena.

Govorec ob procesu izgovorjave določenega glasu dobi povratno informacijo preko svojih slušnih poti. Tako

govorec vedno primerja kvaliteto izgovorjenega glasu z zeleno podobo glasu. Venomer izvaja določene popravke z namenom doseči zeleno izgovorjavo.

Ena izmed osnovnih zahtev za vpis govorno in slušno prizadetih otrok v osnovno in nadaljnje izobraževanje je nedvomno razvoj pravilne izgovorjave in sposobnost razumevanja govora.

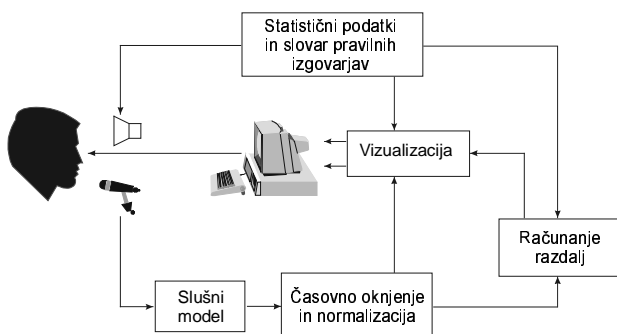
2. Metode učenja govora

Otroci, ki normalno slišijo, uporabljajo za učenje govora predmetno usmerjen pristop – s pomočjo akustične povratne informacije odkrivajo kako kontrolirati svoje govorne organe. Slušno prizadeti otroci imajo običajno težave z uporabo predmetno usmerjenega učenja, zato za učenje govora pri njih največkrat uporabljajo tradicionalno metodo, ki jo imenujemo tudi procesno orientiran pristop. Terapevt daje pacientu navodila, kako uporabljati govorne organe v procesu produkcije glasu. Običajno pri tem uporablja ogledala. Toda takšen način učenja je težek, ker mnoge značilnosti artikulacije niso vidne na ta način. Akustično ima vsak zvok svoje posebne lastnosti, toda vizualno je nekatere glasove v procesu izgovora nemogoče ločiti. Nekateri glasovi imajo skoraj identične premike vidnih govornih organov, medtem ko pri drugih artikulacija sploh ni opazna.

S SPECO projektom želimo namesto tradicionalne procesno orientirane metode učenja ponuditi predmetno orientirano metodo učenja tudi slušno prizadetim otrokom (Vicsi, 1999, Öster, 1999). Razvoj govorne tehnologije omogoča računalniško obdelavo izgovarjave in neposredno primerjavo glasovne izgovarjave s slikovnim prikazom le-te. Sistem je zasnovan tako, da procesira izgovorjene glasove in slikovno prikaže vsebino te izgovarjave. Uporabniku ponuja pomoč pri izgovarjavi s pomočjo primerjanja slikovnih predlog normalno izgovorjenih referenčnih glasov z njegovo lastno izgovarjavo. S slikovno predstavitvijo poskušamo nadomestiti okvarjen slušni kanal pacienta z zdravim vidnim kanalom. Slikovna predstavitev predstavlja povratno informacijo za uporabnika, ki ne more sprejeti akustične povratne informacije (Povel, 1991). Kljub temu je ob slikovni predstavitvi izgovorjenega glasu istočasno podan tudi zvočni posnetek referenčne izgovarjave. Uporabnik s pomočjo slikovne predstavitve odkriva, kako kontrolirati svoje govorne organe in se s tem poskuša naučiti pravilne izgovarjave. Z vidika vizualne predstavitve govora ni najpomembnejši proces izgovarjave, ampak kvaliteta informacije, ki je posredovana drugi osebi. Sistem je zasnovan na osnovi najnovejših tehnologij, vendar sledi korakom tradicionalnega učenja govora. Uporabljen je pristop učenja od najosnovnejših korakov, ki zajemajo predgovorne vaje, do učenja izgovarjave posameznih glasov, vaj glasovnih zvez, izgovarjave besed in kratkih stavkov ter kontrastnih parih.

3. Zasnova sistema

Celotni sistem je zasnovan modularno. Tako je moč ga uporabiti za odpravljanje različnih govornih težav in v različnih jezikih, z uporabo ustreznih govornih baz. Na sliki 1 je prikazana zasnova sistema.



Slika 1: Zasnova sistema

Sistem je sestavljen iz dveh modulov. Glavni del sistema predstavlja splošni jezikovno neodvisni sistem za učenje govora. Drugi modul je jezikovno neodvisen urejevalnik baze podatkov. Uporabnik preko mikrofona posreduje računalniku izgovarjavo. Sledi akustično procesiranje govora. Metoda procesiranja je jezikovno neodvisna in je vgrajena v urejevalnik baze. Opisan sistem se razlikuje od sistemov za učenje izgovarjave, ki so narejeni na osnovi sistemov razpoznavanja govora. Običajno ti sistemi samo razlikujejo med dobro in slabo izgovarjavo. Pri tem je cilj zgolj oceniti izgovarjavo kot pravilno ali nepravilno, kar pa ne omogoča uporabniku samostojnega učenja brez neke

dodatne pomoči. Cilj predstavljenega multimedijskega sistema je pomoč in vodenje pacienta skozi proces izgovarjave od začetne slabe izgovarjave, preko vmesnih popravilanj, do pravilne izgovarjave.

S pomočjo urejevalnika baze lahko ustvarjamo slovarje z različnim naborom besed za različne jezike in različne govorne težave. Trenutno poteka razvoj sistema za štiri jezike: angleški, madžarski, švedski in slovenski. Za vsak jezik so urejeni posebni slovarji izgovarjav, ki vsebujejo izbrane glasovne zveze, besede in stavke. Te izgovarjave uporabljamo kot referenčne predloge. Sistem je zasnovan tako, da lahko slovarji za posamezni jezik vsebujejo izgovarjave od osnovnih glasov do kompleksnih stavkov. Sistem je sestavljen iz dveh modulov: prvi modul je namenjen učenju pripornikov in zapornikov, drugi učenju vokalov.

Uporabnik s pomočjo slikovne predstavitve odkriva, kako kontrolirati govorne organe. S pomočjo uporabe organov vida poskušamo predstaviti pravilno izgovarjavo, vendar ob slikovni informaciji sistem podaja tudi avdio informacijo sinhronizirano s slikovno.

2.1 Metoda procesiranja

Na nivoju procesiranja govornega signala je uporabljen preprost avditorni model. Izhod modela je takoimenovani kohleogram. Kohleogram prikazuje spreminjanje spektralnih značilnosti govora v odvisnosti od časa. Avditorni model pokriva frekvence od 80 Hz do 8kHz. Sistem izračuna parametre, ki so pomembni za kvaliteto izgovarjave: sprememba glasnosti po času, sprememba kohleograma po času, razlike v hitrosti izgovarjav, določitev osnovnih harmonskih komponent (Vicsi, 1981). Vključeni so dodatni algoritmi za določanje začetkov in koncev izgovarjav ter računanje razdalj med referenčnimi in izgovorjenimi glasovi (Vicsi, 1999).

Zelo pomemben je slikovni prikaz izgovarjave, ki mora biti jasen in razumljiv, še posebej otrokom. Program je namenjen za učenje govora otrok v starosti do 10 let. Potekajo tudi raziskave na področju normalizacije govora, ki bi omogočile uporabo starejšim otrokom in mladini (Ogner, 1999).

2.2 Akustično procesiranje

Akustično procesiranje je izvedeno s preprostim avditornim modelom, ki upošteva najpomembnejše značilnosti sluha v frekvenčni domeni, poznanih predvsem iz teorije psihoakustike (Zwicker, 1982). Model uporablja 20 ozkopasovnih filtrov, ki pokrivajo območje od 80 Hz do 8 kHz. Filtri imajo asimptotični nagib za čim natančnejše posnemanje karakteristike človeškega ušesa. Časovne konstante so približno petkrat večje kot je dolžina periode osnovne frekvence. Nadaljnja izračunavanja kohleograma se izvajajo po postopkih opisanih v (Zwicker, 1980).

3. Uporaba sistema

Na sliki 2 je prikazan osnovni zaslon programa, kjer uporabnik izbere predgovorne vaje, vaje učenja vokalov, ali vaje učenja pripornikov in zapornikov.



Slika 2: Osnovni zaslon programa z izbiro vaj

Učenje in vadba sta izvedena v obliki igre, kjer otrok s kontrolo svoje izgovarjave poskuša doseči čim boljše oceno. Kot merilo uspešnosti opravljene vaje je uporabljena rastoča rožica - bolj uspešna je izgovarjava, večja in lepša je rožica. Zelo pomembno je, da otroka motiviramo za sodelovanje pri učenju (Csatari, 1999). Zato je celoten proces učenja zasnovan tako, da vodi uporabnika od osnovnih začetnih predgovornih vaj, preko enostavnih izgovarjav izoliranih fonemov, do besed z različnimi položaji fonemov, in na koncu do krajših stavkov in minimalnih parov. Program sledi in upošteva tradicionalne metode učenja izgovarjave slušno in govorno prizadetih otrok. Na začetku se morajo otroci najprej poskušati naučiti sposobnosti kontrole svojih govornih organov. Zato so v programu dodane predgovorne vaje, ki obsegajo vaje glasnosti, vaje zvokov, vaje višine, ritma in vadbo zvonečih - nezvonečih glasov.

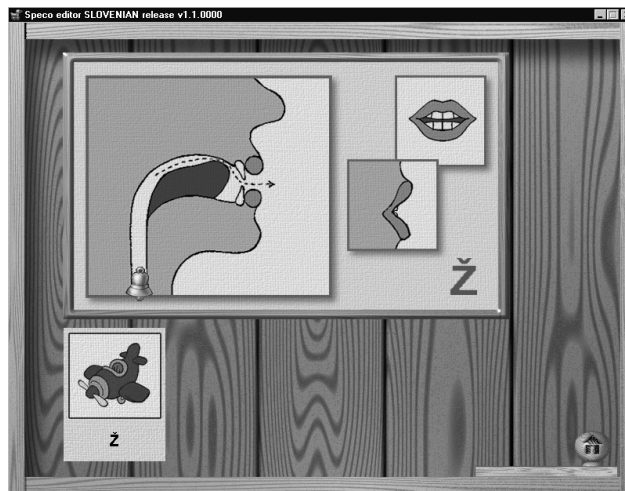


Slika 3: Vaja kontrole glasnosti

Primer predgovornih vaj je prikazan na sliki 3 - pri tem otrok vadi kontrolo glasnosti. S spreminjanjem glasnosti svojega glasu mora nadzorovati premikanje žoge – cilj je pripeljati žogo skozi odprtine v jabolku, ko žoga potuje od leve strani zaslona proti desni.

Vizualno podobno so zasnovane vaje za učenje spreminjanja ritma izgovorjave in višine glasu. Za pomoč

terapevtu pri razlagi oblikovanja govornih organov za pravilne izgovarjave posameznih glasov, so v programu na voljo tudi slike artikulacije za posamezne glasove.



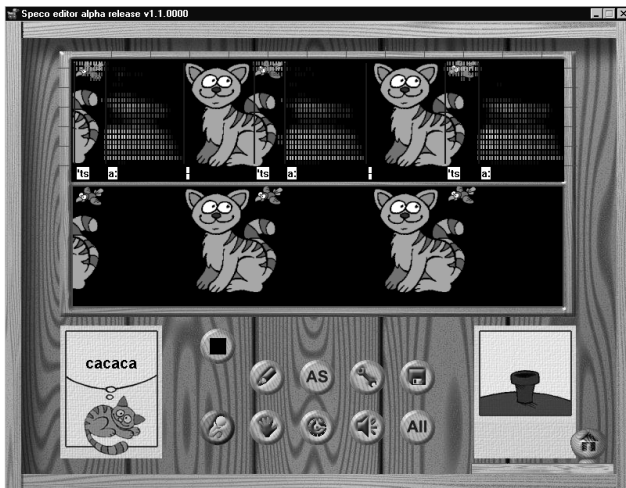
Slika 4: Slika izgovarjave glasu /ž/

Na slikah so podrobno predstavljeni položaji govornih organov pri izgovarjavi posameznih glasov. Na sliki 4 je prikazan primer artikulacije glasu /ž/. Ko pacient obvlada osnovne vaje za kontrolo govornih organov, sledijo vaje izgovarjave posameznih glasov. Pri tem sistem prikaže na zaslonu frekvenčni spekter, ki se izračunava in izrisuje na zaslonu vsakih 20 ms. Oblika spektra je odvisna od posameznih glasov - pri pravilni izgovarjavi mora spekter besede biti znotraj določene predstavljene poti, ki predstavlja največjo in najmanjšo spektralno mejo pravilno izgovorjenih fonemov.



Slika 5: Vaje izgovarjav posameznih glasov

Slika 5 prikazuje primer vadbe izgovarjave glasu /z/. Pri vajah glasovnih zvez uporabnik vadi izgovarjavo fonemov v različnih položajih in kontekstih (pri tem so uporabljene glasovne zveze konzontan - vokal, vokal - konzontan - vokal...). Sledijo vaje izgovarjave besed (slika 5) - pri konzontanah so izbrane takšne besede, da lahko uporabnik vadi izgovarjavo fonema na začetku, na sredini in na koncu besede. Pri vokalih pa lahko vadi v besedah z enim in več samoglasniki v besedi.



Slika 6: Vaje izgovarjav besed

Vsaka slika je predstavljena z jasnimi slikovnimi prikazi pomena besede tako, da tudi otroci, ki ne znajo brati, lahko sami ugotovijo pomen besede in se ob tem tudi učijo branja črk. Za pomoč pri pravilni izgovarjavi besed je na kohleogramu ob spektralnem prikazu izgovorjene besede, prikazana še slika v ozadju. Ta slika določa območje spektra fonema, ki ga uporabnik vadi. S prikazom slike v ozadju omogočimo, da uporabnik sam ugotovi, kako naj interpretira kohleogram, oziroma kateri del kohleograma je pomemben za učenje in pravilno izgovarjavo posameznega fonema – lahko sam določi napačno izgovarjavo fonema in jo skuša popraviti. Po nekaj vajah je uporabnik sam sposoben razlikovati uspešne izgovarjave od neuspešnih. Za uspešno vadbo določenega glasa mora uporabnik s svojo izgovarjavo fonema pokriti določen del slike, ali pa mora nek del slike ostati nepokrit. Na sliki 6 vidimo izgovarjavo glasovne zveze /cacaca/. Ker gre pri tem za učenje izgovarjave pripornika c, mora otrok izgovoriti ta glas tako, da s svojo izgovarjavo pokrije samo sliko ptice. Ptica ponazarja področje spektra pripornika c v fazi odpiranja. Slika muce pa mora ostati pri pravilni izgovorjavi nepokrita.

Uporaba podobnega sistema za nemški in madžarski jezik, ki je služil kot osnova razvoju opisanega sistema, je v praksi pokazala, da je moč z njegovo pomočjo doseči hitrejšo učenje pravilnega izgovora in tudi boljšo kvaliteto v primerjavi s klasičnimi metodami učenja.

Zaključek

Uporaba računalnikov omogoča nov način učenja in vadbe izgovora pri slušno prizadetih otrocih. Ena izmed največjih prednosti je možnost slikovne predstavitve izgovarjav, oziroma zamenjava akustične povratne informacije s slikovno. Program je posebej primeren za otroke z kohlearnim implantom, saj jim omogoča hitrejšo učenje pravilne izgovarjave. Program bi lahko v splošnem uporabili tudi za učenje tujih jezikov, predvsem učenje glasov specifičnih za določeni jezik.

V letu 2000 in 2001 je predviden razvoj drugega dela programa za učenje pravilnega izgovora vokalov in tudi obsežna testiranja sistema v sodelovanju s Centrom za korekcijo sluha v Mariboru.

Literatura

- Povel D.J (1991), "The Visual Speech Apparatus: Theoretical and Practical Aspects" *Speech Communication*, Vol. 10. str. 59-80
- Zwicker E. (1982) *Psychoakustik*. Springer Verlag. Berlin
- Vicsi K (1981) "The most Relevant Acoustic Microsegment and its Duration Necessary for the Recognition of Unvoiced Stops" *ACOUSTICA*. Vol. 48, str. 53-58
- Zwicker E. and Terhardt R. (1980), "Analytical Expressions for Band Rate and Critical Bandwidth as a Function of Frequency", *J. Soc. Am.* Vol. 68. str 1523
- Ogner M. and Kačič Z. (1999) "Speaker Normalization for Audio-Visual Articulation Training", 6th European Conference On Speech Communication And Technology - Eurospeech '99, str. 579-582, Budimpešta, Madžarska
- Vicsi, K. - Csatari, F. - Bakcsi, Zs. - Tantos, A.(1999) "Distance Score Evaluation of the Visualised Speech Spectra at Audio-Visual Articulation Training", 6th European Conference On Speech Communication And Technology - Eurospeech '99, str. 1911-1914, Budimpešta, Madžarska
- Vicsi K., Roach P., Oster A., Kačič Z., Barczikay P., Sinka I., (1999) "SPECO - A Multimedia Multilingual Teaching and training System for Speech Handicaped Children, 6th European Conference On Speech Communication And Technology - Eurospeech '99, str. 859-862, Budimpešta, Madžarska
- Öster M., Vicsi, K., Roach P., Kačič Z. - Barczikay, P. (1999) " A Multimedia Multilingual Teaching and training System for Speech Handicaped Children, FONETIK 99
- Csatari F., Bakcsi Zs., Vicsi K. (1999) "A Hungarian Child Database for Speech Processing Applications" 6th European Conference On Speech Communication And Technology - Eurospeech '99, str. 2231-2234, Budimpešta, Madžarska