

KOMMUNIKATIIVSE TEHISKÕNE POOLE:
ARENUSPROBLEEME JA -VAHENDEID

Imre Siil, Arvo Ott

1.0. Tehiskõne loomisel aastakümneid kestnud pingutused on hakanud vilja kandma. Praegu on olemas seadmeid, mis võivad küllalt arusaadavalt ette lugeda peasegu ortograafiliselt kirja pandud teksti, olgu see kuitahes keerulise struktuuri-ga, sisaldagu lühendeid, numbreid, muid sümbboleid. Tundmatu teksti sisestamisest tehiskõneseadmesse kuni sellest tekstist loodud kõne kuulamiseni väljundis kulub kõige rohkem mõni sekund - töö toimub niisiis loomulikus ajamastaabis (reaalajas).

Väikestel arvutitel on nimetatud tulemusi andnud kõne reegelsünteesi tehnoloogia rakendamine. Reegelsünteesil kujundatakse kõnetrakti mudeli juhtimiseks vajalik sagedus-, energieetilistest ja ajalistest parameetritest koosnev informatsioon tingimuslike tegevusreeglite kujul formaliseeritud lingvistiliste seaduspärasuste alusel (vt. 2.3.1). Tuntumad on D. H. Klatti ja S. R. Hertzi sellealased tööd (1), (2), (3). Valdav osa käesolevast artiklist ongi pühendatud reegelsünteesi arendamisvõimaluste käsitlemisele.

1.1. Suvalisest tekstist kiiresti kõnet genereerivad seadmed annavad hea võimaluse arvutisüsteemide (üldisemalt: automatiseeritud süsteemide) kõneväljundi teostegemiseks, sealjuures üksikute teadete esitamise kõrval ka keerukamate loomulikku keelt kasutavate (formaalse) kommunikatsiooni moodulite loomiseks. Tuleb paraku üldistavalt nentida, et praegu suudavad loomulikku keelt informeerimiseks kavakindlalt appi võtta vähesed praktiliseks tööks pakutavad süsteemid. Mõistetav saab see olla ehk vaid professionaalse ettevalmistuse saanud arvutikasutajale orienteeritud süsteemide korral. Väheste kogemustega programmeerija või süsteemi kasutaja vajab arvutitöös tekkivate raskuste ületamiseks ikkagi loomulikku keelt valdavat assistenti.

Kõnet peetakse efektiivse suhtlemise organiseerimisel vajalikuks kõigi arvutikasutajate jaoks. Tihti viidatakse kinnituseks A. Chapanise'i korraldatud katsetele, mille käigus katseisikutele selgitati tegevusülesandeid mitmesuguseid kommunikatsiooniviise ja nende kombinatsioone kasutades.

Katsealused said ülesannetest kiiresti aru eranditult vaid siis, kui suheldi kõneldes (4).

Niisiis võiksime sedastada, et kõnelev arvuti on samm edasi inimesega kohanduva, intelligentse masina loomisel.

1.2. Arvuti kõneväljundi teostamisel tuleb peale kõneseadme loomist - seadmesse saab kõne genereerimiseks sisestada lähtematerjali sobivas kirjaviisis, aga ka digitaalsel või analoogkujul - tegelda selle seadme kohandamisega konkreetsele arvutisüsteemile, mille väljundiks ta on mõeldud. See ei ole üksnes puhttehniline töö, vaid siin on vaja lahendada ka kõneseadme ümberhäälestamisega ühenduses olevad küsimused. Järgneb kõneväljundi repliikide ettevalmistamine ja nende seostamine teiste väljundkanalite ning akustilise fooniga, seda kõike vastavalt kasutaja poolt seatud ülesannetele ja arvutisüsteemi iseärasustele.

Teostusprotsessis on ilmnunud mitmeid probleeme. Osutame neist olulisematele.

1.2.1. Tehiskõne võib olla peasegu arusaadav, kuid kõne kvaliteet ei rahulda kuulajaid, s.t. pakutavat kõnelemismaneeri ei taheta aktsepteerida.

Vähem ilmutatakse rahulolematust niisuguse tehiskõne korral, mis on saadud loomuliku kõne etalonfragmente kasutavate analüüs-sünteesseadmetega ehk kõnekodeerimisseadmetega (6). Paraku nõuavad need seadmed heakvaliteedilise, naturaalsele kõnele sarnaneva tehiskõne tekitamiseks suurte arvutiressursside kulutamist isegi väikese sõnavaru puhul, samuti ulatuslikku analüütilist ettevalmistustööd. Sellepärast ei ole nende praktiline kasutamine suuremate ülesannete jaoks mõeldav.

Reegelsünteeskõne madala kuulajahinnangu põhjused pole veel päris selged. Oletatavaid põhjusi on:

- sellise tehnoloogiaga saadud kõnesignaali akustilis-foneetiline vaesus loomuliku kõne otse pillava liiasuse kõrval. Seetõttu on kuulaja sunnitud vajaliku informatsiooni kättesaamiseks kõnesignaalist ning selle edasiseks töötlemiseks kõrgematel kognitiivsetel tasemetel tugevasti kontsentreeruma, sealjuures oma mälu koormates ja ühtlasi tavalisest enam aega kulutades (5);

- kõne akustiliste või artikulatoorsete kirjelduste puudulikkus, mis eelkõige tuleneb vastava teadmise väljatöötamise ja kirjeldusprotsessi tehnilisest keerukusest;

- kõne prosoodiat kujundava juhtinformatsiooni vähesus. Kuulajahinnangu parandamise teid on muuhulgas tehiskõne kui loomuliku keele n.ö. masindialekti õpetamine, kuulaja tajumisstrateegiate muutmine spetsiaalsete treeningute käigus.

1.2.2. Kõnelevast suhtluspartnerist kujundab inimene kiiresti tervikliku psühholoogilise ettekujutuse - nimetame seda partnerimudeliks. Järelikult on kõneväljundi loomisel tarvis orienteeruda varem väljaselgitatud kasutajakategooriale ja võimalikult täpselt kujundada kõneleva süsteemi kui partneri roll selle suhtes, ning muidugi saadud mudel järjekindlalt ellu viia - nii et kasutaja hakkaks süsteemi usaldama. Intuitiivselt on mõistetav, et pole arukas öelda alati ühte ja sedasama ühesugusel viisil algajale ja vilunud kasutajale, rääkimata sellest, et täiesti spetsiifilist kõneväljundit vajab pime arvutikasutaja. Lisagem veel, et mõeldav on ka kasutajale mitme suhtluspartneri illusiooni loomine.

Niisiis, kasutajaid tuleb paratamatult diferentseerida, iga kasutajakategooria toob aga kaasa oma nõuded kõneväljundi organiseerimisele. Sellega seoses selgitamist vajavaid küsimusi on rohkesti, näit.:

- mil määral on arvuti partnerirolli loomisel vaja tunda kasutaja sotsioloogilisi ja sotsiolingvistilisi parameetreid;
- kuidas mõjutab leksika, morfoloogiline vormistatus ning süntaktiline struktuur repliigi aktsepteeritavust, milliseid kõnekeele iseärasusi on kasulik arvesse võtta (7). Võib-olla õnnestub aeg-ajalt hakkama saada ilma n.ö. ortogrammatiliste lauseteta ning sel kombel lihtsustada repliikide masingeneereerimist või -kompileerimist, ehkki - ilmselt tuleb sel juhul kompensatsiooniks rohkem tegelda kõne prosoodilise kujundamisega;

siia liitub osaliselt ka küsimus sellest, kuidas kõneväljundi kaudu realiseeritavate vahenditega - aga need on peamiselt just lingvistilised vahendid koos akustiliste taustnähtustega, ajafaktori arvestamisega - kujundada inimeses süsteemi suhtes selle edukat kasutamist soodustavaid hoiakuid;

- kuidas mingi kasutajakategooria suhtes lahendada tehiskõne kvaliteedi probleeme. Kui on vaja säästa arvutiressursse ja kõne genereerimise aega, võib pärast kuulaja kohanemist tehiskõne iseärasustega kõne kvaliteeti mõningal määral alandada. Siiski tuleb arvestada, et ebatavalises situatsioonis töötavale ning keerulist ülesannet lahendavale kasutajale esitatavad teated peavad olema hästi tajutavad, küllastatud

mitmesuguse informatsiooniga (5). Peaaegu alati on äsjakirjeldatud olukorras nn. vähekogenud arvutikasutaja. Lisame, et kõne kvaliteedi modifitseerimisega võib väga hästi tegelda ka puhtalt stilistilistel kaalutlustel.

Osutame veel, et mitme väljundkanali rakendamisel tuleb ilmselt uurida ka seda, kuidas jaotada ülesandeid nende vahel. Nii võib keerulise struktuuriga informatsiooni esitava visuaalse väljundi kõrval tehiskõne funktsiooniks olla kasutaja tähelepanu suunamine, nõuannete ja selgituste jagamine, suhteliselt lihtsama ülesehitusega lühinfo andmine, samuti süsteemi n.ö. emotsionaalsete reaktsioonide esitamine, abistamaks kasutajat süsteemiga töötamise võtete otstarbekuse hindamisel.

Kujundades ainult informeeriva väljundiga süsteemi asemel konversatsioonisüsteemi, hõlmatakse eespool toodud probleemid kasutajakategooriat arvestavate suhtlusstrateegiate väljatöötamisel, kommunikatsiooni dünaamika modelleerimisel. Konversatsioonisüsteemis on võimalik ka interaktiivne käitumine. Interaktsioon tähendab vahelesegamist partneri plaanipärasesse suhtlusstrateegiasse, kusjuures vahelesegamise tulemusel võib partner olla sunnitud oma strateegiat muutma (8). Interaktiivses süsteemis ei pea olema rangelt vahelduvat teadete vahetust.

1.2.3. See, milliseid ülesandeid nende hulgast, mis teatud partneri puhul oleksid vajalikud, saab anda kõneväljundile ning milliseid kommunikatsioonistruktuure õnnestub ühtlasi luua, sõltub praktikas eelkõige arvutisüsteemi võimsusest. Seni on kõnet kasutatud peamiselt vaid üksikute, omavahel nõrgalt seotud infofragmentide edasiandmiseks. Süsteemi mingit sisemist seisundit kirjeldab tavaliselt üks ja seesama teade. Loomuliku suhtlemise illusiooni suurendavate suhtlus-tsüklite loomine kasvatab süsteemi suhtlemisega tegeleva osa, interfeisiosa töömahukust tunduvalt. Tekib reaalne oht, et psühholoogiliselt vastuvõetav suhtlusmoodul nõuab endale suurema osa süsteemi ressursidest.

Eelnevas me ei rääkinud arvutisüsteemi sisendi kujust. Tuleb märkida, et loomuliku keele kasutamine praktilistes süsteemides on seni väga täpselt reglementeeritud, osalt just viimati nimetatud põhjusel. Süsteem saab kasutajalt infot peamiselt formaalsetes konstruktsioonides, etteantud vastuseformaadis või graafiliselt. Loomulikuma väljundi võib luua sellest hoolimata.

2.0. Arvutisüsteemide kõneväljundi väljatöötamisel vajalike lingvistilist laadi teadmiste hankimise töömahukus ja nende väärtus sõltuvad suuresti sellest, kui võrd otseselt konkreetse keele professionaalne ekspert saab osaleda tööprotsessis. Et vähendada inseneride ja programmeerijate vahendajarolli keelespetsiifilise teadmise saamisel ja arvutitöötuseks ettevalmistamisel, läheb tarvis kõnetehnoloogilist arendussüsteemi, mis annaks lingvistile võimaluse lihtsal, hästi jälgitaval kujul oma teadmisi esitada, neid kohe tehiskõne genereerimiseks rakendada ning kui vaja, kogutud teadmistesse korrektiive viia. Paindlike arendussüsteemide puudumises nähakse hea tehiskõne loomise peamist takistust (3). Märkimisväärne ühtlasi, et niisugused süsteemid tooksid lingvistidele kasu ka kui uurimisvahendid (allpool kirjeldatav süsteem on näiteks kasulik vahend foneetiliste või fonoloogiliste kirjelduste adekvaatsuse selgitamiseks).

Järgnevalt käsitleme ENSV TA Küberneetika Instituudis personaalarvutile loodud arendussüsteemi kõne reegelsünteesil kasutatavate teadmiste väljatöötamiseks ja esitamiseks. Saadud teadmiste alusel valmistab süsteem ette parameetrilise sisendspetsifikatsiooni kõnetrakti formantmudelile.

2.1. Kõnetrakti formantmudelil FS-05 (vt. joon.1), millel on mitmeid ühisjooni hästi tuntud mudeliga OVE-III (9)), koosneb 3 reguleeritavast (F1, F2, F3) ja 2 fikseeritud formantfiltrist (F4, F5), reguleeritavast frikatiivi- ja fikseeritud nasaaliresonaatorist (vastavalt FF ja FN), 5 regulaatorist vokaalide (AV), nasaalide (AN), frikatiivse (AF) ja aspiratoorse müra (AH) ning heliliste klusiilide madalsagedusliku komponendi (MO) amplituudi muutmiseks. Muuta saab ka põhitooni (PT) ning põhitooni ja amplituudide siirdeid (TP, TF, TA). Üksiku vokaali a saab genereerida näiteks järgmise parameetrikogumi alusel (näited on eestikeelse kõne sünteesilt):

F1=660 Hz, F2=950 Hz, F3=3000 Hz, FF=3500 Hz,

AN=0, AH=0, AF=0, AV=100%, MO=0

Juurde võib lisada hääliku kui kõnesegmendi kestuse D millisekundites. Häälikuühendi esitamisel tuleb kindlaks teha häälikute kontekstuaalseid ja positsioonilisi variatsioone kujundavad algparameetrite modifikatsioonid ning ühtlasi kirjeldada sagedus- ja amplituudikarakteristikute siirded.

2.2. Praktilise süsteemi loomiseks, mis sünteesib kõne parameetrilise esituse traditsioonilisel viisil kirja pandud teksti alusel, peab lingvistil olema võimalus töötada nii konkreetse foneetilise kirjelduse kui ka fonoloogiliste abstraktsioonide tasemel. Foneetiliste kirjeldustega töötamiseks on sobiv kasutada kõnetrakti mudeli juhtparameetreid eksplitseerivat aktiivset arvutigraafikat (vt. näit. joonis 2). Käsitletavas arendussüsteemis saab graafiliselt esitatud parameetreid tajueksperimentide käigus manuaalselt muuta.

2.3. Põhiliselt fonoloogilise iseloomuga teadmised, sealhulgas samuti üldistatud teadmised kõne prosodia kujundamise kohta võib ekspert koondada arendussüsteemi teadmisbaasi. Teadmisbaas on teadmiste - faktide, nendevaheliste suhete ja reeglite spetsialiseeritud kogum arvuti mälus. Teadmisbaasi kui häälikusüsteemi abstrahdeeritud mudelit kasutades sünteesivad süsteemi täitevmodulid kõnetrakti juhtimiseks vajalikke parameetreid.

Kõnetehnoloogiliste arendussüsteemide väljatöötamise kogemused kinnitavad, et eksperdile mugava teadmiste esitamise viisi väljatöötamine teadmisbaasi komplekteerimiseks ei ole kerge ülesanne. Tuleb ju ühtaegu pakkuda võimsaid ja sealjuures ka programmeerimisoskusteta lingvistil vaatepunktist küllalt arusaadavaid vahendeid teadmisbaasi viimiseks sobivasse seisundisse. Teiselt poolt on vaja arvestada realiseeritava sünteesisüsteemi tehniliste piirangutega.

2.3.1. Vaadeldav süsteem töötab teadmistega, mille staatiline osa esitatakse lihtsate deklaratsioonidena ja protseduuriline osa produktsioonireeglites, s.o. konditsionaalsetes struktuurides, milles kirjeldatud operatsioonid kuuluvad täitmisele vaid juhul, kui antud eeltingimused osutuvad tõesteks (10). Produktsioonireegli üldkuju on järgmine:

KUI tingimus SIIS tegevus

Niisugustes reeglites teadmiste väljendamine peaks olema üsna loomulik. Reeglisüsteemil on veel mitmeid häid jooni, näiteks: iga reegel defineerib väikese ja suhteliselt sõltumatu teadmiste osa; reegleid võib teadmisbaasis muuta ja uusi reegleid lisada sõltumatult teistest reeglitest. Reegleid töötlevas produktsioonisüsteemis on pealegi hõlpus genereerida selgitusi reeglite kasutamise kohta.

2.3.2. Teadmisbaas on struktureeritud selliselt, et ühtaegu

oleks arvestatud fonoloogilise kirjelduse loogika ja tagatud teadmisaasi töötlemine reaalses. Lingvistil on vaja tutvuda teadmisaasi põhistruktuuriga, teadmiste organiseerimise ja produktsioonisüsteemi juhtstrateegiaga, kuid ainult kontseptuaalsel tasemel.

2.3.2.1. Teadmisaas koosneb kolmest põhilisest osast. Esimeses osas säilitatakse fonemaatiliste üksuste kirjeldused (loobume siin terminist foneem, sest selle lingvistilises mõttes ranged tõlgendamiskatsed võivad tekitada asjatut segadust) ja reeglid, mis määravad ühikute positsioonilised variatsioonid ning koartikulatsiooni.

Fonemaatiliste üksuste kõik kirjeldused peavad mahtuma süsteemi püsimalu puhvrise, mille maht on praegu piiratud (256 baiti). Seega, mida komplitseeritumad need kirjeldused on, seda vähem mahub neid mälu puhvrise.

Kui fonemaatiline üksus a kirjeldatakse nii, nagu ülal esitatud (vt. 2.1.), moodustub kirjelduse kodeerimise ja kokkupakkimise järel 4-baidine juhtvektor. Nimetame seda baasvektoriks. Allotandasil on iga ühiku kohta vaja juba arvukalt erinevaid juhtvektoreid, mida aga võib käsitleda kui baasvektori modifikatsioone. Nende kõigi säilitamine ja kasutamise vajaduse igakordne eksplitsiitne esitamine ei ole ilmselt otstarbekas. Selle asemel võib ära näidata tingimused, mille kehtimise korral juhtvektor tuletatakse automaatselt baasvektorist ja lülitatakse seejärel sünteesiprotsessi. Niisiis peab ekspert lahendama järgmise optimeerimisülesande: leida tehniliste piirangute keskel sellised fonemaatilised üksused, mis koos oma reaalses saadavate teisenditega tagaksid võimalikult loomutruult varieeruva kõne genereerimise.

Juhtvektoriga määratud sageduse, amplituudi ja siirete kontekstuaalsed muutused esitatakse produktsioonireeglitega. Nende kirjapanemiseks loodud lihtsaid keeli oleme tutvustanud varasemates töödes, vt. näit. (7). Esitame siin kolm näidet eestikeelse kõne sünteesil reeglendamist vajavaks osutunud fonemaatiliste üksuste variatsioonide ülesmärgimise kohta.

- Tähistatagu S0, S1, S2 järjestikuseid sisendmärke;
- %1 olgu suvaline fonemaatiline üksus;
- %2 olgu suvaline vokaal;
- %3 olgu fonemaatiline üksus hulgest p, k, t, f, h, s, s', š .

Kiire amplituudsiire klusiilide järel:

(1) KUI (SO=k V SO=p V SO=t V SO=t') & S1=%1 SIIS S1 par TA=Oms

Selgitus: kui esimene vaadeldav fonemaatiline üksus (sisendmärk) on k või p või t', ja teine on ükskõik milline määratud üksustest, siis amplituudsiirde kestus teisele üksusele üleminekul on minimaalne (0 millisekundit).

/m/-i kvaliteedi muutus vokaalide vahel:

(2) KUI SO=%2 & S1=m & S2=%2 SIIS S1 par F3=3300 Hz

Selgitus: kui esimene vaadeldav fonemaatiline üksus kuulub vokaalide hulka, teine on m ja kolmas kuulub vokaalide hulka, siis teise üksuse kolmanda formandi sagedus on 3300 Hz.

/e/ aeglane formantsiire helitute konsonantide järel:

(3) KUI SO=%3 & S1=e & S2=%1 SIIS S1 par TF=100 ms

Selgitus: kui esimene fonemaatiline üksus kuulub hulka %3 (helitud konsonandid) ja teine on e ning kolmas ükskõik milline, siis formantsiire esimeselt üksuselt teisele üleminekul kestab 100 ms.

Esitatud näidete põhjal tuleb ilmsiks vajadus fonemaatiliste üksuste funktsionaalse jaotuse järele - sellel alusel, kuidas nad mitmesugustes kontekstides käituvad, s.t. millist mõju vastastikku avaldavad, milliseid variaablussuheteid tingivad. Käsitletavas süsteemis on kujundatud puukujuline struktuur, mida arvuti saab ökonoomselt töödelda. Puu sõlmi võib käsitleda fonemaatiliste üksuste distinktiivtunnustena ja anda neile olemasoleva terminoloogilise süsteemiga kooskõlalisi nimetusi. Nii saab reegleid küllalt ülevaatlikul viisil kirjeldada. Esitame näiteks uuesti reeglid (1), (2) ja (3):

(1-2) KUI SO=SULG & S1=FON & S2=FON SIIS S1 par TA=Oms

(2-2) KUI SO=VOKAAL & S1=m & S2=VOKAAL SIIS S1 par F3=3300 Hz

(3-2) KUI SO=HELITU & S1=e & S2=FON SIIS S1 par TF=100 ms

Eestikeelse kõne sünteesi jaoks välja töötatud fonemaatiliste üksuste komplekti rühmitamist tutvustab joonis 3. Lisame, et sünteesimisel rakendatakse praegu kontekstiefekte kirjeldavat 45 teisendusreeglit.

2.3.2.2. Teadmisaasi teine osa on prosoodia kujundamise kohta käivate teadmiste jaoks. Esialgu saab formeerida vaid kõne ajalise mudelit (kestusmudelit), edaspidi aga ka põhitooni ja intensiivsust reguleerivaid seaduspärasusi.

Sünteeskõne kestusmudel peab samuti nagu äsjakirjeldatud kontekstimudel arvestama mitte ainult inimese poolt tekitatud akustilise kõnesignaali mõõtmisel saadud tulemusi, vaid alati silmas pidama ka erinevate mõjufaktorite pertseptiivset tähtsust. Uuringutulemuste operatiivse kasutamise kestusmudeli loomisel tagab taas produktsioonireeglite süsteem.

Kestusmudelit kirjeldab valem

$$D = D_1 f_1 + D_2 f_2 + \dots + D_n f_n,$$

kus D - segmendi kestus;

$f_1 \dots f_n$ - muutuja, mis kirjeldab segmendi kestust mõjutavat faktorit produktsioonireegli tingimuses;

$D_1 \dots D_n$ - kestuse muutus millisekundites, mis määratakse produktsioonireegli operatsiooniosas.

Produktsioonireegel kirjeldab valemi aditiivset komponenti $D_i f_i$: kui sisendtekstis esineb tingimuseks seatud situatsioon, siis tuleb segmendi algkestusele lisada kestus $+D_i$. Tingimusi võib kirjeldada analoogiliselt eespool esitatuga (vt. 2.3.2.1.).

2.3.2.3. Teadmisaasi kolmandas osas esitatakse seaduspärasused, mille alusel saab loodud fonoloogilist süsteemi suhtestada tavapärase või kasutaja poolt sobivaks peetud kirjaväisiga. Deklareeritakse vastavused fonemaatiliste elementide ja grafeemide vahel ning grafeemide grupid. Produktsioonireeglites kirjeldatakse vajalikud teisendused, näiteks:

MÜÜA - MÜI'JA

(4) KUI SO=Ü & S1=Ü & S2=A SIIS S1=I'J

Eriotstarbelisi teisendusi saab esitada suhteliselt omaette seisvate reeglikompleksidena. Nii on eesti keele jaoks ette valmistatud sõnarõhu automaatse määramise reeglid.

Teadmisaasi võib paigutada ka juhendeid sisendtekstis ette tulevate lühendite ja märkide lugemiseks.

Teadmisaasi üldine struktuur on esitatud joonisel 4. Ekspordile esitatakse lähtematerjalina fonemaatiliste üksuste ja grafeemide kogumid.

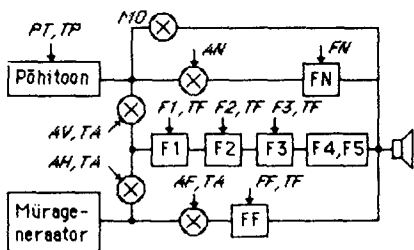
2.4. Arendussüsteemi juurde kuulub moodul selgituste andmiseks teadmisaasi kasutamise kohta sünteesiprotsessis. Õigupoolest loobki lingvist ülalkirjeldatud vahendite abil spetsialiseeritud ekspertsüsteemi - nimetagem seda näiteks häälduseksperdik - , mis lisaks teksti lugemisele on suuteline jagama selgitusi kõnemoodustamise protsessi üksikseikade kohta. Selline ekspertsüsteem vastab küsimustele: miks sõna või häälikukombinatsiooni hääldatakse just niimoodi, või: kuidas, milliste seaduspärasuste alusel selline tulemus on saadud. Süsteem esitab vastavad teadmised nende rakendamise järjekorras ning toob ühtlasi välja kõik vahetulemused. Niisugune tagasiside on vajalik esmajoones reeglitekompleksi töö kontrollimisel, kuid olulisemaid teisendusi teel õigekirjalt hääldusele selgitav partner võib anda abi võõrkeele õppijatelegi.

K i r j a n d u s

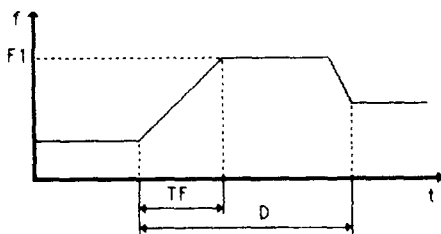
1. Klatt, D. H., The KLATTTalk text-to-speech conversion system. - In: Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing, 1982, pp. 1589-1592.
2. Hertz, S. R., From text to speech with SRS. - In: J. Acoust. Soc. Am., 1982, vol. 72, No.4, pp. 1155-1170.
3. Hertz, S. R., Kadin, J., Karplus, K. J., The DELTA rule development system for speech synthesis from text. - In: Proc. of the IEEE, 1985, vol.73, No.11, pp. 1589-1601.
4. Chapanis, A., Parrish, R. N., Ochsmann, R. B., Weeks, G. D., The effect of four communication modes on the linguistic performance of teams during cooperative problem solving. -In: Human Factors, 1977, vol. 19, No.2, pp. 101-126.
5. Pisoni, D. B., Nusbaum, H. C., Greene, B. G., Perception of Synthetic Speech Generated by Rule. - In: Proc. of the IEEE, 1985, vol. 73, No.11, pp. 1665-1676.
6. Voiers, W. D., Diagnostic evaluation of speech intelligibility. - In: M. Hawley, Ed., Benchmark Papers in Acoustics, vol. II, Stroudsburg, PA: Dowden, Hutchinson and Ross, 1977.
7. Ott, A., Siil, I., Real-time speech synthesis - development and employment. - In: Computers and Artificial Intelligence, Bratislava, 1987, vol. 6, No.2.
8. Kupka, I., Wilsing, N., Dialogsprachen. Stuttgart: Teubner, 1975.

<p>ELEM fonemaatiliste üksuste vestavad juhtparameetrid</p> <hr/> <p>fonemaatiliste üksuste rühmad</p>	<p>MOODUL B</p>	<p>RULE juhtparameetrite teisendusreeglid</p>
<p>ELEM</p>	<p>C</p>	<p>TIME kestusreeglid (põhitoonireeglid) (intensiivsusreeglid)</p>
<p>DEF fonemaatiliste üksuste ja grafeemide vahelised vestavused</p> <hr/> <p>grafeemide rühmad</p>	<p>A</p>	<p>ORFRULE asendusreeglid grafeem → fonemaatiline üksus</p>
<p>CLIT kliitivate sõnastik</p>		<p>rõhureeglid j.m.</p>
<p>ABBREV lühendite sõnastik</p>		

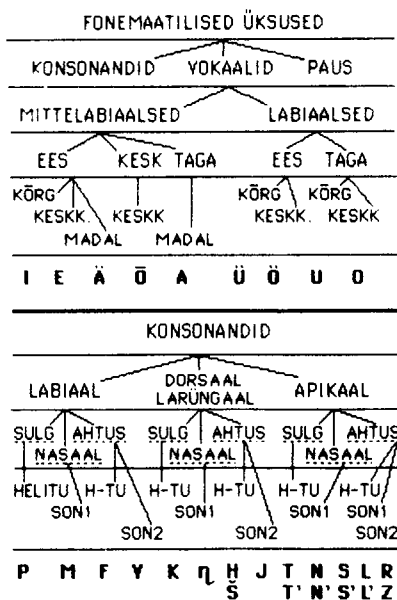
JOOINIS 4



JOOINIS 1



JOOINIS 2



JONIS 3

9. Liljencrants, J., The OVE III speech synthesizer. - In: STLQPSR, 1967, No.2-3, pp. 76-81.
10. Nilsson N. J., Principles of Artificial Intelligence. SRI International. Tioga Publishing Company, Palo Alto, California, 1980.

К КОММУНИКАТИВНОЙ ИСКУССТВЕННОЙ РЕЧИ:
ПРОБЛЕМЫ И СРЕДСТВА РАЗРАБОТКИ

И. Сийл, А. Отт

Р е з ю м е

В настоящее время развитие устройства обработки речи, в частности синтезаторы речи, могут озвучивать любые тексты. Однако, при реализации речевого ответа микропроцессорных систем возникли многие трудности:

- качество синтезируемой речи не удовлетворяет пользователей;
- при разработке речевого ответа мало учтены категории пользователей и в этой связи возникает широкий круг сложных задач: определение ролевого статуса системы речевого ответа как партнера общения; формирование узуса системы речевого общения; выявление оптимальных качественных характеристик синтезируемой речи; установление подходящей корреляции между разными выходными каналами системы в содержательных, объемных и функциональных аспектах; проектирование динамики общения;
- на реализацию коммуникативных систем с речевым ответом оказывает влияние ограничения вычислительных ресурсов.

Дальнейший прогресс в области речевой технологии, очевидно, зависит от гибких и мощных средств проектирования и отладки. Полезно постепенно уменьшить роль инженеров синтеза речи в процессе получения разного рода лингвистических знаний и дать лингвистам возможность подборно моделировать и непосредственно использовать для синтеза речи их профессиональные знания.

Рассматривается система разработки знаний для синтеза речи. На основе полученных знаний система подготавливает параметрическую входную спецификацию для формантной модели речевого тракта (модель ФС-05). Лингвисту предлагается возможность составлять базу знаний согласно своей концепции, используя при этом некоторые простые и прозрачные формализмы. База знаний состоит из трех основных частей (рис. 4). Первая

сохраняет параметрические описания фонематических единиц и продукционные правила, определяющие позиционные варианты и коартикуляцию единиц. Вторая часть – для знаний о формировании просодики речи. В третьей части запоминаются объяснения соответствий между символами (для определения способа правописания вводимого текста). Используя активную компьютерную графику, лингвист может работать с конкретным фонетическим описанием. Допускается изменять параметры управления непосредственно в ходе экспериментов восприятия. Система разработки имеет модули отладки и объяснения для проверки и изменения знаний на разных уровнях синтеза речи.