

**НАУЧНЫЙ СОВЕТ DSc.13/30.12.2019.Т.07.01 ПО ПРИСУЖДЕНИЮ  
УЧЕНЫХ СТЕПЕНЕЙ ПРИ ТАШКЕНТСКОМ УНИВЕРСИТЕТЕ  
ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ**

---

**ТАШКЕНТСКИЙ УНИВЕРСИТЕТ ИНФОРМАЦИОННЫХ  
ТЕХНОЛОГИЙ**

**АБДУЛЛАЕВА МАЛИКА ИЛЬХАМОВНА**

**РАЗРАБОТКА ИНФОРМАЦИОННОЙ СИСТЕМЫ РЕАБИЛИТАЦИИ  
ПАЦИЕНТОВ С НАРУШЕНИЕМ СЛУХА НА БАЗЕ РЕЧЕВЫХ  
ТЕХНОЛОГИЙ**

05.01.10 – Информационные системы и процессы

**АВТОРЕФЕРАТ ДИССЕРТАЦИИ  
ДОКТОРА ФИЛОСОФИИ (PhD) ПО ТЕХНИЧЕСКИМ НАУКАМ**

**Ташкент – 2023**

**Texnika fanlari bo'yicha falsafa doktori (PhD)  
Dissertatsiyasi avtoreferati mundarijasi**

**Оглавление автореферата диссертации доктора философии (PhD)  
по техническим наукам**

**Contents of dissertation abstract of doctor of philosophy (PhD)  
on technical sciences**

**Abdullayeva Malika Ikhamovna**

Nutq texnologiyalar asosida eshitish qobiliyati cheklangan bemorlarni reabilitatsiya qilish axborot tizimini ishlab chiqish 3

**Абдуллаева Малика Ильхамовна**

Разработка информационной системы реабилитации пациентов с нарушением слуха на базе речевых технологий 21

**Abdullaeva Malika Ikhamovna**

Development of information system for rehabilitation of patients with hearing loss based on speech technologies 40

**Список опубликованных работ**

E'lon qilingan ishlar ro'yxati  
List of published works 43

**TOSHKENT AXBOROT TEXNOLOGIYALARI UNIVERSITETI  
HUZURIDAGI ILMIY DARAJA BERUVCHI**

**DSc.13/30.12.2019.T.07.01 RAQAMLI ILMIY KENGASH**

---

**TOSHKENT AXBOROT TEXNOLOGIYALARI UNIVERSITETI**

**ABDULLAYEVA MALIKA ILXAMOVNA**

**NUTQ TEXNOLOGIYALAR ASOSIDA ESHITISH QOBILIYATI  
CHEKLANGAN BEMORLARNI REABILITATSIYA QILISH AXBOROT  
TIZIMINI ISHLAB CHIQISH**

05.01.10 - Axborot olish tizimlari va jarayonlari

**TEXNIKA FANLARI BO'YICHA FALSAFA DOKTORI (PhD)  
DISSERTATSIYASI AVTOREFERATI**

**Toshkent – 2023**

**Texnika fanlari bo'yicha falsafa doktori (PhD) dissertatsiyasi mavzusi O'zbekiston Respublikasi Vazirlar Mahkamasi huzuridagi Oliy attestatsiya komissiyasida B2023.2.PhD/T3546 raqam bilan ro'yxatga olingan.**

Dissertatsiya Toshkent axborot texnologiyalari universitetida bajarilgan.

Dissertatsiya avtoreferati uch tilda (o'zbek, rus, ingliz (rezyume)) Ilmiy kengash veb-sahifasida (www.tuit.uz) va «Ziyonet» axborot-ta'lim portalida (www.ziyonet.uz) joylashtirilgan.

**Ilmiy rahbar:**

**Musayev Muhammadjon Mahmudovich**  
texnika fanlari doktori, professor

**Rasmiy opponentlar:**

**Saidov Abdusobir Abdurahmonovich**  
texnika fanlari doktori, professor

**Elov Jamshid Bekturodovich**  
texnika fanlari bo'yicha falsafa doktori (PhD), dotsent

**Yetakchi tashkilot:**

**Sharof Rashidov nomidagi Samarqand davlat universiteti**

Dissertatsiya himoyasi Toshkent axborot texnologiyalari universiteti huzuridagi DSc.13/30.12.2019.T.07.01 raqamli Ilmiy kengashning 2023-yil «23» iyun da soat 14<sup>00</sup> dagi majlisida bo'lib o'tadi. (Manzil: 100084, Toshkent shahri, Amir Temur ko'chasi, 108-uy. Tel.: (99871) 238-64-43; e-mail: iktuit@tuit.uz).

Dissertatsiya bilan Toshkent axborot texnologiyalari universitetining Axborot-resurs markazida tanishish mumkin ( 278 raqam bilan ro'yxatga olingan). (Manzil: 100084, Toshkent shahri, Amir Temur ko'chasi, 108-uy. Tel.: (99871) 238-64-70).

Dissertatsiya avtoreferati 2023-yil « 10 » iyun kuni tarqatildi.  
(2023-yil « 06 » iyun dagi 17 raqamli reyestr bayonnomasi).



**M.A. Raxmatullayev**  
Ilmiy darajalar beruvchi ilmiy  
kengash raisi o'rinbosari, texnika fanlari  
doktori, professor

**N.O. Raximov**  
Ilmiy darajalar beruvchi ilmiy  
kengash ilmiy kotibi, texnika fanlari  
doktori, dotsent

**U.R. Xamdamiyov**  
Ilmiy darajalar beruvchi  
ilmiy kengash huzuridagi ilmiy  
seminar raisi, texnika fanlari  
doktori, dotsent

## KIRISH (falsafa doktori (PhD) dissertatsiyasi annotatsiyasi)

**Dissertatsiya mavzusining dolzarbligi va zaruriyati.** Jahonda nutq texnologiyalari asosidagi axborot tizimlarini yaratishga alohida e'tibor qaratilmoqda, sababi ular qo'llanish sohasi jihatdan keng qamroqlidir. Bugungi kunning asosiy tadqiqot sohasiga aylanib borayotgan nutq interfeys tizimlarini yaratish bevosita nutqni tahlil qilish va sintez qilish bilan bog'liqdir. Ma'lumki, nutq dialogi inson muloqotining asosiy usuli bo'lganligi sababli, nutq komponentlari bazasida intellektual tizimni ishlab chiqish sun'iy intellektning asosiy yo'nalishlaridandir. Xorijiy mamlakatlarda, jumladan, AQSH, Rossiya Federatsiyasi, Xitoy, Yaponiya va boshqa mamlakatlarda nutqni tahlil qilish va sintez qilishning mavjud usullari va algoritmlarini takomillashtirishga qaratilgan nazariy va amaliy ishlarni, hamda nutq signallarini intellektual qayta ishlash usullarini ishlab chiqish sohasida ilmiy izlanishlar olib borilmoqda. Bugungi kunda nutqdan foydalangan holda odam va mashina o'rtasida teng huquqli dialogni ta'minlash qobiliyatiga ega tizimlarni ishlab chiqish, raqamli signallarni qayta ishlash, sun'iy intellekt sohasidagi asosiy vazifalardan biridir.

Jahonda eshitish va nutq nuqsonlari mavjud bo'lgan bolalar tug'ilishi oshib bormoqda. Shu munosabat bilan, koxlear implant bozori 2027 yilga kelib o'rtacha 8,57 foizga o'sishi kutilmoqda. Shuning uchun koxlear implantatsiyadan so'ng bemorlarni reabilitatsiya qilish uchun axborot tizimini ishlab chiqish dolzarb masala hisoblanadi. Bemorlarning eshitish va nutq imkoniyatlarini reabilitatsiya qilish zamonaviy axborot tizimlari nutqni tahlil va sintez qilish vositalari asosida qurilgan nutq interfeysi yordamida amalga oshiriladi.

Respublikamizda Prezidentning farmoniga<sup>1</sup> muvofiq «...aholiga ko'rsatilayotgan xizmatlar sifatini oshirish maqsadida sun'iy intellekt texnologiyalaridan keng foydalanish» vazifalari belgilab berilgan. Axborot texnologiyalari va sun'iy intellektni rivojlantirishni qo'llab-quvvatlash tufayli respublikamizda nutq signallarini intellektual qayta ishlash, talaffuz etilgan tovushlar, bo'g'inlar va so'zlarni avtomatik aniqlash, diktorni ovozi orqali autentifikatsiya qilish, ovozli boshqarish tizimlarini yaratish, o'zbek nutqining tahlili va sintezi bo'yicha ilmiy tadqiqotlar olib borilmoqda. «Raqamli O'zbekiston – 2030» strategiyasi doirasida belgilangan ustuvor yo'nalishlarga ko'ra, respublikada zamonaviy axborot-kommunikatsiya texnologiyalarini faol rivojlantirish va keng joriy etish, xususan, axborot-kommunikatsiya texnologiyalarini o'rganish va amaliyotga tatbiq etish bo'yicha kompleks chora-tadbirlar amalga oshirilmoqda, shu jumladan, sun'iy intellekt texnologiyalaridan, mashinali o'qitish algoritmlaridan foydalanish, barcha sohalarda, birinchi navbatda, sog'liqni saqlashda, shuningdek, davlat boshqaruvi, ta'lim va qishloq xo'jaligida katta ma'lumotlarni tahlil qilish imkoniyatlarini o'rganishga katta e'tibor qaratilmoqda<sup>2</sup>.

O'zbekiston Respublikasi Prezidentining 2021 yil 17 fevraldagi PQ-4996-son «Sun'iy intellekt texnologiyalarini jadal joriy etish uchun shart-sharoitlar yaratish chora-tadbirlari to'g'risida»gi Farmoni, 2020 yil 5 oktyabrdagi PF-6079-son «Raqamli O'zbekiston – 2030» strategiyasi va uni samarali amalga oshirish chora-tadbirlarini tasdiqlash to'g'risida»gi Qarori, hamda mazkur faoliyatga tegishli me'yoriy-huquqiy

---

<sup>1</sup> O'zbekiston Respublikasi Prezidentining 2021 yil 17 fevraldagi PQ-4996-son «Sun'iy intellekt texnologiyalarini jadal joriy etish uchun shart-sharoitlar yaratish chora-tadbirlari to'g'risida»gi Farmoni

<sup>2</sup> O'zbekiston Respublikasi Prezidentining 2020 yil 5 oktyabrdagi PF-6079-son «Raqamli O'zbekiston – 2030» strategiyasi va uni samarali amalga oshirish chora-tadbirlarini tasdiqlash to'g'risida»gi qarori

hujjatlarda belgilangan vazifalarni amalga oshirishda ushbu dissertatsiya tadqiqoti muayyan darajada xizmat qiladi.

**Tadqiqotning respublika fan va texnologiyalari rivojlanishining ustuvor yoʻnalishlariga mosligi.** Mazkur tadqiqot respublika fan va texnologiyalar rivojlanishining IV. «Axborotlashtirish va axborot-kommunikatsiya texnologiyalarini rivojlantirish» ustuvor yoʻnalishi doirasida bajarilgan.

**Muammoning oʻrganilganlik darajasi.** Bugungi kunga qadar rivojlangan mamlakatlarning ilmiy markazlari nutq signallarini qayta ishlash usullari va modellarini ishlab chiqish, matnli axborotni samarali konvertatsiya qilish uchun neyron tarmoqlari arxitekturasi qurish, shuningdek, nutqni generatsiyalash tizimlari uchun koʻp soatlik nutq korpuslarini yaratish boʻyicha qator tadqiqotlar olib bormoqda. Nutqni tahlil qilish va sintez qilish usullarini ishlab chiqish va takomillashtirish, shuningdek ularni amaliy qoʻllashda L.Rabiner, J.Shen, R.Yamomoto, J.Miller, D.Klatt, O.Arik, V.Ping, P.Dyuhamel, I.Kipyatkova, L.Ronjin va boshqalar kabi olimlarning hissasi katta.

Nutqni tahlil qilish va sintez qilish asosida algoritm va tizimlar yaratish boʻyicha oʻzbek olimlaridan M.Aripov, M.Musayev, Oʻ.Xamdamov, N.Mamatov va boshqa olimlarning ishlarini qayd etish mumkin.

Hozirgi kunda nutqni tanib olish (tahlil qilish) va nutq sintezatori yordamida ovozli javoblar generatsiya qilish imkonini aks ettiruvchi axborot tizimlarini yaratish boʻyicha tadqiqotlar olib borilmoqda. Oʻtkazilgan tadqiqotlar tahlili shuni koʻrsatadiki, sunʼiy intellekt algoritmlari asosida oʻzbek nutqini tahlil va sintez qilish muammolari toʻliq hal etilmagan.

**Dissertatsiya tadqiqotining dissertatsiya bajarilgan oliy taʼlim muassasasining ilmiy-tadqiqot ishlari rejalari bilan bogʻliqligi.** Dissertatsiya tadqiqoti Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universiteti ilmiy – tadqiqot ishlar rejasining №-ILM-202101222 «Sud tizimi uchun sunʼiy intellekt algoritmlari asosida oʻzbek tilidagi nutqni qayta ishlash dasturlarini ishlab chiqish» (2021-2023) va №-Φ3-2020103164 «Zamonaviy kompyuter spektral analitik texnologiyalari asosida eshitish qobiliyati zaif boʻlgan bolalarning eshitish va soʻzlashish-nutq qobiliyatini (koxlear implantatsiya tizimida) reabilitatsiya qilish-tiklash va monitoring qilish texnologiyasini ishlab chiqish» (2021-2024) mavzusidagi loyiha ishlari doirasida bajarilgan.

**Tadqiqotning maqsadi** chuqur neyron tarmoqlardan foydalangan holda oʻzbek tili nutqini tahlil qilish va sintez qilish algoritmlari asosida eshitish nuqsoni mavjud bemorlarni reabilitatsiya qilish uchun axborot tizimini ishlab chiqishdan iborat.

**Tadqiqotning vazifalari:**

koxlear implantatsiyadan soʻng bemorlarni reabilitatsiya qilish zamonaviy texnologiyalarini tahlil qilish, eshitish va nutq qobiliyatlarini tiklash uchun arxitektura, texnik va dasturiy taʼminotni tanlash;

koxlear implantatsiyadan soʻng bemorlarning nutqini tiklash jarayoni uchun axborot modellari va multimedia maʼlumotlar bazasini yaratish;

oʻzbek tilining koʻp soatlik nutq bazasini yaratish uchun maʼlumotlarni normallashtirish, sinxronlashtirish va sinovdan oʻtkazish algoritmlarini ishlab chiqish;

neyron tarmoqlar asosida oʻzbek tili nutqini sintez qilish modellari va algoritmlarini ishlab chiqish;

dialog interfeysi uchun oʻzbek tilining nutq birliklarini tahlil qilish algoritmlarini ishlab chiqish;

bemorlarni reabilitatsiya qilish uchun qulay foydalanuvchi interfeysli axborot tizimini yaratish.

**Tadqiqotning obyekti** sifatida koxlear implantatsiyadan soʻng bolalarning eshitish va nutq qobiliyatlarini tiklash jarayonining axborot tizimi qaralgan.

**Tadqiqotning predmeti** koxlear implantatsiyadan soʻng bemorlarni reabilitatsiya qilish uchun axborot tizimi, diktor nutqini tahlil qilish va autentifikatsiya qilish maqsadida oʻzbek nutqini tahlil qilish algoritmlari, oʻzbek nutqini sintez qilish uchun neyron tarmoq arxitekturasi modellari, oʻzbek tili matnini normallashtirish algoritmi va oʻzbek tilining koʻp soatlik nutq bazasini yaratish uchun dasturiy vositalar toʻplami tashkil etadi.

**Tadqiqotning usullari.** Dissertasiya ishi tadqiqotlarni olib borishda quyidagi usullardan foydalanildi: ehtimollar nazariyasi; maʼlumotlarga intellektual ishlov berish; matematik statistika; chiziqli algebra; signallarga raqamli ishlov berish; relyatsion maʼlumotlar bazasini boshqarish tizimlari; obyektga yoʻnaltirilgan dasturlash; qaror qaʼbul qilish algoritmlari.

**Tadqiqotning ilmiy yangiligi** quyidagilardan iborat:

eshitish nuqsoni mavjud bemorlarni reabilitatsiya qilish dasturiy komponentlari va MVT modeli asosida axborot tizimining arxitekturasi ishlab chiqildi;

nutq tahlili va sintez modellari asosida nutq interfeysi ishlab chiqildi, chuqur neyron tarmoqlar asosida yaratilgan oʻzbek tili TTS tizimining IDEF0 diagrammasi qurildi;

oʻzbek tili nutq sintezi jarayonining nutq bazasi va bemorni reabilitatsiya maxsus lugʻatning nutq birliklari etalonlari maʼlumotlar bazasi shakllantirildi;

bemorning nutq qobiliyatlarini tiklash uchun oʻzbek tili fonemalari, boʻgʻinlari va soʻzlarini tahlil qilish jarayoni uchun algoritm ishlab chiqildi.

**Tadqiqotning amaliy natijalari** quyidagilardan iborat:

maktab yoshigacha boʻlgan bemorlarning nutqini tiklash jarayoni uchun foydalanuvchi nutq interfeysi ishlab chiqildi;

nutqida nuqsoni boʻlmagan 3 yoshdan 7 yoshgacha boʻlgan bolalar tomonidan talaffuz qilingan oʻzbek tilining tovushlari, boʻgʻinlari va soʻzlaridan iborat yagona nutqiy etalonlar bazasi shakllantirildi;

reabilitatsiya texnologiyasida qoʻllaniladigan statik va dinamik multimediya maʼlumotlar bazasi yaratildi;

SpeechKIDS axborot tizimi asosida koxlear implantatsiyadan keyin bolalarni nutqini reabilitatsiya qilish uchun axborot tizimi (trenajori) ishlab chiqilgan;

koxlear implantatsiyadan keyin bemorlarda nutqni tiklash jarayonida axborot tizimidan foydalanish samaradorligi koʻrsatkichlari aniqlandi.

**Tadqiqot natijalarining ishonchliligi.** Tadqiqot natijalarining ishonchliligi koxlear Implantatsiyadan keyin bolalarni reabilitatsiya qilish jarayonida ishlab chiqilgan axborot tizimidan foydalanish natijasida olingan test xulosalari, shuningdek, xalqaro ilmiy konferentsiyalarda muhokama qilingan ilmiy maqolalarda dissertatsiyadan taklif qilingan algoritmlar va arxitekturalarni qoʻllash asosida olingan tajribalar natijalari bilan tasdiqlanadi.

**Tadqiqot natijalarning ilmiy va amaliy ahamiyati.** Tadqiqot natijalarining ilmiy ahamiyati ishlab chiqilgan quyi tizimlar sunʼiy intellekt algoritmlari asosida oʻzbek tili nutqini tahlil qilish va sintezining nazariy asoslarini yanada rivojlantirishga qoʻshgan hissasi bilan belgilanadi.

Tadqiqot natijalarining amaliy ahamiyati ishlab chiqilgan axborot tizimi eshitish qobiliyati cheklangan maktab yoshigacha boʻlgan bemorlarni reabilitatsiya qilish jarayonini 1,5-2 martaga tezlashtirishi bilan belgilanadi.

**Tadqiqot natijalarini joriy etish.** Sun'iy intellekt algoritmlari yordamida nutqni tahlil qilish va sintez qilish usullari, algoritmlari va dasturiy vositalariga asoslangan axborot tizimi:

Respublika ixtisoslashtirilgan pediatriya ilmiy-amaliy tibbiyot markazida (RIPITM), LOR a'zolarining tug'ma va orttirilgan kasalliklari bo'limida bemorlarning koxlear implantatsiyadan so'ng eshitish va nutqini reabilitatsiya qilish uchun joriy etilgan (O'zbekiston Respublikasi raqamli texnologiyalar vazirligining 2023 yil 11 maydagi 33-8/3119-son ma'lumotnomasi). Axborot tizimidan foydalanish bolalarni reabilitatsiya qilish jarayonini surdopedagog rahbarligida an'anaviy reabilitatsiya usuliga nisbatan 1,5-2 baravar tezlashtirish imkonini berdi. Ilmiy ishning samarasi, reabilitatsiya jarayonida ishlab chiqilgan nutq texnologiyalaridan foydalanish orqali erishiladi; Shu o'rinda men sizning fikringizni ushbu dissertatsiya ishining ijtimoiyligiga qaratmoqchiman.

“Nova research and resolution” MChJda nutq buzilishi bo'lgan bolalarda nutq nuqsonlarini bartaraf etish uchun joriy etilgan (O'zbekiston Respublikasi raqamli texnologiyalar vazirligining 2023 yil 11 maydagi 33-8/3119-son ma'lumotnomasi ). Natijada, taklif etilgan quyi tizimlar asosida ishlab chiqilgan axborot tizimi logoped bilan bolalarning mashg'ulotlar sonini qisqartirish imkonini berdi, shu bilan bir qatorda axborot tizimi onlayn ko'rinishga ega bo'lganligi bolalarning istalgan vaqtda shug'ullanishiga imkon beradi. Bu esa o'z navbatida nutq nuqsonlari mavjud bo'lgan bolalarni reabilitatsiya qilish jarayonini tezlashtirishdi.

**Tadqiqot natijalarining aprobatsiyasi.** Mazkur tadqiqot natijalari 8 ta xalqaro va 2 ta Respublika ilmiy-amaliy anjumanlarida ma'ruza qilingan va muhokamadan o'tkazilgan.

**Tadqiqot natijalarining e'lon qilinganligi.** Tadqiqot mavzusi bo'yicha jami 20 ta ilmiy ish chop etilgan, shulardan, O'zbekiston Respublikasi Oliy attestatsiya komissiyasining dissertatsiyalar asosiy ilmiy natijalarini chop etish tavsiya etilgan ilmiy nashrlarda 6 ta maqola, shulardan 3 ta xorijiy jurnallarda va 3 ta respublika jurnallarda chop etilgan hamda 4 ta EHM uchun ishlab chiqilgan dasturiy vositalarni qayd qilish guvohnomalari olingan.

**Dissertatsiyasining tuzilishi va hajmi.** Dissertatsiya ishi kirish, to'rtta bob, xulosa, foydalanilgan adabiyotlar ro'yxati va ilovalardan iborat. Dissertatsiyaning umumiy hajmi 116 betni tashkil etadi.

## **DISSERTATSIYANING ASOSIY MAZMUNI**

**Kirish** qismida dissertatsiya mavzusining dolzarbligi va zaruriyati asoslangan, tadqiqotning O'zbekiston Respublikasi fan va texnologiyalar taraqqiyotining ustuvor yo'nalishlariga mosligi ko'rsatilgan, tadqiqotning maqsad va vazifalari belgilab olingan hamda tadqiqot obyekti va predmeti aniqlangan, olingan natijalarning ishonchliligi asoslab berilgan, ularning nazariy va amaliy ahamiyati, tadqiqot natijalarini amalda joriy qilish holati, nashr etilgan ishlar va dissertatsiya tuzilishi bo'yicha ma'lumotlar keltirilgan.

Dissertatsiyaning «**Intellektual axborot tizimlarida nutq texnologiyalarining modellari, usullari va algoritmlari**» deb nomlangan birinchi bobida intellektual axborot tizimlarining asosiy tarkibiy qismlari hisoblanmish nutq texnologiyalari, nutq signallarini tahlil qilish jarayonining algoritmlari o'rganilgan va matnni nutqqa o'girish usullari keltirilgan.

Koxlear implantatsiyadan so'ng bemorlarning nutqini reabilitatsiya qilish axborot tizimlari arxitekturasiga qo'yiladigan talablarni tahlil qilish asosida foydalanuvchi va mashina o'rtasidagi nutq dialogini ta'minlash uchun nutqni tahlil qilish va sintez qilish



quyi tizimlarini ishlab chiqish muhimligi ko'rsatilgan. Nutq signallarini tahlil qilish algoritmlari, shuningdek yuqori darajadagi TTS (eng. Text-to-Speech) o'zbek nutqini yaratish tizimlari o'rganilgan.

Dissertatsiyaning «O'zbek tili nutqini sintez qilish tizimini ishlab chiqish» deb nomlangan ikkinchi bobida TTS tizimlarining neyron tarmoq arxitekturalariga, ya'ni zamonaviy akustik modellar va neyron vokoderlarning arxitekturalariga, nutq bazasi asosida o'zbek nutqini sintez qilish algoritmini ishlab chiqishga, o'zbek tilining nutq bazasini shakllantirish algoritmiga, o'zbek tilining shakllangan nutq bazasining statistik xususiyatlariga, shuningdek ishlab chiqilgan o'zbek tili TTS tizimini baholashga bag'ishlangan.

Umuman olganda, yuqori darajadali matnni nutqqa o'girish texnologiyasi quyidagi bosqichlarni o'z ichiga oladi:

- tilning nutq bazasini shakllantirish;
- akustik modelni o'qitish;
- akustik parametrlarga asoslangan neyron tarmoq vokoderini o'qitish.

Shuningdek, ushbu bobda yuqori darajali TTS tizimi uchun nutq bazasini yaratish usullari, o'zbek tilining nutq bazasini shakllantirish ketma-ketligi, nutq bazasining asosiy xususiyatlari va shakllangan ko'p soatlik nutq bazasining statistik parametrlari keltirilgan.

Ma'lumki, zamonaviy yuqori darajali TTS tizimlarining asosiy tarkibiy qismi katta hajmli nutq bazasi hisoblanadi. TTS tizimlarining nutq bazasi - bu audio ma'lumotlar to'plami va tegishli matnli fayllardan tashkil topgan ma'lumotlar bazasi hisoblanadi. Audio fayllar nutq birliklarining (tovushlar, bo'g'inlar, so'zlar, jumlar) namunalaridan iborat bo'lib, matnli fayllar ushbu nutq birliklariga mos keladigan transkripsiyalarni o'z ichiga oladilar. Shakllantirilgan TTS tizimi nutq bazasi o'zbek tilidagi har hil uzunlikdagi jumalarning audio va matn fayllaridan tashkil topgan

Bugungi kunda ochiq manbaalarda dunyo tillari uchun qator nutq bazalari mavjud, shu bilan birga, nutq sintezatorni o'qitish uchun o'zbek tilining nutq bazasi mavjud emas.

Amaliyotda, sintetik nutqning sifati nutq bazasining sifatiga bog'liq. Bu, ayniqsa, neyron tarmoq arxitekturalariga asoslangan nutq sintezi tizimlari uchun tasdiqlangan. TTS tizimlari uchun nutq bazasini shakllantirishning keng tarqalgan quyidagi usullari tasdiqlaydi:

- oldindan tayyorlangan matn, materialni o'qiyotgan diktorni ovozini yozib olish;
- ixtiyoriy nutq, rivoyatlar va boshqalarni aytadigan diktorning ovozini yozib olish.

Nutq bazasini yaratishning eng oddiy shakli to'rtta asosiy bosqich asosida amalga oshiriladi: matnli ma'lumotlarni tayyorlash, matnni ovozlashtirish, audio signalni uning matni bilan sinxronlash, audio fayl va uning transkripsiyasini nutq bazasiga qo'shish.

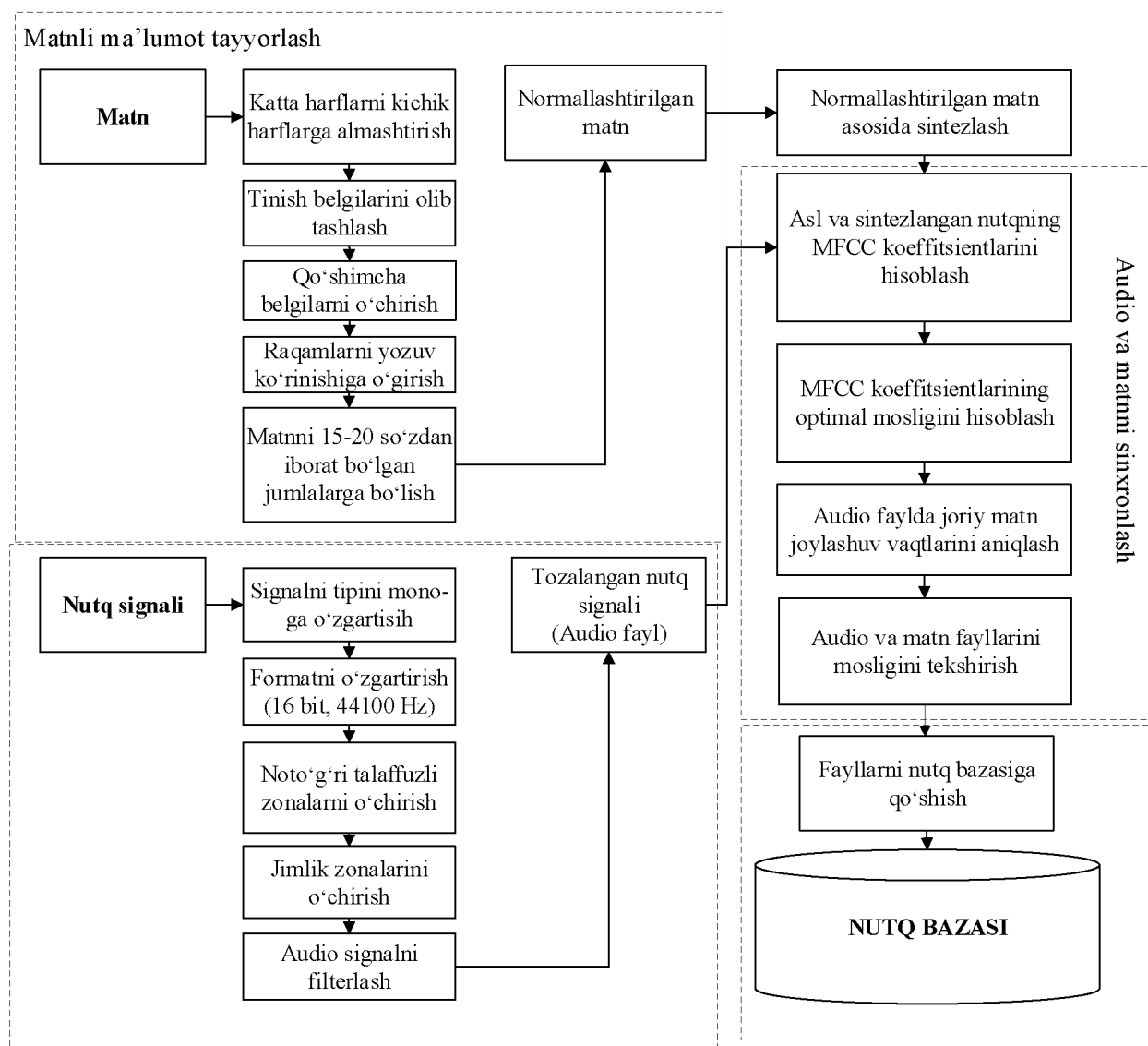
**1- jadval.**

**Nutq bazasining xususiyatlari**

<b>Kategoriya</b>	<b>O'qitish (Training)</b>	<b>Parametrlarni sozlash (Validation)</b>	<b>Umumiy</b>
Nutq bazasining hajmi (soatda)	26	3	29
Jumlalar	11107	584	11691
So'zlar	143902	7574	151476
Takrorlanmaydigan so'zlar	21585	1136	22721

Taklif etilayotgan usul asosida o‘zbek tilining nutq bazasi shakllantirildi. Nutqni sintez qilish tizimining nutq bazasi hajmi ~ 29 soatni tashkil etdi. Nutq bazasi ayol-diktor ovozidan tashkil topgan. O‘qish uchun taqdim etilgan matnlarda jami 11 mingdan ortiq jumlar ishlatilgan. Jumalarda ~ 151,500 o‘zbek tilidagi so‘zlar ishlatilgan, ulardan 22721 tasi takrorlanmaydigan so‘zlardir.

O‘zbek tili uchun nutq bazasini shakllantirishning kengaytirilgan BPMN (eng. Business Process Model and Notation) modeli asosidagi blok diagrammasi 1-rasmda keltirilgan.



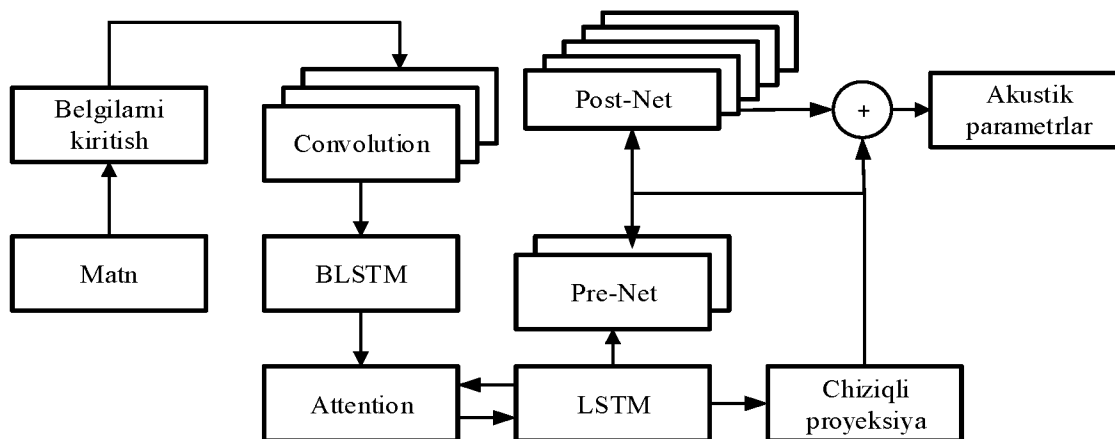
**1-rasm. O‘zbek tili uchun nutq bazasini shakllantirish jarayonining BPMN modeli**

Quyida yuqori darajadali nutq sintezatorlarining arxitekturasini ko‘rib chiqilgan. Yuqori darajadali TTS tizimlar ikkita asosiy neyron tarmoq arxitekturasidan iborat – akustik model va vokoder.

Akustik model formantlar, ohang chastotasi, energiya, chastota-kepstral koeffitsientlarini mel-shkalasi (MFCC, eng Mel-frequency cepstral coefficients), chiziqli bashorat qilish koeffitsientlari (LPC, eng. Linear predictive coefficients), chiziqli spektrogramma, mel-spektrogrammasi va boshqalar kabi akustik parametrlarni hisoblash

uchun mo'ljallangan. Mashhur zamonaviy akustik modellar bo'lib DeepVoice, Clarinet, Tacotron va Fast Speech hisoblanadi.

Tacotron2 15 soatdan ortiq nutq bazalar uchun mo'ljallangan akustik model bo'lib, u ikkita tarmoqdan tashkil topgan: xususiyatlarni bashorat qilish tarmog'i va vokoderdan. Xususiyatlarni bashorat qilish tarmog'i asosiy tarmoq sifatida qaraladi va WaveNet vokoderi akustik parametrlarga asoslangan nutqni yaratish tizimi vazifasini bajaradi. Tacotron2 asosida taklif qilingan akustik modelning tuzilishi 2-rasmda va uning parametrlari 2-jadvalda keltirilgan.



2-rasm. Tacotron2 arxitekturasi asosida ishlab chiqilgan akustik modelning tuzilishi

2-jadval.

Akustik model neyron tarmog'ining qatlamlari va parametrlari

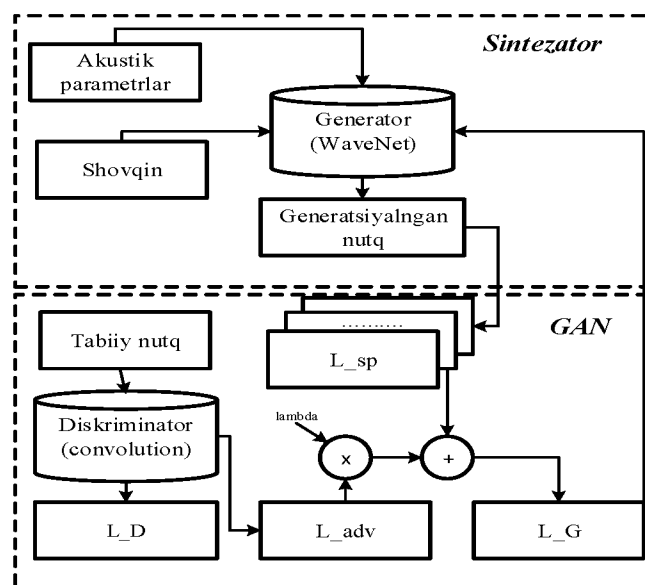
№	Qatlam nomi	Qatlam xususiyatlari	Qatlamlar soni
1	Kirish qatlami	Transkripsiya shaklidagi normallashtirilgan matn (26x1)	1
2	Belgilarni joylashtirish	Qatlam o'lchami = 512 (x5) Faollashtirish funksiyasi = ReLU Dropout=0,5	1
3	Svyortka qatlami	Qatlam o'lchami = 512 Faollashtirish funksiyasi = ReLU Dropout=0,5	3
4	BLSTM	Qatlam o'lchami = 512 Faollashtirish funksiyasi = ReLU Dropout=0,5	1
5	Attention	Xotira yacheykalari soni = 512 Batch Size=32	1
6	LSTM	Xotira yacheykalari soni = 1024 Zoneout=0,1	2
7	PreNet	Xotira yacheykalari soni = 256	2
8	PostNet	Xotira yacheykalari soni = 512 Filtr o'lchami =5 Faollashtirish funktsiyasi = tanh	5
9	O'qitish parametrlari	Optimallashtirish algoritmi: Adam o'qitish qadami uzunligi =0.001, Epoch=160	

Yaratilgan akustik model parametrlari keyingi komponentga - vokoderga uzatiladi. Vokoder - TTS tizimning akustik modelidan olingan akustik parametrlar asosida to'liq shaklini yaratish uchun mo'ljallangan tizim. Amalda qo'llanilgan vokoderlar chuqur neyron tarmoq hisoblanadi. Parallel to'liq hosil qiluvchi neyron vokoderlar orasida arxitektura o'z-o'zini o'qitish qobiliyati mavjudligi, shuningdek, qolganlar vokoderlarga nisbatan quyidagi afzalliklar mavjudligi sababli GAN (eng. Generative Adversarial Network) asosidagi modellar alohida o'rin tutadi:

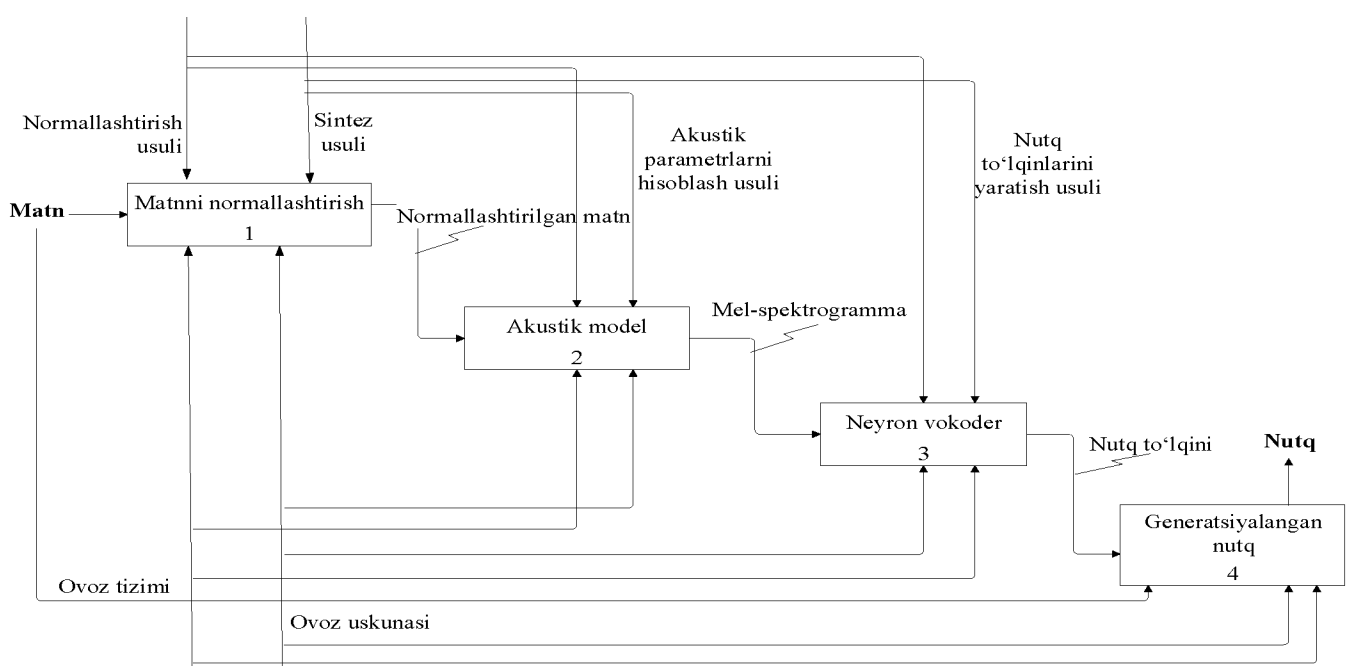
- tezlik: o'qitish va ma'lumotlarni chiqarish tezligi an'anaviy parallel WaveNet (ya'ni ClariNet)dan 4,82 (1,96) baravar yuqori;

- yuqori sifat: model 4.16 MOS (transformator asosidagi TTS da) aniqlikka ega bo'lib, filterlashga asoslangan eng yaxshi ClariNetga raqobatbardosh bo'la oladi.

Shu sababli, ushbu ilmiy tadqiqotda GAN modeli asosida o'zbek nutqini sintez qilish uchun vokoder tanlangan. Parallel Wave GAN (PWG) asosidagi neyron tarmoq vokoderining arxitekturasi 3-rasmda keltirilgan.



3-rasm. Ishlab chiqilgan PWGga asoslangan neyron tarmoq vokoder tuzilishi



4-rasm. O'zbek tili nutqini sintez qilish tizimi IDEF0 diagrammasi

Dissertatsiyaning «**Reabilitatsiya jarayonining dialog interfeysi komponentlarini yaratish**» deb nomlangan uchinchi bobida beshta paragraf mavjud. Ularda mel-spektrogrammasi asosida o‘zbek tili nutqini tahlil qilish uchun ishlab chiqilgan algoritmlar, taklif etilayotgan modelni o‘qitish va sinovdan o‘tkazish jarayoni tavsifi, formant va nutq signallarini taqqoslash asosida o‘zbek tovushlarini tahlil qilish, shuningdek, diktorni autentifikatsiya qilish uchun ishlab chiqilgan algoritmlar keltirilgan.

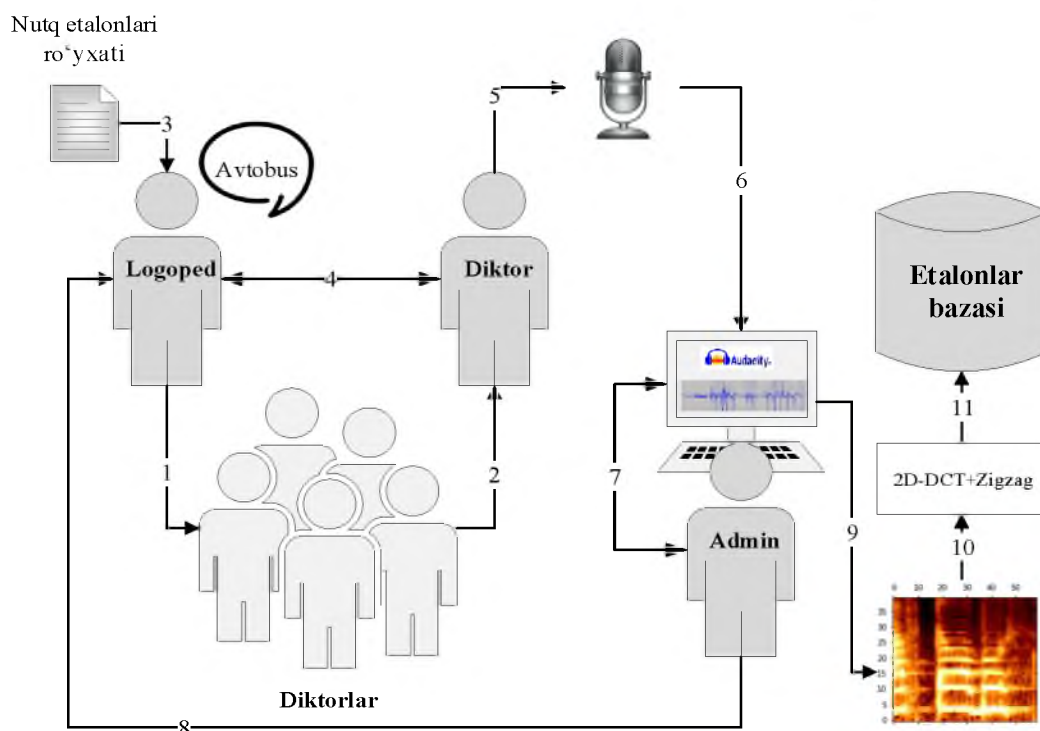
Nutqni reabilitatsiya qilish jarayonining muhim tarkibiy qismi nutqni sintez qilish va tahlil qilish protseduralaridan iborat dialog interfeysidir. Nutq signallarini tahlil qilish jarayoni oldindan ishlov berish, spektrogramma tuzish, nutqning xususiyatlarini hisoblash va tasniflashdan iborat. Nutqni tahlil qilish tizimlarida eng muhimi nutq signalini parametrlash bosqichidir. Parametrik xususiyatlar quyidagi talablarga javob berishi kerak: informativlik, ma’lumotlar miqdorini minimallashtirish, diktorga bog‘liq emaslik, bir xillik.

Nutqni tahlil qilish vazifalarida Furrye va mel spektrogrammalari informativ belgilar sifatida ishlatilgan. Furrye spektridan farqli o‘laroq, mel spektrogrammasi gertsda emas, balki mel-shkalasida ifodalangan chastotadir. Mel-spektrogrammasiga o‘tish kiruvchi spektrogrammaga mel-filtrlarini qo‘llash orqali amalga oshiriladi. Shunday qilib, mel-spektrogrammasi signaldan muhim xususiyatlarni ajratib oladi. Bu bilan uning siqilish darajasini oshirishga imkon beradi.

Quyida esa reabilitatsiya tizimining alohida tarkibiy qismlari ko‘rib chiqilgan.

#### **Bolalar nutq etalonlarini shakllantirish.**

Dissertatsiya doirasida cheklangan lug‘at asosida maktabgacha yoshdagi bolalar ovozlari etalonlari bazasi yaratildi. Belgilangan o‘zbek tili lug‘ati uchta toifasidagi nutq birliklarining tashkil topgan: tovushlar, bo‘g‘inlar va uchta qiyinlik darajasidagi so‘zlar.



**5-rasm. Bolalar etalonlari bazasini shakllantirish jarayoni**

Har bir nutq birligi mikrofon orqali maxsus, shovqinsiz muhitda yozilishi kerak. 5-rasmda har bir nutq etalonini yozish olish jarayonining batafsil ketma-ketligi keltirilgan, unga ko‘ra:

1-2 qadamlar. Logoped bolaning nutq qobiliyatini tekshirish asosida bolalar (diktorlar) orasidan nomzodlarni tanlaydi.

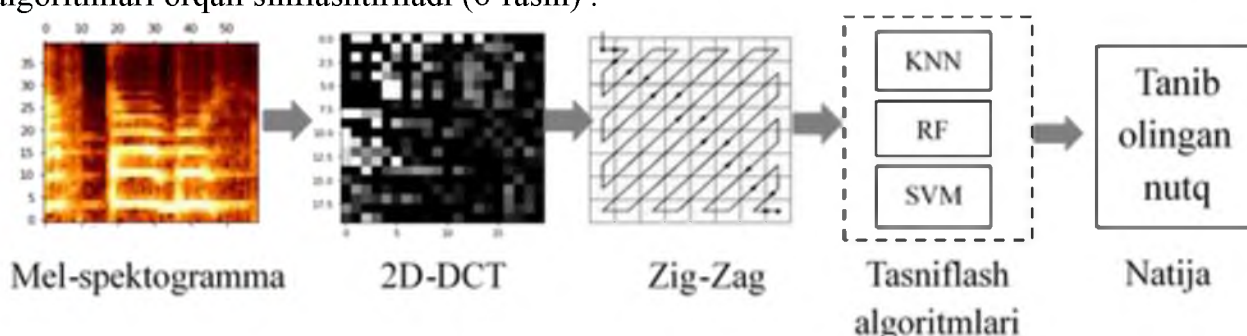
3-4 qadamlar. Logoped nutq birligi (tovush, bo‘g‘in, so‘z)ni talaffuz qiladi. Keyin diktor logoped ketidan joriy nutq birligini takrorlaydi.

5-7 qadamlar. Administrator diktorning nutqini mikrofon orqali yozib oladi va saqlaydi.

8-11 qadamlar. Administrator yozib olingan audio etalonlarni kerakli formatga o‘zgartiradi va nutq bazasiga saqlaydi.

### **O‘zbek tili bo‘g‘inlari va so‘zlarini tahlil qilish jarayonining quyi tizimi.**

Taklif etilgan bo‘g‘inlar va so‘zlarni tahlil qilish algoritmidan mel-spektrogrammalari maxsus 2D-Diskret kosinus o‘zgartirishi (2D-DCT) dan o‘tkaziladi. Shu tariqa tasvirlardan informativ xususiyatlar ajratib olinadi. So‘ngra ular “Zig-zag” skanerlash algoritmi bilan saralanadi va ketma-ket ishlov berish natijasida KNN (eng.k-nearest neighbour), RF (eng. Random Forest), SVM (eng. Support Vektor Machine) mashinali o‘qitish tasniflash algoritmlari orqali sinflashtiriladi (6-rasm).



**6-rasm. Mel-spektrogramma asosida nutqni tahlil qilish algoritmining jarayonlar ketma-ketligi**

Nutq tahlilining aniqligi to‘g‘ridan-to‘g‘ri mashinali o‘qitish algoritmlarini o‘quv tanlanmasi hajmi va sifatiga bog‘liq. Taklif etilayotgan algoritmgamuvofiq nutq signalini tasniflash uchun o‘quv tanlanmasi har bir yozib olingan audio faylga “Zig-zag” skanerlash bilan 2D-DCT asosida ishlab chiqilgan algoritmni qo‘llash natijasida olingan etalon xususiyatlar vektoridan iborat bo‘lishi kerak. Etalon-tovush (bo‘g‘in, so‘z)ni to‘g‘ri talaffuzining asosiy namunasi bo‘lgan audio faylning xususiyatlari to‘plamidan tashkil topgan.

**3-jadval.**

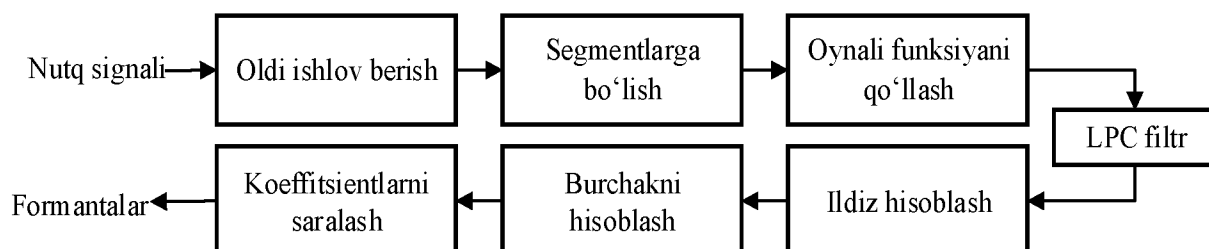
**Nutq tahlili modelining test natijalari**

	KNN	RF	SVM	Sinflashning o‘rtacha aniqligi
Tovushlar	95,6%	92,3%	93,5%	93,8%
Bo‘g‘inlar	94,3%	94,8%	92,5%	93,9%
So‘lar	95,7%	93,4%	93,8%	94,3%
Modellar bo‘yicha sinflashtirishning o‘rtacha aniqligi	95,2%	93,5%	93,3%	94%

Sinov natijalariga ko‘ra (3-jadval) tanlangan o‘zbek tovushlari, bo‘g‘inlari va so‘zlarining eng aniq tasnifiga KNN qo‘llanilganda erishildi. Uning o‘rtacha tasniflash aniqligi RF (93,5%) va SVM (93,3%) dan oshib, 95,2% ni tashkil qiladi.

### Unli tovushlarni tahlil qilish quyi tizimi.

Dissertatsiya ishida o'zbek tili tovushlarini formant to'plam yordamida tahlil qilish algoritmi ham taklif etilgan. Formanta chastota, kenglik va amplituda bilan tavsiflanadi. Formanta - bu spektr grafigidagi ba'zi amplituda ko'tarilishi, uning chastotasi esa bu ko'tarilishning eng yuqori chastotasidir. Chiziqli bashorat qilish usuli asosida formantlarni hisoblash uchun taklif qilingan ketma-ketlik 7-rasmda keltirilgan.



7-rasm. Formantani hisoblash jarayoni

Odatda F1, F2, ... ,F6 mos ravishda uchdan oltitagacha formantalar hisoblab chiqiladi va belgilanadi. Shunga qaramay, unli tovushlar birinchi uchta formantada (F1, F2, F3) yetarlicha yuqori aniqlik bilan farqlanadi. Birinchi ikkita formanta {F1, F2} ko'proq ma'lumotga ega va bir-biri bilan chambarchas bog'liq, F3 esa intonatsiya ma'lumotlarini o'zida aks etadi. Olingan formanta xususiyatlarni tasniflash KNN (95%), RF (94%), SVM (95%) algoritmlari yordamida amalga oshirildi.

### Nutq signallarini taqqoslash quyi tizimi.

Nutqni tahlil qilish tizimlarida signallarni taqqoslash va ularning o'xshashligini foizda baholash vazifasi muhim vazifalardan biridir. Nutq signallarini taqqoslash uchun bir qator algoritmlar mavjud, bular DTW (eng. Dynamic Time Wrapping), korrelyatsiya, diskret masofa, Evklid masofa kabilardir. Ushbu algoritmlarning ba'zilarini natijalarini foiz yoki ball shkalasida talqin qilish imkoni mavjud emas.

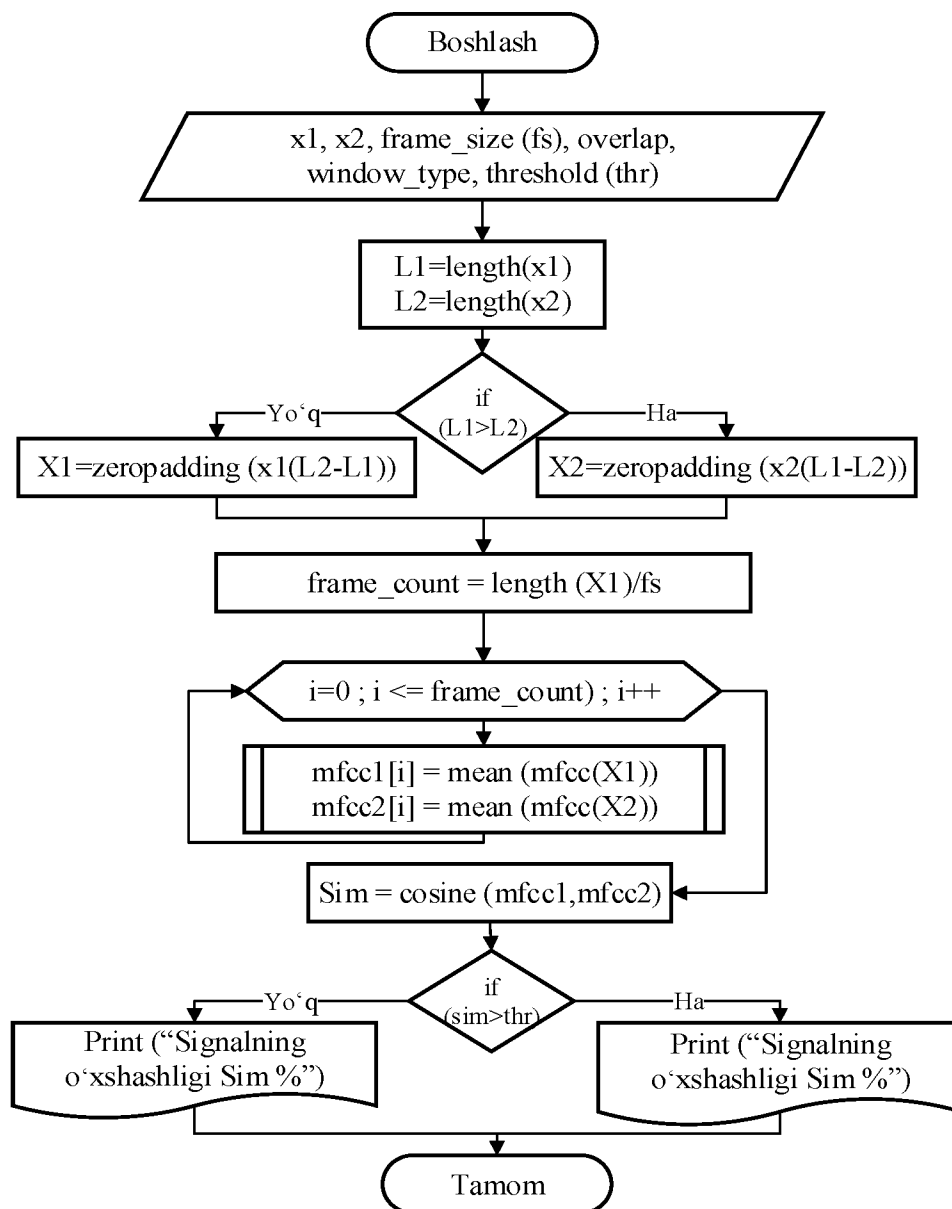
Ma'lumki, ikkita nutq signalini taqqoslash informativ xususiyatlarni taqqoslash orqali amalga oshiriladi, bu mel-koeffitsientlari, chiziqli bashorat qilish usulining koeffitsientlari, formantlar, nutq signallaridan hisoblangan kepsstral koeffitsientlar bo'lishi mumkin.

MFCC koeffitsientlarining kosinus o'xshashligi asosida ikkita signalni taqqoslash uchun algoritmi taklif qilingan. Taklif qilingan algoritmning blok-sxemasi 8-rasmda keltirilgan.

Kosinus o'xshashligining asosiy g'oyasi quyidagi formula bo'yicha koeffitsientlarning ikkita MFCC vektorlari orasidagi burchak kosinusini hisoblashdan iborat:

$$\text{CosineSimilarity} = S_c(A, B) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$$

bu erda A va B ikkita nutq signalining MFCC koeffitsientlari.



8-rasm. Nutq signallarini taqqoslash algoritmining blok-sxemasi

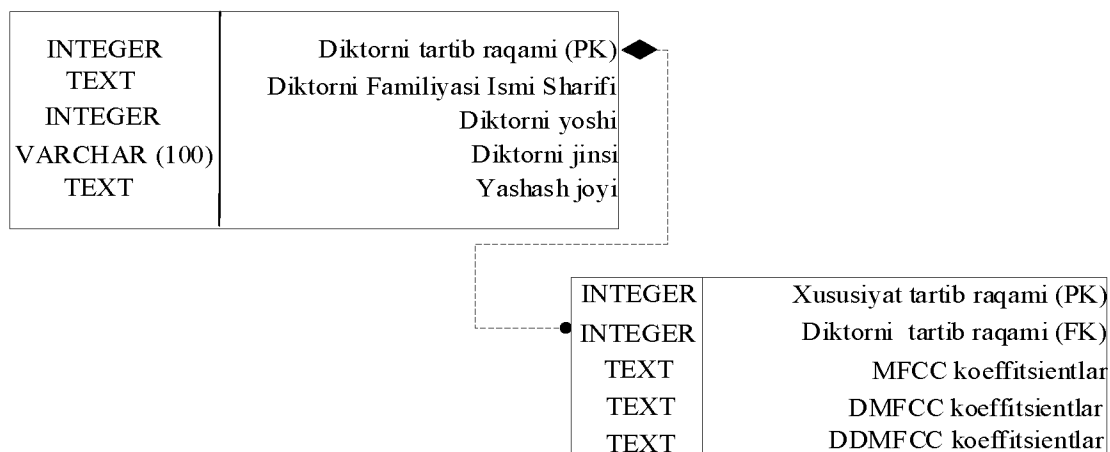
### Autentifikatsiya quyi tizimi.

Bemorlarning ma'lumotlar bazalarini yaratish va saqlash uchun diktorlarni (bemorlarni) autentifikatsiya qilish quyi tizimi ishlab chiqilgan. Uchta mustaqil qismdan iborat algoritm taklif qilindi:

1. Axborot xususiyatlarini tanlash (tavsiflangan algoritmgaga muvofiq kirish nutq signalini parametrlash).
2. Xususiyatlar vektori shaklida ma'lum diktor uchun etalonni yaratish.
3. Etalonlar bilan taqqoslash uchun GMM asosida qaror qabul qilish (eng. Gaussian Mixture Model).

9-rasmda IDEF1x asosida ma'lumotlar bazasi jadvallari jismoniy bog'lanish sxemasi keltirilgan. IDEF1x metodologik sxemasi IDEF1 modelning takomillashtirilgan versiyasi bo'lib, relyatsion ma'lumotlar bazalarini modellashtirish uchun ishlatiladi.





**9-rasm. Ma'lumotlar bazasi jadvallarini jismoniy bog'lanishi IDEF1x metodologik sxemasi**

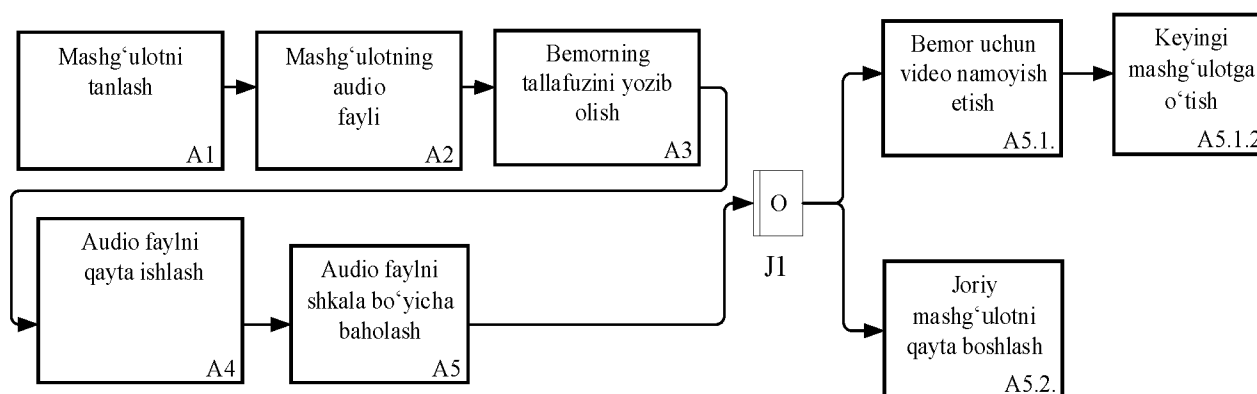
Taklif etilayotgan algoritm asosida diktor autentifikatsiyasining aniqligi 96,4% ni tashkil etdi.

Dissertatsiyaning «Eshitish qobiliyati buzilgan bemorlarni reabilitatsiya qilish uchun axborot tizimini yaratish» deb nomlangan to'rtinchi bobida bemorlarni reabilitatsiya qilish texnologiyasi, ishlab chiqilgan axborot tizimining tuzilishi, nutqni reabilitatsiya qilish uchun ishlab chiqilgan axborot tizimining interfeysi va uni amaliy qo'llash samaradorligi ko'rsatkichlari keltirilgan.

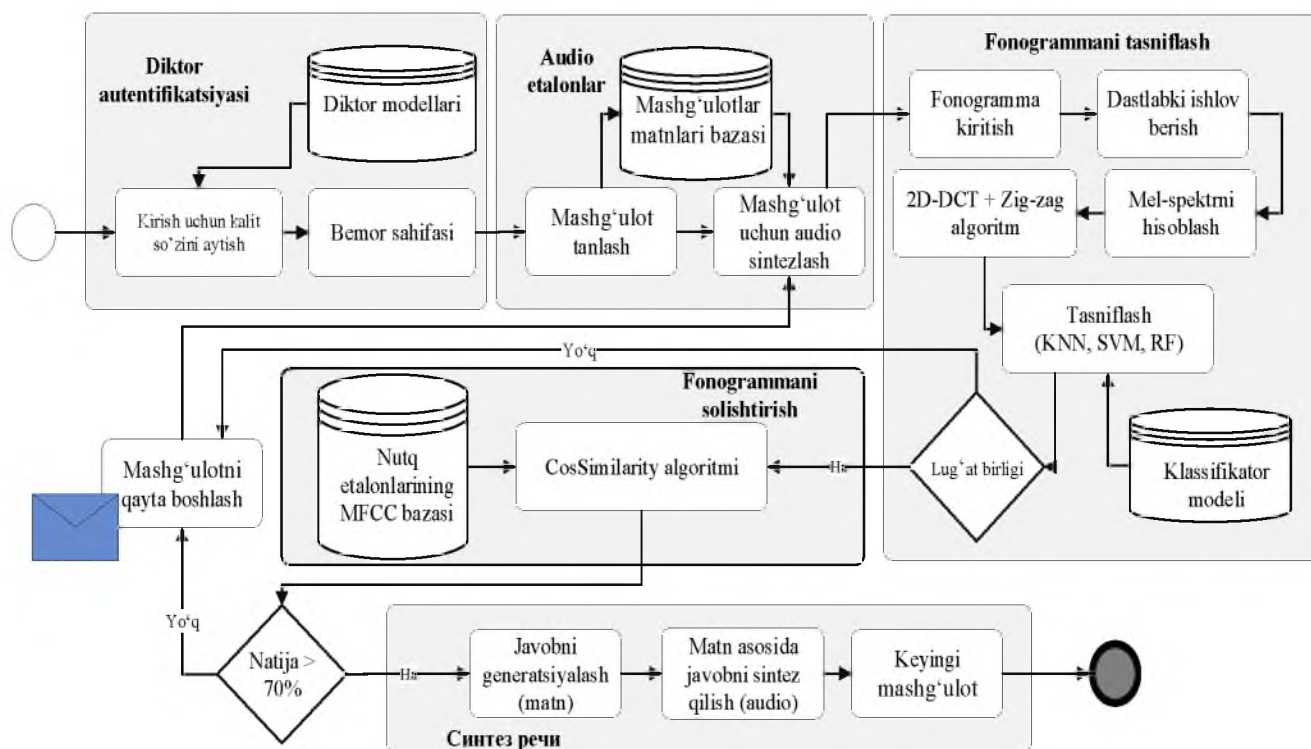
Axborot tizimi koxlear implantatsiyadan keyin bemorlarni reabilitatsiya qilish uchun mo'ljallangan. Koxlear implantatsiya - bu eshitish qobiliyati cheklangan odamlarni reabilitatsiya qilish bo'yicha chora-tadbirlar tizimi bo'lib, u uch bosqichni o'z ichiga oladi.

1. Jarrohlik uchun nomzodlarni tanlash;
2. Jarrohlik amaliyotini o'tkazish;
3. Operatsiyadan keyingi nutq va eshitish reabilitatsiyasi.

Strukturaviy usuldan foydalangan holda harakatlar o'rtasidagi bog'liqliklarni namoyon etadigan IDEF3 modellashtirish metodologiyasiga asoslangan bemor tomonidan har bir blokning mashqlarini bajarish jarayonining bosqichlari ketma-ketligi 10-rasm va 11-raslarda BPMN modellashtirishga asoslangan axborot tizimi doirasida bemorlarning nutqini reabilitatsiya qilishning batafsil jarayoni keltirilgan.



**10-rasm. Bemorlarning nutqini reabilitatsiya qilish mashqlarini bajarish jarayonining funksional IDEF3 modeli**



**11-rasm. Axborot tizimi doirasida bemorlarning nutqini reabilitatsiya qilishning BPMN diagrammasi**

Nutqning eng kichik birligi tovushdir. Surdopedagoglar o‘zbek tilida 17 ta tovushlarni ajratishadi va ularni o‘z o‘rnida ikkita konsentrga guruhlashadi (4-jadval). Ushbu konsentrlarga asoslangan ma’lumotlar bazasi yordamida nutq reabilitatsiya jarayoni olib borilmoqda.

**4-jadval**

**O‘zbek tovushlarining konsertlar bo'yicha tasnifi**

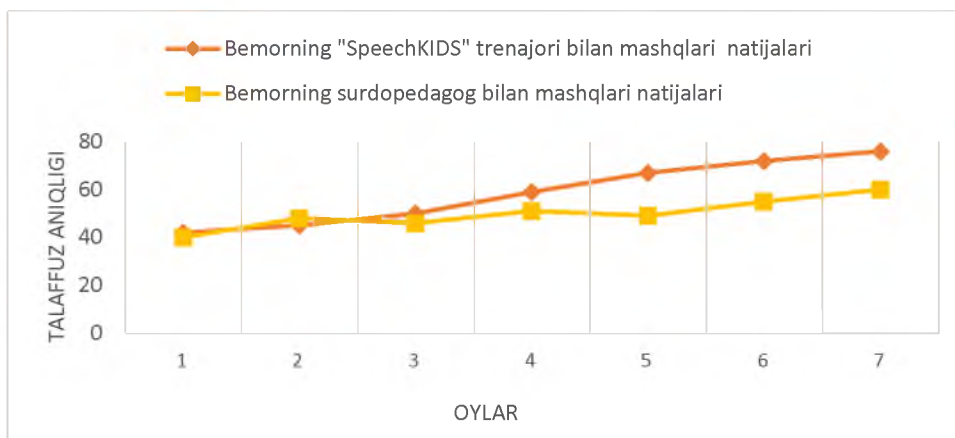
<b>1-guruh konsetri</b>	A	O	U	E	I	O‘	P	T	K	F	S	Sh	V	L	R	M	N
<b>2- guruh konsetri</b>					Y		B	D	Q G G‘		Z	X H	Ch J				NG

Ma’lumki, o‘zbek tilida “Q”, “G”, “G’” harflari eng qiyin talaffuz qilinadigan tovushlardir.

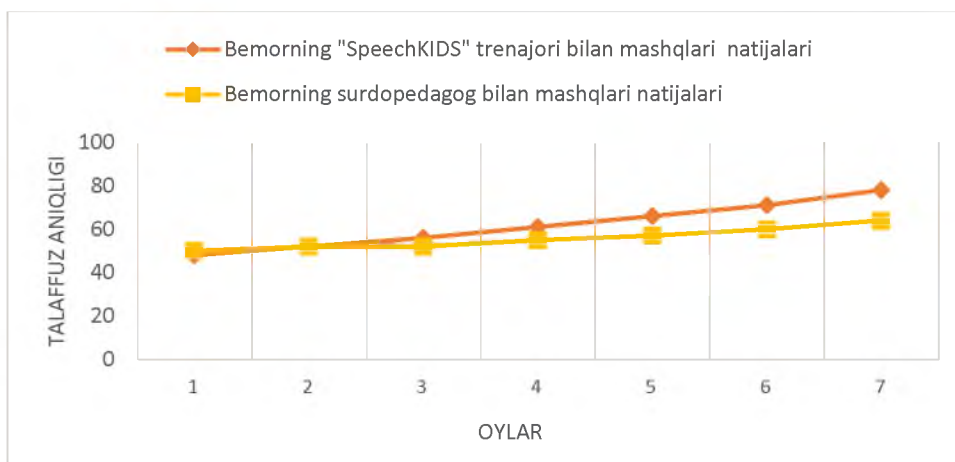
Axborot tizimini baholash uchun qator koxlear operatsiyani o‘tkazgan bolalar jalb etilgan. Bolalarni reabilitatsiya qilish 7 oy davomida amalga oshirilgan. Shu bilan birga, axborot tizimini baholash uchun reabilitatsiyaning ikkita amaldagi usuli tanlangan: surdopedagog bilan an’anaviy; axborot tizimi (trenajor).

Quyidagi 12-14-rasmlarda tavsiya etilgan va mavjud reabilitatsiya texnologiyalari bo‘yicha talaffuz aniqligini taqqoslash natijalari taqdim etilgan.

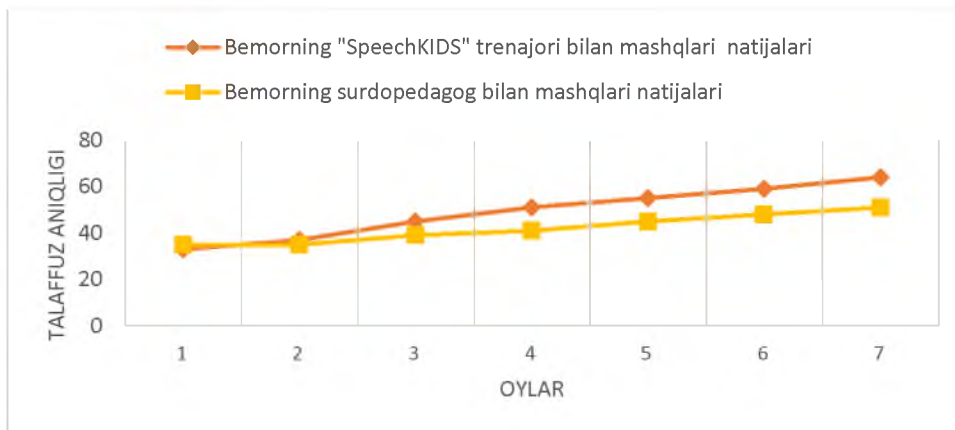
Diagrammalarga ko‘ra, ishlab chiqilgan axborot tizimi (SpeechKids) reabilitatsiya jarayonini o‘rtacha 1,5-2 baravarga tezlashtirishadi.



**12-rasm. «Q» tovushini talaffuz qilish natijalari**



**13-rasm. «G» tovushini talaffuz qilish natijalari**



**14-rasm. «G'» tovushini talaffuz qilish natijalari**

## XULOSA

“Nutq texnologiyalari asosida eshitish qobiliyati cheklangan bemorlarni reabilitatsiya qilish axborot tizimini ishlab chiqish” mavzusidagi olib borilgan tadqiqotning asosiy natijalari quyidagilardan iborat:

1. Nutqni qayta ishlash modellari, usullari va algoritmlari tahlil qilindi, o‘zbek nutqini tahlil qilish va sintez qilishning intellektual algoritmlari asosida eshitish

qobiliyati cheklangan bemorlarni reabilitatsiya qilish axborot tizimining arxitekturasi yaratildi.

2. Nutqni sintez qilish tizimi uchun BPMN modeli, audio signalni matn va uning transkripsiyasi bilan sinxronlashtirish algoritmlari asosida o'zbek tili nutq bazasi yaratildi. Nutq bazasi diktor tomonidan o'qilgan 11 ming jumla asosida shakllantirilgan bo'lib, uning umumiy hajmi 29 soatni tashkil etadi.

3. Spektral tahlil algoritmlari asosida nutq parametrlarini aniqlash, kepstral koeffitsientlarni hisoblash uchun maxsus akustik model ishlab chiqildi. Akustik model konvolyutsion va rekurrent chuqur o'rganish neyron tarmoqlari asosida qurilgan.

4. Tacotron2 va Parallel WaveGAN neyron arxitekturalari asosida o'zbek nutqini sintez qilish uchun TTS tizimi ishlab chiqilgan. Ushbu sintez tizimi o'zbek tilining ko'p soatlik nutq bazasida asosida o'qitilgan va MOS sifatni baholash shkalasi bo'yicha 5 dan 4,36 ball bilan matnni nutqqa aylantirish imkonini berdi.

5. Bemor bolalarni reabilitatsiya qilish jarayonining samarali dialog interfeysini yaratish uchun o'zbek tilidagi tovushlar, bo'g'inlar va so'zlarni tahlil qilish quyi tizimi ishlab chiqilgan. Nutqning talaffuz qilinadigan tarkibiy qismlarini tahlil qilish uchun ishlab chiqilgan quyi tizimda 2D Diskret kosinus o'zgartirishi (xususiyatlarni ajratib olish uchun) va "Zig-zag" skanerlash algoritmlari yordamida amalga oshirilgan. Olingan o'zgartirish natijalari ketma-ket KNN, RF SVM tasniflash protseduralaridan o'tkaziladi va ilgari tayyorlangan nutq bazasi etalonlariga nisbatan to'g'ri talaffuz mezoniga muvofiq baholandi. Sinov tovushlar, bo'g'inlar va so'zlarni tahlil qilish algoritmlarining o'rtacha aniqligini 94% tashkil etdi.

6. Ommaviy reabilitatsiya jarayonida bemorlarning ma'lumotlar bazalarini shakllantirish va qo'llab-quvvatlash uchun diktorlarni (bemorlarni) autentifikatsiya qilish uchun dasturiy modul ishlab chiqilgan. Umumiy algoritm uchta mustaqil qismni o'z ichiga oladi: - talaffuz qilingan parolni parametrlash; - xususiyat vektori shaklida namunani qurish; - kiruvchi nutqni Gauss aralashmalari asosida etalon bilan taqqoslash.

7. Reabilitatsiya texnologiyasi asosida koxlear implantatsiyadan keyin bolalarni reabilitatsiya qilish uchun trenajor shaklida axborot tizimi yaratildi. Mashqlarni bajarish jarayonining bosqichlari ketma-ketligi IDEF3 modellashtirish metodologiyasiga muvofiq amalga oshirildi. Multimediya rejimida bemorlar va o'qituvchilarning o'zaro ta'sir rejimlari uchun foydalanuvchi interfeyslari, shu jumladan grafik interfeyslar ishlab chiqilgan. Buning uchun bolalarga xos rangli rasmlardan tashkil topgan tasvirlarning maxsus multimediya bazasi yaratildi.

8. Tadqiqot ishida ishlab chiqilgan reabilitatsiya axborot tizimidan amaliy foydalanish samaradorligi ko'rsatkichlari keltirilgan. Tizim relyatsion ma'lumotlar bazasi, nutq interfeysi, amaliy nutqni tahlil qilish va sintez qilish quyi tizimlaridan iborat. Respublika ixtisoslashtirilgan pediatriya ilmiy-amaliy tibbiyot markazida reabilitatsiya jarayonining samaradorligini sinovdan o'tkazish va baholash shuni ko'rsatdiki, ishlab chiqilgan axborot tizimi yordamida maktabgacha yoshdagi bolalarni nutqini reabilitatsiya qilish muddatlari 1,5 – 2 baravarga qisqartirilishi kuzatildi.

**НАУЧНЫЙ СОВЕТ DSc.13/30.12.2019.Т.07.01 ПО ПРИСУЖДЕНИЮ  
УЧЕНЫХ СТЕПЕНЕЙ ПРИ ТАШКЕНТСКОМ УНИВЕРСИТЕТЕ  
ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ**

---

**ТАШКЕНТСКИЙ УНИВЕРСИТЕТ ИНФОРМАЦИОННЫХ  
ТЕХНОЛОГИЙ**

**АБДУЛЛАЕВА МАЛИКА ИЛЬХАМОВНА**

**РАЗРАБОТКА ИНФОРМАЦИОННОЙ СИСТЕМЫ РЕАБИЛИТАЦИИ  
ПАЦИЕНТОВ С НАРУШЕНИЕМ СЛУХА НА БАЗЕ РЕЧЕВЫХ  
ТЕХНОЛОГИЙ**

05.01.10 – Информационные системы и процессы

**АВТОРЕФЕРАТ ДИССЕРТАЦИИ  
ДОКТОРА ФИЛОСОФИИ (PhD) ПО ТЕХНИЧЕСКИМ НАУКАМ**

**Ташкент – 2023**

Тема диссертации доктора философии (PhD) по техническим наукам зарегистрирована в Высшей аттестационной комиссии при Кабинете Министров Республики Узбекистан за B2023.2.PHD/T3546

Диссертация выполнена в Ташкентском университете информационных технологий.

Автореферат диссертации на трех языках (русский, узбекский, английский (резюме)) размещен на веб-странице ([www.tuit.uz](http://www.tuit.uz)) и на Информационно-образовательном портале «Ziynet» ([www.ziynet.uz](http://www.ziynet.uz)).

**Научный руководитель:** Мусаев Мухаммаджон Махмудович  
доктор технических наук, профессор

**Официальные оппоненты:** Саидов Абдусобир Абдурахмонович  
доктор технических наук, профессор

Элов Жамшид Бекмуродович  
доктор философии технических наук (PhD), доцент

**Ведущая организация:** Самаркандский государственный  
университет имени Шарофа Рашидова

Защита диссертации состоится «23» ИЮНЬ 2023 г. в 14<sup>00</sup> часов на заседании научного совета DSc.13/30.12.2019.T.07.01 при Ташкентском университете информационных технологий. (Адрес: 100084, г. Ташкент, ул. Амира Темура, 108. Тел.: (99871) 238-64-43; e-mail: [iktuit@tuit.uz](mailto:iktuit@tuit.uz)).

С диссертацией можно ознакомиться в Информационно-ресурсном центре Ташкентского университета информационных технологий (регистрационный номер № 278). (Адрес: 100084, г. Ташкент, ул. Амира Темура, 108. Тел.: (99871) 238-64-70).

Автореферат диссертации разослан «10» ИЮНЬ 2023 года.  
(протокол рассылки № 17 от «06» ИЮНЬ 2023 г.).



**М.А.Рахматуллаев**  
Заместитель председателя научного совета по  
присуждению ученых степеней,  
доктор технических наук, профессор

**Н.О.Рахимов**  
Ученый секретарь научного совета по  
присуждению учёных степеней,  
доктор технических наук, доцент

**У.Р.Хамдамов**  
Председатель научного семинара  
при научном совете по присуждению  
ученых степеней, доктор технических наук,  
доцент

## **ВВЕДЕНИЕ (аннотация диссертации доктора философии (PhD))**

**Актуальность и востребованность темы диссертации.** В мире особое внимание уделяется разработке информационных систем на базе речевых технологий, так как они охватывают широкие области применения. Создаваемые в настоящее время системы речевого интерфейса непосредственно связаны с анализом и синтезом речи. Известно, что речевой диалог является основным способом человеческого общения, то разработка интеллектуальной системы на основе речевых компонентов является одним из основных направлений искусственного интеллекта. Поскольку речевой диалог – главный метод общения людей, разработка интеллектуальной системы с речевыми компонентами являются основными задачами искусственного интеллекта. Речевой интерфейс состоит из таких базовых речевых технологий, как анализ и синтез речи. В развитых странах мира, как США, Российская Федерация, Китай, Япония и другие ведутся научные исследования в области развития теоретических и практических работ, направленных на совершенствование существующих методов и алгоритмов анализа и синтеза речи, на разработку методов интеллектуальной обработки речевых сигналов. Сегодня, разработка систем с возможностью обеспечения полноценного диалога человека и машиной с помощью речи является одной из главных задач в области цифровой обработки сигналов и искусственного интеллекта.

В мире увеличивается рождение детей с нарушениями слуха и речи. В связи с этим ожидается, что к 2027 году рынок кохлеарных имплантов вырастет в среднем на 8,57%. Поэтому разработка информационной системы для реабилитации пациентов после кохлеарной имплантации является актуальным вопросом. Информационная система реабилитации слухоречевых навыков пациента реализуется на базе речевого интерфейса, с инструментами анализа и синтеза речи.

В республике в соответствии с постановлением Президента Республики Узбекистан<sup>3</sup> определены задачи «...широкого применения технологий искусственного интеллекта для улучшения качества оказания услуг в интересах населения». Благодаря поддержке развития информационных технологий и искусственного интеллекта, в нашей республике проводятся научные исследования по интеллектуальной обработке речевых сигналов, автоматическому анализу произносимых звуков, слогов и слов, голосовой аутентификации, созданию систем голосового управления, анализу и синтезу узбекской речи. Согласно приоритетным направлениям поставленным в рамках стратегии «цифровой Узбекистан-2030», в республике реализуются комплексные меры по активному развитию и широкому внедрению современных информационно-коммуникационных технологий, а именно изучению и применению на практике возможностей использования технологий искусственного интеллекта, алгоритмов машинного обучения, анализа больших данных во все отрасли и сферы, прежде всего, в здравоохранение, а также в государственное управление, образование и сельское хозяйство<sup>4</sup>.

---

<sup>3</sup> Постановление Президента Республики Узбекистан от 17.02.2021 г. № ПП-4996 «О мерах по созданию условий для ускоренного внедрения технологий искусственного интеллекта»

<sup>4</sup> Указ Президента Республики Узбекистан от 05.10.2020 г. № УП-6079 «О мерах по созданию условий для ускоренного внедрения технологий искусственного интеллекта»



Результаты исследований данной диссертации в значительной степени способствует в реализации Постановлений Президента Республики Узбекистан ПП-4996 от 17 февраля 2021 года «О мерах по созданию условий для ускоренного внедрения технологий искусственного интеллекта», Указ Президента Республики Узбекистан УП-6079 от 5 октября 202 года «Об утверждении стратегии «Цифровой Узбекистан 2030» и мерах по её эффективной реализации».

**Соответствие исследования приоритетным направлениям развития науки и технологий Республики.** Данное диссертационное исследование выполнено в соответствии с приоритетным направлением развития науки и технологии республики IV. «Информатизация и развитие информационно-коммуникационных технологий».

**Степень изученности проблемы.** На сегодняшний день в научных центрах развитых стран проводятся ряд исследований по разработке методов и моделей обработки речевых сигналов, построению нейросетевых архитектур для эффективного преобразования текстовой информации, созданию многочасовых речевых корпусов для систем генерации речи. Такие ученые, как L.Rabiner, J.Shen, R.Yamamoto, J.Miller, D.Klatt, O.Arik, W.Ping, P.Duhamel, И.Кипяткова, Л.Ронжин и другие внесли значимый вклад в развитие и совершенствование методов анализа и синтеза речи, а также в их практическое применение.

Среди узбекских ученых по разработке алгоритмов и систем на базе анализа и синтеза речи можно отметить работы М.Арипова, М.Мусаева, У.Хамдамова, Н.Мамадова и других.

В настоящее время ведутся исследования по созданию информационных систем позволяющих распознавать (анализировать речь) и генерировать ответы с помощью синтезатора речи. Анализ проведенных исследований показывает, что не до конца решены задачи анализа и синтеза узбекской речи на базе алгоритмов искусственного интеллекта.

**Связь темы диссертационного исследования с планами научно-исследовательских работ высшего образовательного учреждения, где выполнена диссертация.** Диссертационное исследование выполнено в рамках научно исследовательской работы Ташкентского университета информационных технологий имени Мухаммада ал-Хоразми по проектам №-ILM-202101222 - «Разработка программ обработки узбекской речи на основе алгоритмов искусственного интеллекта для судебной сферы» (2021-2023) и №-ФЗ-2020103164 - «Разработка технологий реабилитации и мониторинга слуха и речи слабослышащих детей на основе современных компьютерных спектрально-аналитических технологий» (2021-2024).

**Целью исследования** является создание информационной системы реабилитации пациентов с нарушением слуха на базе алгоритмов анализа и синтеза узбекской речи с помощью глубоких нейронных сетей.

**Задачи исследования:**

исследование современных технологий реабилитации пациентов после кохлеарной имплантации, выбор архитектур, аппаратных и программных средств для восстановления слухоречевых навыков;

создание информационных моделей и мультимедийной базы данных для процесса реабилитации речи пациентов после кохлеарной имплантации;



разработка алгоритмов нормализации, синхронизации и тестирования данных для создания многочасовой речевой базы узбекского языка;

разработка моделей и алгоритмов синтеза узбекской речи на базе нейронных сетей;

разработка алгоритмов анализа речевых единиц узбекского языка для обеспечения диалогового интерфейса;

создание информационной системы реабилитации пациентов с удобным пользовательским речевым интерфейсом.

**Объектом исследования** является информационная система процесса реабилитации слухоречевых навыков детей после кохlearной имплантации.

**Предметом исследования** являются программный комплекс реабилитации пациентов после кохlearной имплантации, алгоритмы анализа узбекской речи в целях анализа и аутентификации диктора, модели нейросетевых архитектур для синтеза узбекской речи, алгоритм нормализации узбекского текста, программные средства формирования многочасовой речевой базы узбекского языка.

**Методы исследования.** Для проведения исследования в диссертационной работе были использованы следующие методы: теория вероятностей; интеллектуальная обработка данных; математическая статистика; линейная алгебра; цифровая обработка сигналов; системы управления реляционными базами данных; объектно-ориентированное программирование; алгоритмы принятия решений.

**Научная новизна исследования** состоит в следующем:

разработаны программные компоненты информационной системы реабилитации пациентов с нарушением слуха на основе технологии MVT;

разработан речевой интерфейс на базе моделей анализа и синтеза речи, построена IDEF0 диаграмма разработанной TTS системы узбекского языка на основе глубоких нейронных сетей;

разработаны алгоритмы процесса синтеза узбекской речи на основе речевой базы, а также создана база эталонов речевых единиц специализированного словаря реабилитации пациентов;

разработан алгоритм процесса анализа фонем, слогов и слов процесса реабилитации слухоречевых навыков пациента.

**Практические результаты** исследования состоят из следующих:

разработан пользовательский речевой интерфейс процесса реабилитации речи пациентов дошкольного возраста;

сформирована единая речевая база эталонов, состоящая из звуков, слогов и слов узбекского языка, произнесенных детьми без дефекта речи от 3 до 7 лет;

создана база статических и динамических мультимедийных данных, используемых в технологии реабилитации;

разработана информационная система (тренажер) реабилитации детей после кохlearной имплантации на основе информационной системы «SpeechKIDS»;

определены показатели эффективности применения информационной системы в процессе реабилитации речи у пациентов после кохlearной имплантации.

**Достоверность результатов исследования** подтверждается заключениями испытаний, полученных на основе использования разработанной информационной системы в процессе реабилитации детей после кохlearной имплантации, а также результатами экспериментов, полученных на основе применения предложенных алгоритмов и архитектур из диссертации в научных статьях, обсужденных на международных научных конференциях.

### **Научная и практическая значимость результатов.**

Научная значимость результатов исследования определяется тем, что разработанные алгоритмы являются определенным вкладом в дальнейшее развитие теоретических основ анализа и генерации узбекской речи на основе алгоритмов искусственного интеллекта.

Практическая значимость результатов исследования определяется тем, что разработанная информационная система позволяет значительно ускорить процесс реабилитации пациентов дошкольного возраста с нарушениями слуха.

**Внедрение результатов исследования.** Разработанная информационная система на основе методов, алгоритмов и программных средств анализа и синтеза речи с помощью алгоритмов искусственного интеллекта внедрена:

в Республиканский специализированный научно-практический медицинский центр педиатрии (РСНПМЦ Педиатрии), в Отделении врожденных и приобретенных заболеваний ЛОР-органов для реабилитации слуха и речи пациентов после кохлеарной имплантации (Справка Министерства цифровых технологий Республики Узбекистан от 11 мая 2023 года № 33-8/3119). Использование тренажера позволило ускорить процесс реабилитации детей на 1,5-2 раза по сравнению с традиционным методом реабилитации под руководством сурдопедагога. Эффект научной работы достигнут за счет использования разработанных речевых технологий в процессе реабилитации;

в ООО «Nova research and resolution» для устранения дефектов речи у детей с нарушением речи (Справка Министерства цифровых технологий Республики Узбекистан от 11 мая 2023 года № 33-8/3119). В результате разработанная информационная система на основе предложенных алгоритмов позволила сократить количество уроков с логопедом, таким образом ускорить процесс реабилитации детей с дефектами речи. Главным преимуществом информационной системы является возможность онлайн реабилитации.

**Апробация результатов исследования.** Основные теоретические и практические результаты диссертации обсуждались на 8 международных и 2 республиканских научно-практических конференциях.

**Опубликованность результатов исследования.** Основные результаты исследования опубликованы в 20 научных работах, 6 из которых опубликованы в журналах, рекомендованных Высшей аттестационной комиссией Республики Узбекистан для публикации основных научных результатов докторских диссертаций, в том числе 3 в зарубежных журналах и 3 в республиканских журналах, а также получены 4 свидетельства об официальной регистрации программ для ЭВМ.

**Структура и объем диссертации.** Диссертация содержит 116 страниц и состоит из введения, четырех глав, заключения, списка условных сокращений, списка литературы и приложений.

## **ОСНОВНОЕ СОДЕРЖАНИЕ ДИССЕРТАЦИИ**

**Во введении** обоснованы актуальность и востребованность темы диссертации, определены цель и задачи, объект и предмет исследования, соответствие исследования приоритетным направлениям развития науки и технологий

Республики Узбекистан, изложены научная новизна, практические результаты исследования, обоснована достоверность полученных результатов, раскрывается теоретическая и практическая значимость результатов исследования, приведены результаты внедрения проведенных исследований, сведения об опубликованности результатов и структуре диссертации.

В первой главе диссертации «**Модели, методы и алгоритмы речевых технологий в интеллектуальных информационных системах**» приведены речевые технологии, как главные компоненты интеллектуальных информационных систем, изучены алгоритмы процесса анализа речевых сигналов и проведен обзор методов преобразования текста в речь. На основе анализа требований к архитектурам информационных систем реабилитации речи пациентов после кохлеарной имплантации показана важность разработки подсистем анализа и синтеза речи для обеспечения речевого диалога между пользователем и машиной. Проведен широкий обзор алгоритмов анализа речевых сигналов, а также высокоуровневых TTS (eng. Text-to-Speech) систем для генерации узбекской речи.

Вторая глава диссертации «**Разработка системы синтеза узбекской речи**» посвящена нейросетевым архитектурам TTS систем, а именно архитектурам современных акустических моделей и нейронных вокодеров, разработке алгоритма синтеза узбекской речи на основе речевой базы, алгоритму формирования речевой базы узбекского языка, статистическим характеристикам сформированной речевой базы узбекского языка, а также оценке разработанной TTS системы генерации узбекской речи.

В целом технология высокоуровневого преобразования текста в речь включает следующие этапы: формирование речевой базы языка; обучение акустической модели; обучение нейросетевого вокодера для на базе акустических параметров.

Также в данной главе рассмотрены приемы создания речевой базы для TTS системы высокого уровня, предложена последовательность формирования речевой базы узбекского языка, его основные свойства и приведены основные статистические параметры сформированной многочасовой речевой базы.

Известно, что главным составляющим современных высококачественных TTS систем является речевая база с большим объемом. Речевая база TTS-систем - это база данных, состоящая из набора аудиоданных и соответствующих текстовых файлов. Аудиофайлы состоят из образцов речевых элементов (звуков, слогов, слов, предложений), а текстовые файлы содержат транскрипции, соответствующие этим речевым элементам. Сформированная речевая база TTS систем состоит из аудио-и текстовых файлов предложений узбекского языка с разной длиной.

На сегодняшний день существует ряд речевых баз для мировых языков в открытом доступе, вместе с тем нет речевой базы узбекского языка для обучения синтезатора. На практике качество синтетической речи зависит от качества речевой базы. Это особенно подтверждается для синтеза речи на основе нейросетевых архитектур. Распространенными методами формирования речевой базы для TTS систем являются:

- запись диктора, читающего заранее подготовленный текстовый, материал;
- запись диктора, произносящего спонтанную речь, нарративы и тд.

Простейшая форма создания грамотной речевой базы реализуется на базе

четырёх главных этапов: подготовка текстовой информации, озвучивание текста, синхронизация сигнала с его текстом и добавление аудио файла и его транскрипции в речевую базу.

Таблица 1.

Спецификации сформированной речевой базы

Категория	Обучение (Training)	Настройки параметров (Validation)	Общий
Объем речевой базы (в часах)	26	3	29
Выражения	11107	584	11691
Слова	143902	7574	151476
Неповторяющиеся слова	21585	1136	22721

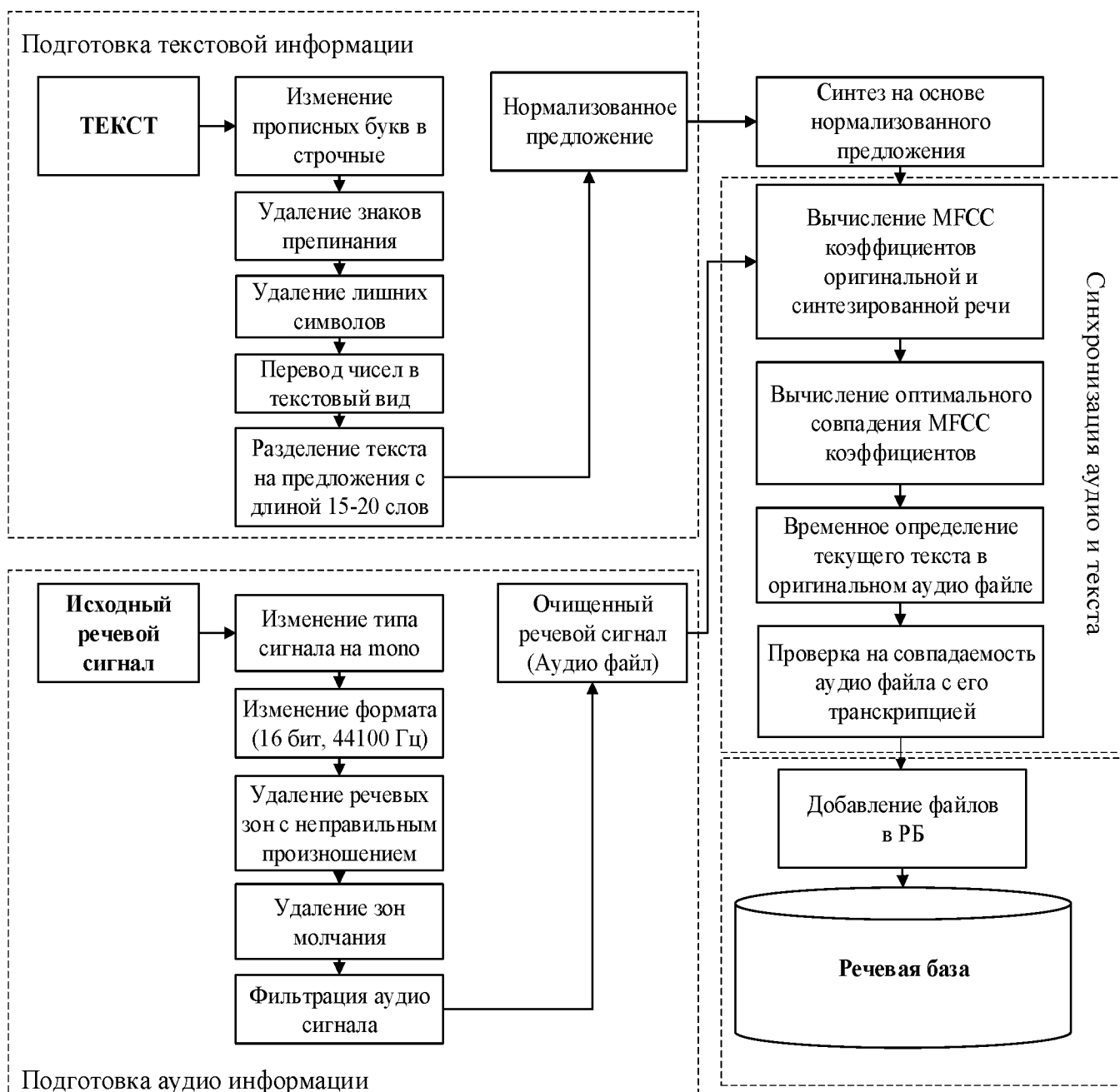


Рис.1. BPMN модель процесса формирования речевой базы для узбекского языка

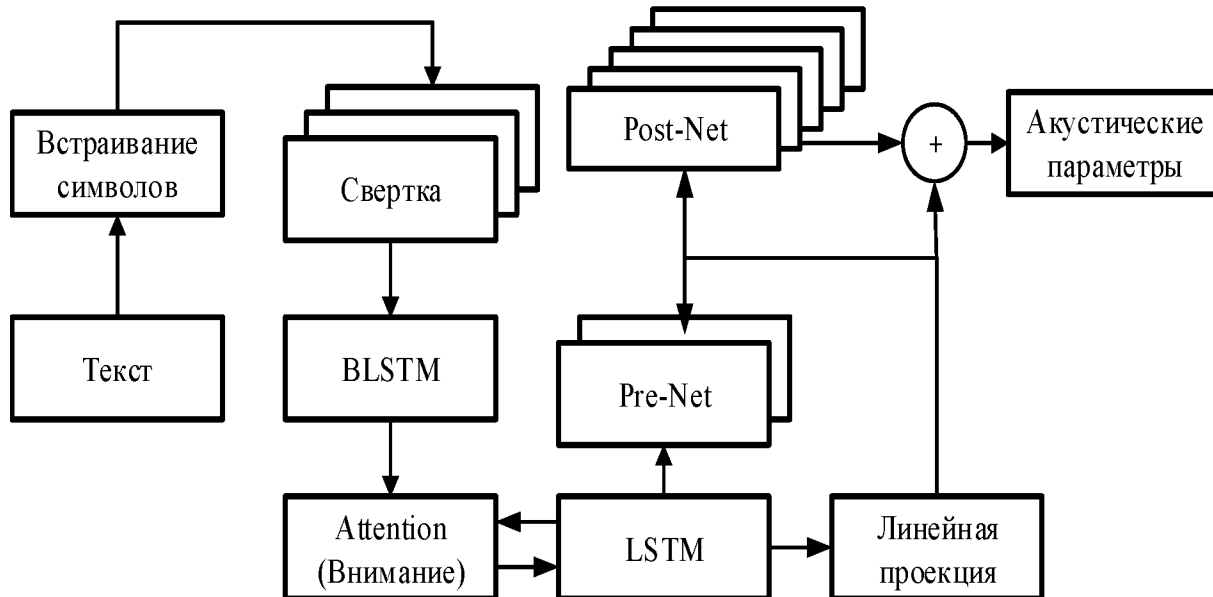
На базе предложенного метода сформирована речевая база узбекского языка. Объем речевой базы системы синтеза речи составил ~ 29 часов. Речевая база была сформирована одним диктором-женщиной. Всего в текстах, предоставленных для чтения, использовано более 11 тысяч предложений. В предложениях использовано ~ 151 500 слов узбекского языка, 22721 из которых являются неповторяющимися словами.

Развернутая структурная схема (eng. BPMN, Business Process Model and Notation модель) формирования речевой базы для узбекского языка приведена на рис.1.

Рассмотрим архитектуру высокоуровневых синтезаторов речи. TTS системы высокого уровня состоят из двух главных нейросетевых архитектур – акустической модели и вокодера.

Акустическая модель предназначена для вычисления акустических параметров, таких как форманты, частота основного тона, энергия, частотно-кепстральные коэффициенты мел-шкалы (MFCC), коэффициенты линейного предсказания (LPC), линейная спектрограмма, мел-спектрограмма и другие. Популярными современными акустическими моделями являются DeepVoice, Clarinet, Tacotron и FastSpeech.

Tacotron2 – это акустическая модель, способная генерировать речь на основе речевой базы объемом более 15 часов и представляет из себя две сети: сеть предсказания признаков и вокодер. Сеть предсказания признаков рассматривается как основная сеть, а вокодер WaveNet выступает в роли системы генерации речи на основе акустических параметров. Структура предложенной акустической модели на основе Tacotron2 приведена на рис.2, а её параметры в таблице 2.



**Рис.2. Структура разработанной акустической модели на основе Tacotron2**

Сгенерированные параметры акустической модели передаются следующему компоненту - вокодеру. Вокодер – система, предназначенная для генерации формы

волны на основе акустических параметров, полученных из акустического модуля TTS системы. Применяемые вокодеры представляют из себя глубокую нейронную сеть.

Среди нейронных вокодеров параллельной генерации волн особое место занимают модели на основе GAN (eng. Generative adversarial network) по причине присутствия самообучаемости архитектуры, а также из-за наличия следующих преимуществ по отношению к остальным:

-скорость: скорость обучения и вывода данных в 4,82 (1,96) раза выше, чем у обычного параллельного Wavenet (т.е. ClariNet);

-высокое качество: модель достигает 4,16 по шкале MOS (eng. Mean Opinion Score), что конкурентоспособно по сравнению с лучшей ClariNet на основе фильтрации. По этой причине, в данном научном исследовании для синтеза узбекской речи выбран вокодер на основе модели GAN.

**Таблица 2.**

**Слои и параметры нейронной сети акустической модели**

№	Название слоев	Параметры слоев	Кол-во слоев
1	Входной слой	Нормализованный текст в транскрипционной форме (26x1)	1
2	Встраивание символов	Размер слоя = 512 (x5) Функция активации = ReLU Dropout=0,5	1
3	Сверточный слой	Размер слоя= 512 Функция активации = ReLU Dropout=0,5	3
4	BLSTM	Размер слоя= 512 Активация = ReLU Dropout=0,5	1
5	Attention	Количество ячеек памяти = 512 Batch Size=32	1
6	LSTM	Количество ячеек памяти = 1024 Zoneout=0,1	2
7	PreNet	Количество ячеек памяти = 256	2
8	PostNet	Количество ячеек памяти = 512 Размер фильтра =5 Функция активации = tanh	5
9	Параметры обучения	Алгоритм оптимизации: Длина шага обучения Адама =0.001, Epoch=160	

Структура предлагаемого нейросетевого вокодера на базе Parallel Wave GAN (PWG) приведена на рис.3.

Общая разработанная архитектура синтеза узбекской речи на основе вышеизложенных стандартных пакетных средств описана с помощью диаграммы IDEF0 и приведена на рис.4. Разработанная TTS система узбекского языка по шкале MOS составило 4,36 баллов из 5.

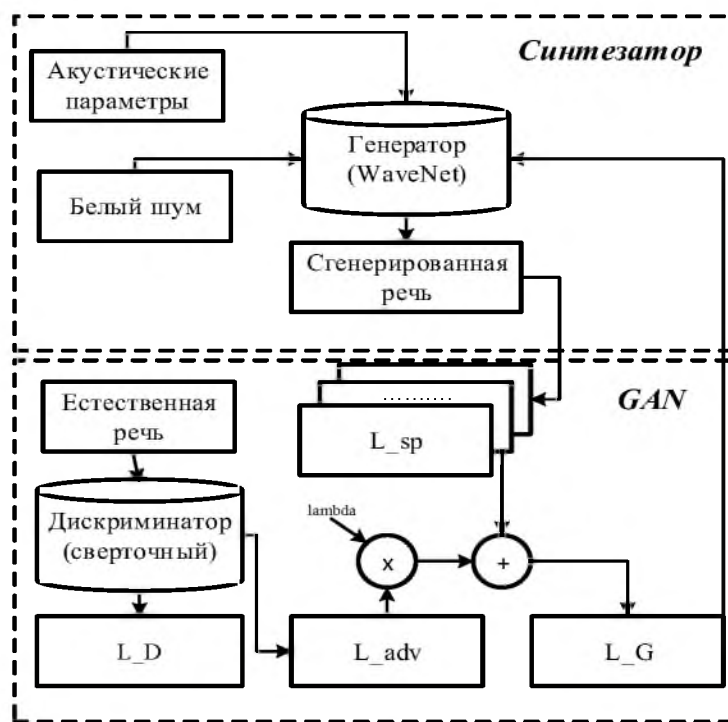


Рис.3. Структура разработанного нейросетевого вокодера на базе PWG

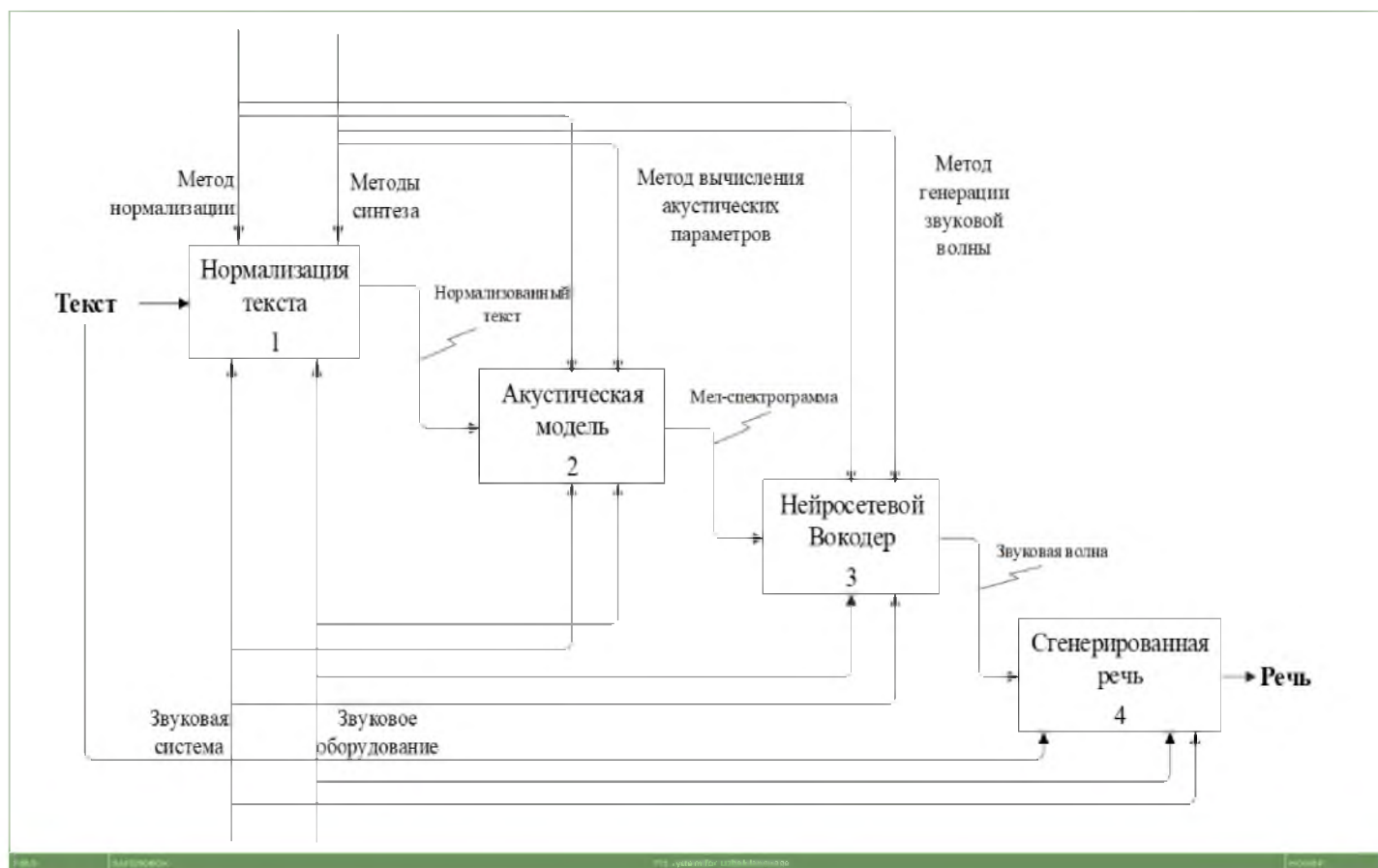


Рис.4. IDEF0 диаграмма системы синтеза узбекской речи

Третья глава диссертации «Создание компонентов диалогового интерфейса процесса реабилитации» содержит пять параграфов. В них представлены разработанные алгоритмы анализа узбекской речи на основе мел-спектрограммы,

описание процесса обучения и тестирования предлагаемой модели, разработанные алгоритмы анализа узбекских звуков на основе формант и сравнения речевых сигналов, а также разработанные алгоритмы аутентификации диктора.

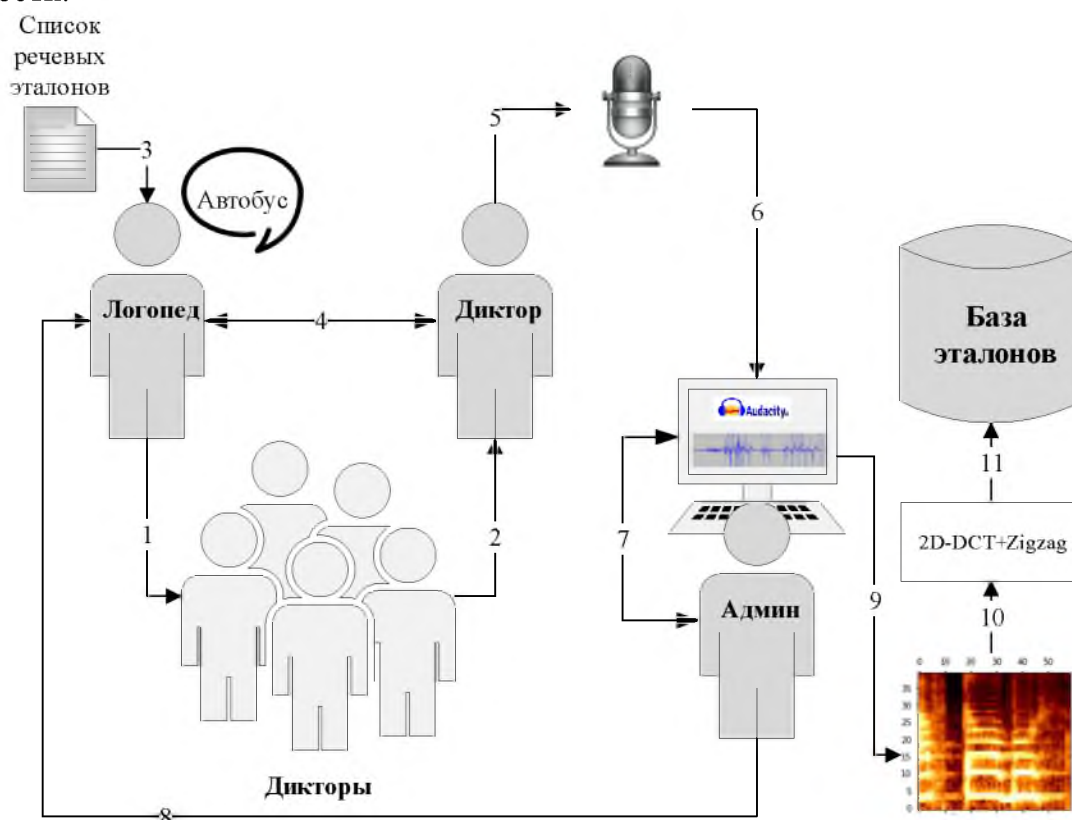
Важнейшим компонентом процесса реабилитации речи является диалоговый интерфейс, состоящий из процедур синтеза и анализа речи. Процесс анализа речевых сигналов состоит из предобработки, построения спектрограммы, вычисления признаков и классификации речи. В системах анализа речи наиболее важным является этап параметризации речевого сигнала. Параметрические признаки должны отвечать следующим требованиям: информативность, минимизация объема данных, дикторнезависимость, единообразие.

В задачах анализа речи в качестве информативных признаков использованы спектрограммы Фурье и мел-спектрограммы. В отличие от Фурье-спектра мел-спектрограмма представляет собой частоту, выраженную в мел-шкалах, а не в Гц. Переход к мел-спектрограмме осуществляется путем применения мел-фильтров к входящей спектрограмме. Таким образом, мел-спектрограмма выделяет из сигнала важные особенности, позволяет увеличить степень его сжатия.

Рассмотрим отдельные компоненты системы реабилитации.

### Формирование речевых эталонов детей.

В рамках диссертационной работы была создана база эталонов детских голосов дошкольного возраста на основе ограниченного словаря. Словарь состоит из трех категорий речевых единиц узбекского языка: звуки, слоги, слова трех уровней сложности.



**Рис.5. Цикл процесса формирования базы эталонов детей**

Каждая речевая единица записана через микрофон в специальной, незашумлённой среде. На рис.5 приведен подробный цикл процесса записи каждого речевого эталона, согласно которому:



1-2 шага. Логопед осуществляет отбор кандидатов среди детей (дикторов) на основе проверки речевых навыков ребенка.

3-4 шага. Логопед произносит речевую единицу (звук, слог, слова). Следом диктор повторяет текущую речевую единицу за логопедом.

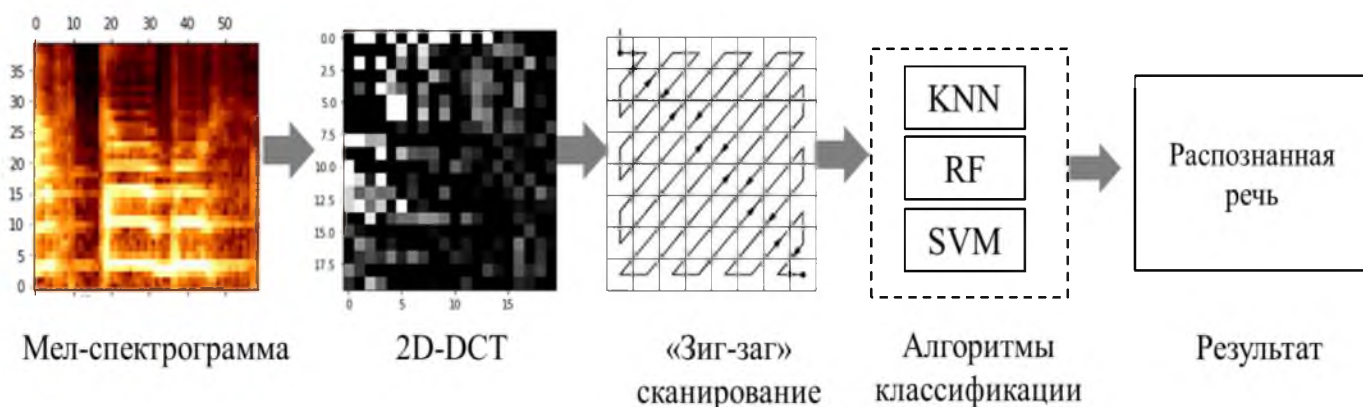
5-7 шага. Администратор записывает речь диктора через микрофон и сохраняет.

8-11 шага. Администратор преобразует записанные аудио эталоны в необходимый формат и сохраняет в речевой базе.

#### Подсистема процесса анализа слогов и слов узбекского языка.

В предложенном алгоритме анализа слов, мел-спектрограммы подвергаются специальному 2D-Дискретному косинусному преобразованию (2D-DCT), извлекающему из изображений информативные признаки, которые далее сортируются алгоритмом «зиг-заг» сканирования и классифицируются алгоритмами машинного обучения.

В результате последовательной обработки получена оценка средней точности по KNN (eng.k-nearest neighbour), RF (eng. Random Forest), SVM (eng. Support Vector Machine) алгоритмам классификации машинного обучения (Рис.6.)



**Рис.6. Последовательность процессов предлагаемого алгоритма анализа речи на базе мел-спектрограммы**

Точность анализа напрямую зависит от объема и качества обучающей выборки алгоритмов машинного обучения. Для классификации речевого сигнала, согласно предлагаемому алгоритму, обучающая выборка должна состоять из вектора признаков эталонов, полученных на основе применения разработанного алгоритма на основе 2D-DCT с «Зиг-заг» сканированием к каждому исходному аудио файлу. Эталон – набор признаков аудио файла, являющийся главным образцом правильного произношения звука (слога, слова).

**Таблица 3.**

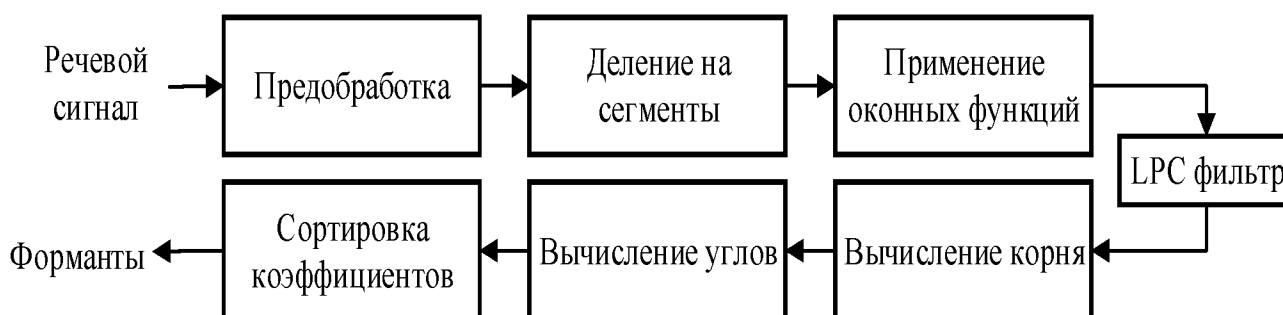
**Усредненные результаты тестирования предлагаемой модели**

	KNN	RF	SVM	Средняя точность классификации
Звуки	95,6%	92,3%	93,5%	93,8%
Слоги	94,3%	94,8%	92,5%	93,9%
Слова	95,7%	93,4%	93,8%	94,3%
Средняя точность классификации по моделям	95,2%	93,5%	93,3%	94%

Согласно результатам тестирования (таблица 3) наиболее точная классификация выбранных узбекских звуков, слогов и слов достигается при применении KNN, его средняя точность классификации составляет 95,2%, превосходя RF (93,5%) и SVM (93,3%).

#### Подсистема анализа гласных звуков.

В диссертационной работе также предложен алгоритм анализа узбекских звуков с помощью формантного набора. Форманта характеризуется частотой, шириной и амплитудой. Формантой является некоторый амплитудный всплеск на графике спектра, а его частота – частота пика этого всплеска. Описание предлагаемого алгоритма вычисления формант на основе метода линейного предсказания приведено на рис.7.



**Рис.7. Процесс вычисления форманты**

Обычно рассчитывают и обозначают от трех до шести формант соответственно F1, F2, ... ,F6. Несмотря на это, гласные звуки различаются с достаточно высокой точностью по первым трем формантам {F1, F2, F3}. Первые две форманты {F1, F2} являются более информативными и имеют тесную взаимосвязь, а F3 несет информацию о интонации. Классификации полученных формантных признаков осуществляется на базе матрицы ошибок алгоритмами KNN (95%), RF (94%), SVM (95%).

#### Подсистема сравнения речевых сигналов.

В системах анализа речи задача сравнения сигналов и оценка их схожести в процентах является одной из важных. Существует ряд алгоритмов сравнения речевых сигналов, т.к. алгоритм динамической трансформации временной шкалы (DTW-алгоритм), корреляция, дискретное расстояние, Эвклидово расстояние. Эти алгоритмы имеют свои недостатки, например их результаты не могут быть интерпретированы в процентной или балловой шкале.

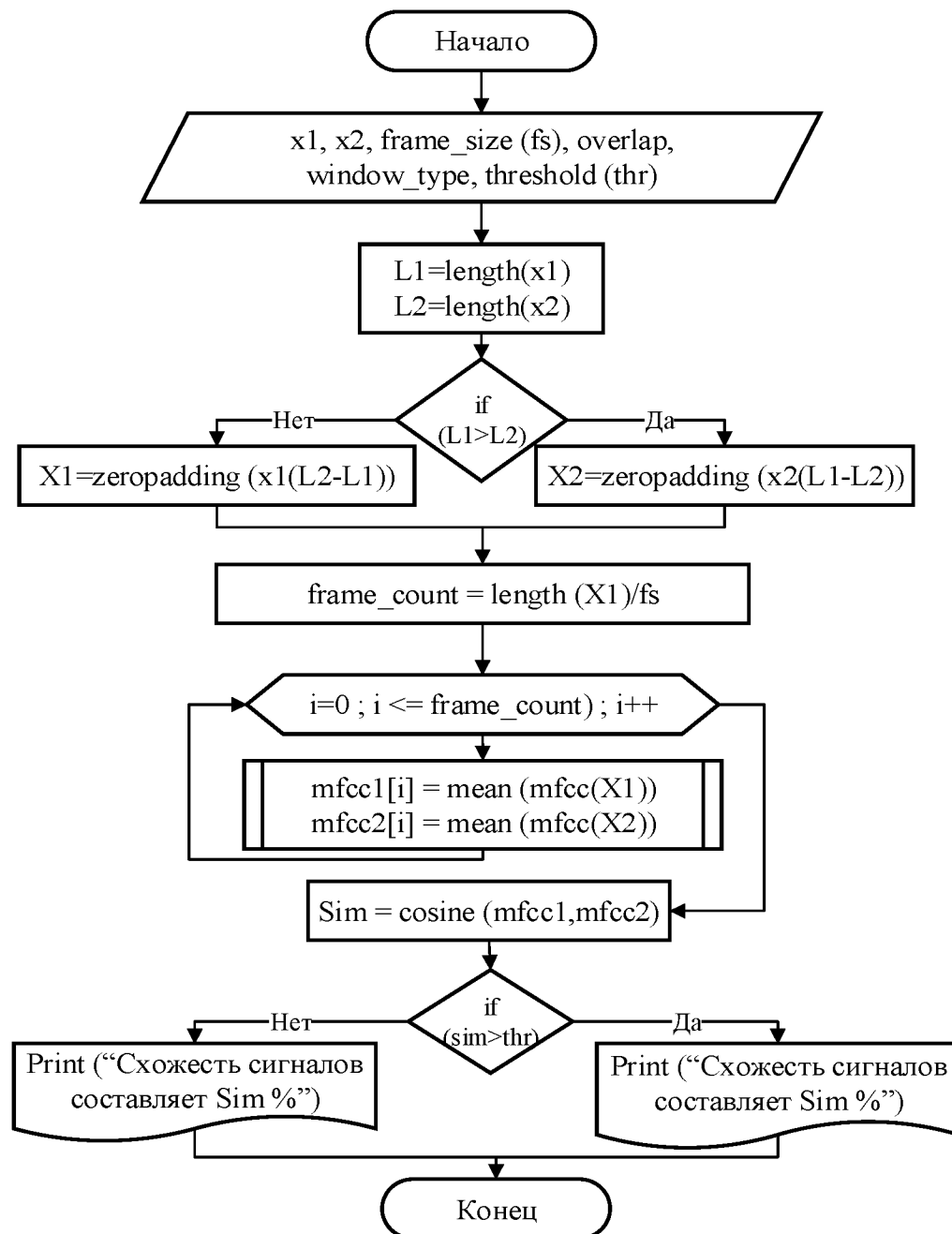
Известно, что сравнение двух речевых сигналов осуществляется через сравнение информативных признаков, таковыми могут являться мел-коэффициенты, коэффициенты метода линейного предсказания, форманты, кепстральные коэффициенты, вычисленные из речевых сигналов.

На рис.8 приведена блок-схема предлагаемого алгоритма сравнения двух сигналов на основе косинусного подобия MFCC коэффициентов.

Основная идея подобия косинусов — это вычисление косинуса угла между двумя векторами MFCC коэффициентов по следующей формуле:

$$\text{CosineSimilarity} = S_C(A, B) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$$

где  $A$  содержит MFCC коэффициенты входной речевой фонограммы и  $B$  представляет MFCC коэффициенты речевых эталонов.



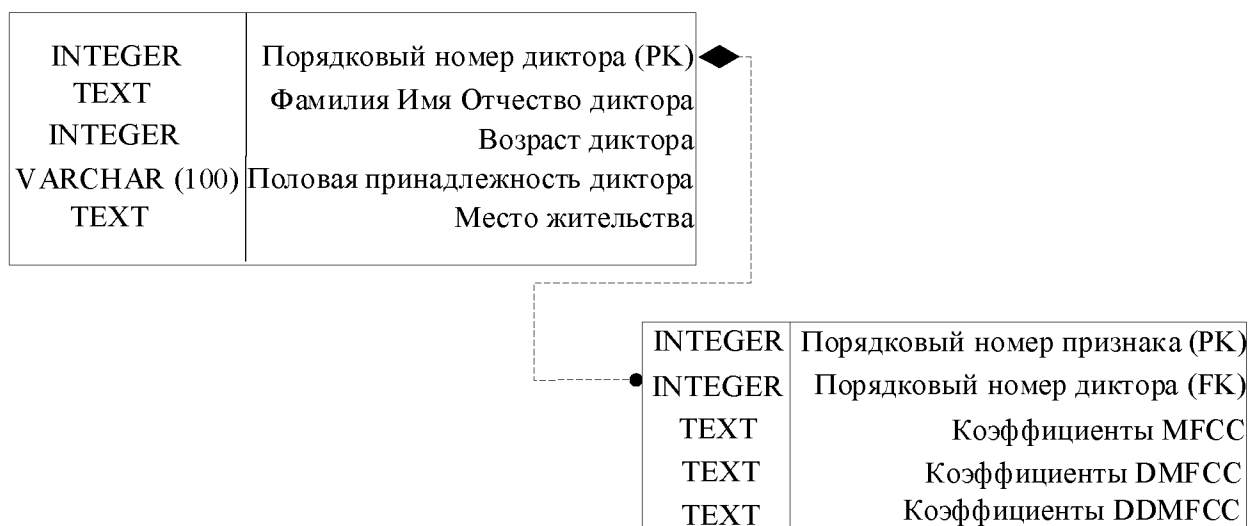
**Рис.8. Блок-схема алгоритма сравнения речевых сигналов**

#### **Подсистема аутентификации.**

Для создания и ведения баз данных пациентов была разработана подсистема аутентификации дикторов (пациентов). Был предложен алгоритм, состоящий из трех относительно независимых частей:

1. Выделение информативных признаков (параметризация входного речевого сигнала по описанному алгоритму).
2. Построение эталона для определенного диктора в виде вектора признаков.
3. Принятие решения на основе сравнения с эталонами на базе GMM (eng. Gaussian Mixture Model).

На рис.9 приведена схема физического соединения таблиц базы данных эталонов на основе IDEF1x. Методологическая схема IDEF1x является усовершенствованной версией модели IDEF1. IDEF1x используется для моделирования реляционных баз данных.



**Рис.9. Методологическая схема IDEF1x физического соединения таблиц БД**

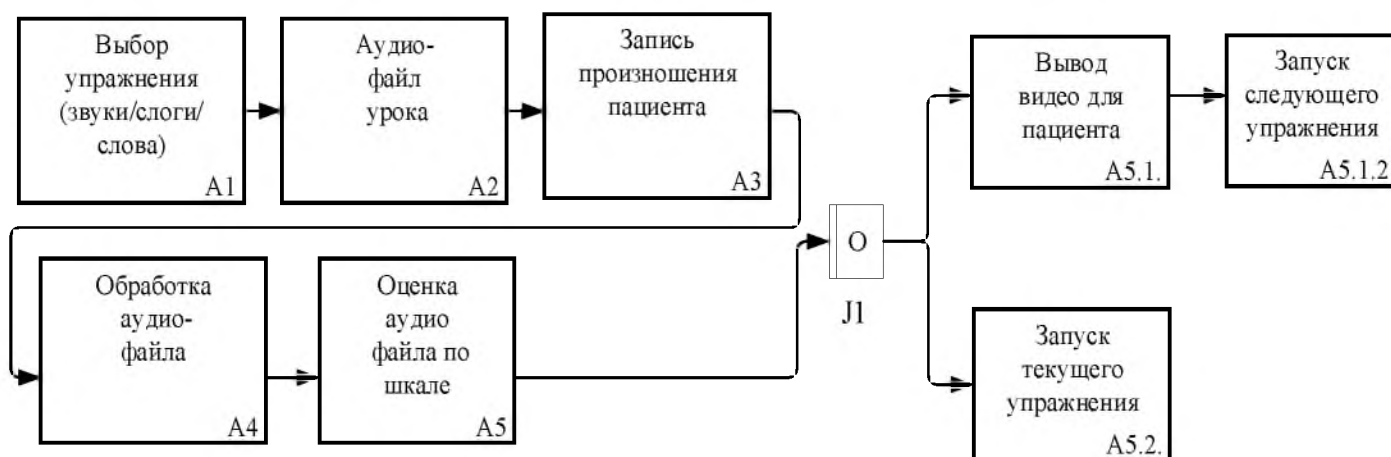
Точность аутентификации диктора на основе предложенного алгоритма составило 96,4%.

В четвертой главе диссертации «Создание информационной системы реабилитации пациентов с нарушением слуха» приведены технология реабилитации пациентов, структура разработанной информационной системы, интерфейс разработанной информационной системы реабилитации речи и показатели эффективности её практического применения.

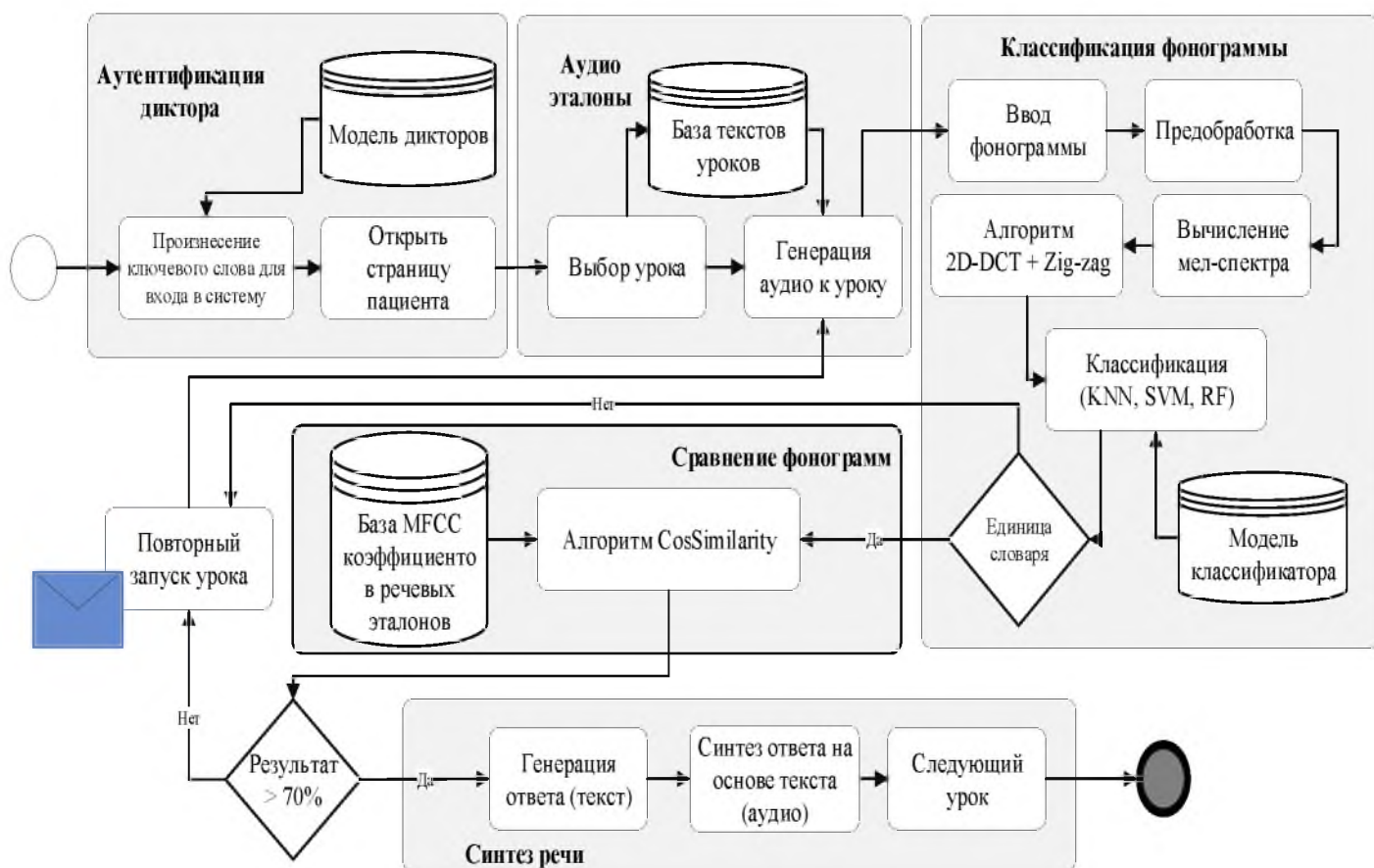
Информационная система предназначена для реабилитации пациентов после кохлеарной имплантации. Кохлеарная имплантация – это система мероприятий по реабилитации людей с нарушением слуха, включающая три этапа:

1. Отбор кандидатов для проведения операции;
2. Оперативное вмешательство (проведение хирургической операции);
3. Послеоперационная слухоречевая реабилитация.

Последовательность этапов процесса выполнения упражнений каждого блока пациентом на базе методологии моделирования IDEF3 (рис.10), на котором показаны причинно-следственные связи между действиями с использованием структурного метода. На рис.11 приведен развернутый процесс реабилитации речи пациентов в рамках информационной системы на базе BPMN моделирования.



**Рис.10. Функциональная IDEF3 модель процесса выполнения упражнений реабилитации речи пациентов**



**Рис.11. BPMN диаграмма информационной системы реабилитации**

Самая маленькая единица речи - звуки. Сурдопедагоги выделяют 17 главных звуков узбекского языка, которые группируются в два концентра (таблица 4). На базе данных концентров ведется процесс реабилитации.

Таблица 4

## Классификация узбекских звуков по концентрам

1-группа концентр	A	O	U	E	I	O'	P	T	K	F	S	Sh	V	L	R	M	N
2-группа концентр				Y		B	D	Q G G'		Z	X H	Ch J					NG

Известно, что звуки латинского алфавита узбекского языка, такие как «Q», «G», «G'» являются наиболее сложно произносимыми звуками.

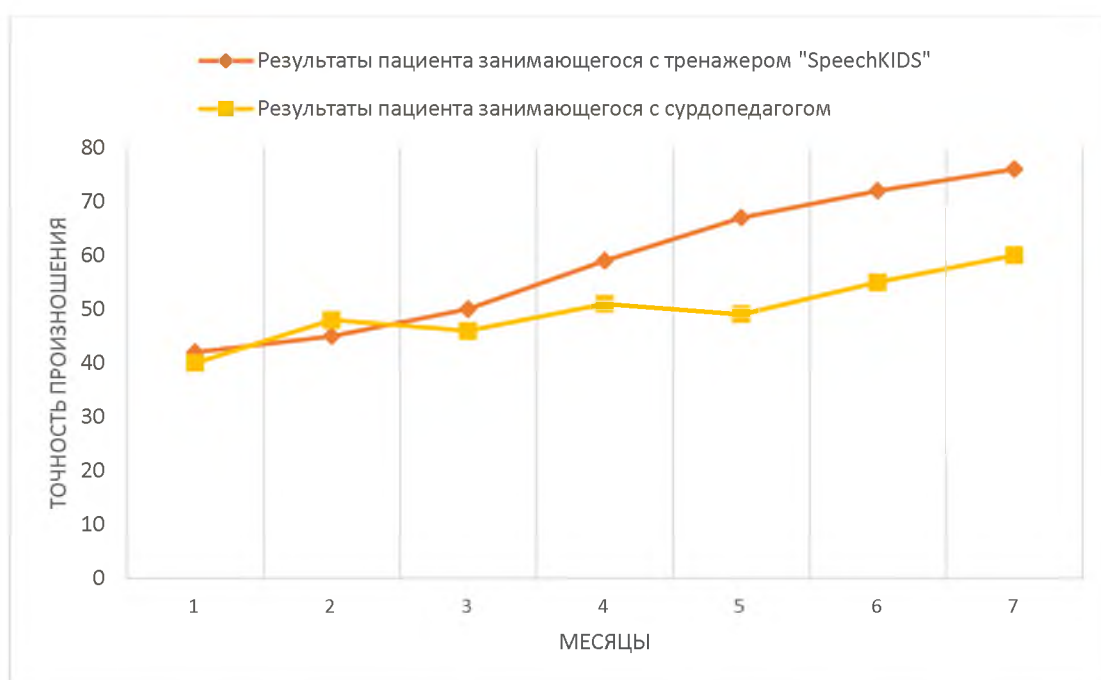


Рис.12. Результаты произношения звука «Q»

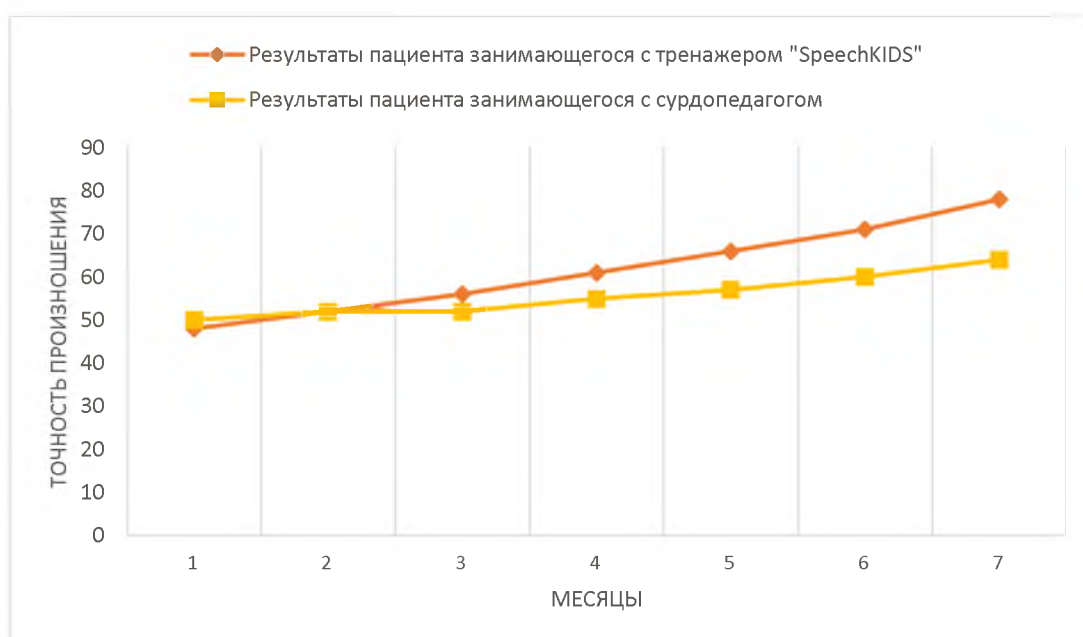
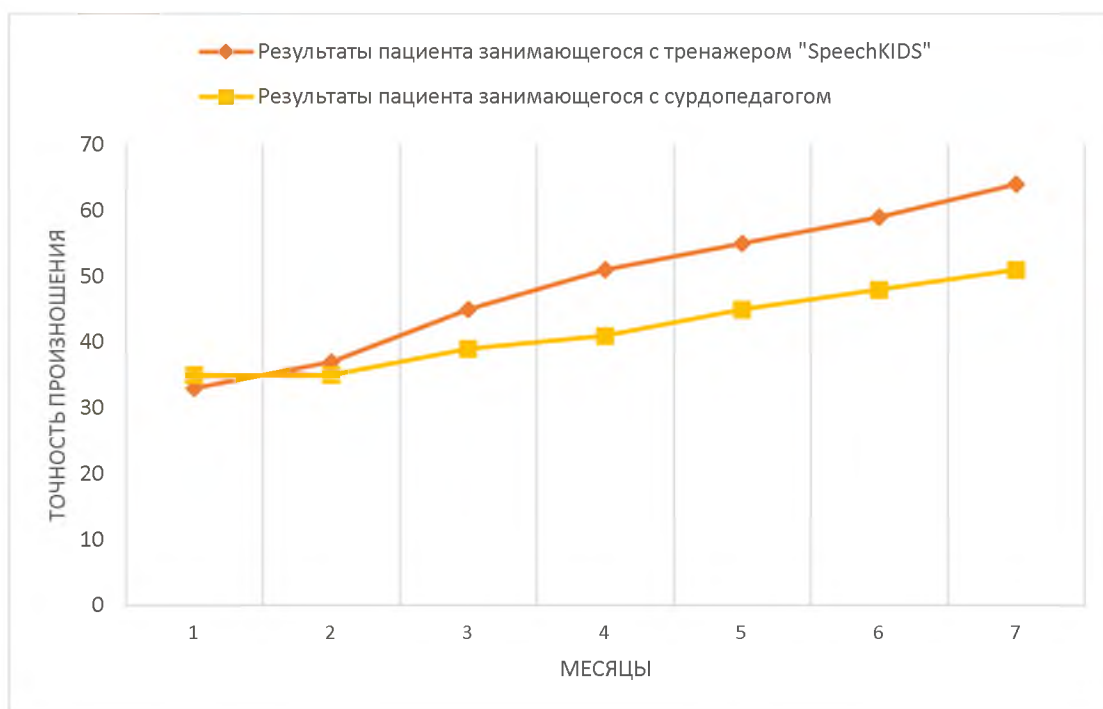


Рис.13. Результаты произношения звука «G»



**Рис.14. Результаты произношения звука «G»**

Для оценки информационной системы были выбраны дети послекохлеарной операции. Реабилитация детей проводилась на протяжении 7 месяцев. При этом были выбраны два действующих метода реабилитации: традиционный с сурдопедагогом; информационная системы (тренажер).

На рис.12-14 представлены результаты сравнения точности произношения по предлагаемой и ныне существующим технологиям реабилитации. Согласно диаграммам видно, что разработанная информационная система (SpeechKids) в среднем ускоряет процесс реабилитации в 1,5-2 раза.

## ЗАКЛЮЧЕНИЕ

На основе исследований по диссертационной работе на тему «Разработка информационной системы реабилитации пациентов с нарушением слуха на базе речевых технологий» можно сформулировать следующие выводы:

1. Проведен анализ моделей, методов и алгоритмов обработки речи, создана архитектура информационной системы реабилитации пациентов с нарушениями слуха на базе интеллектуальных алгоритмов анализа и синтеза узбекской речи.

2. Для системы синтеза речи создана речевая база узбекского языка на основе ВРММ-модели и алгоритмов синхронизации аудиосигнала с текстом и его транскрипции. Объем речевой базы составляет 11 тысяч предложений в объеме 29 часов чтения диктором.

3. Разработана акустическая модель для определения параметров речи на базе алгоритмов спектрального анализа, вычисления кепстральных коэффициентов. Акустическая модель построена на базе сверточных и рекуррентных нейронных сетей глубокого обучения.

4. Разработана TTS-система синтеза узбекской речи на основе нейронных архитектур Takotron2 и Parallel WaveGAN. Данная система синтеза обучена на многочасовой речевой базе узбекского языка и позволяет преобразовать текст в речь с показателем 4,36 из 5 по шкале оценивания качества MOS.

5. Для создания эффективного диалогового интерфейса процесса реабилитации детей-пациентов разработана подсистема анализа звуков, слогов и слов узбекского языка. В разработанной подсистеме анализа произносимых компонентов речи реализуется с использованием алгоритмов 2D-дискретного косинусного преобразования (извлечение признаков) и «зиг-заг» сканирования. Полученные файлы преобразования подвергались последовательным процедурам классификации KNN, RF SV и оценивались по критерию правильного произношения в сравнении с базой ранее подготовленных эталонов. Тестирование показало среднюю точность алгоритмов анализа по звукам, слогам и словам 94%.

6. Для формирования и сопровождения баз данных пациентов в процессе массовой реабилитации был разработан программный модуль аутентификации дикторов (пациентов).

Общий алгоритм содержит три независимых части:

- параметризация произносимого пароля;
- построение образца в виде вектора признаков;
- сравнение входной речи с эталоном на основе гауссовых смесей.

7. На основе технологии реабилитации создана в виде тренажера информационная система реабилитации детей после кохлеарной имплантации. Последовательность этапов процесса выполнения упражнений реализованы по методологии моделирования IDEF3. Разработаны пользовательские интерфейсы, в том числе графические, для режимов взаимодействия пациентов и педагогов в режиме мультимедиа. Для этого создана специальная мультимедийная база цветных изображений из детских рисунков.

8. В работе представлены показатели эффективности практического применения созданной информационной системы реабилитации. Система в целом состоит из реляционной базы данных, речевого интерфейса, прикладных программ анализа и синтеза речи. Тестирование и оценка эффективности процесса реабилитации в Республиканском специализированном научно-практическом медицинском центре педиатрии показало, что сроки реабилитации пациентов - детей дошкольного возраста - могут быть сокращены в 1,5 - 2 раза.



**SCIENTIFIC COUNCIL AWARDING SCIENTIFIC DEGREES  
DSc.13/30.12.2019.T.07.01 AT TASHKENT UNIVERSITY OF  
INFORMATION TECHNOLOGIES**

---

**TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES**

**ABDULLAEVA MALIKA ILKHAMOVNA**

**DEVELOPMENT OF INFORMATION SYSTEM FOR REHABILITATION  
OF PATIENTS WITH HEARING LOSS BASED ON SPEECH  
TECHNOLOGIES**

05.01.10 – Information systems and processes

**DISSERTATION ABSTRACT OF THE DOCTOR OF PHILOSOPHY (PhD)  
ON TECHNICAL SCIENCES**

**Tashkent – 2023**

The theme of doctor of philosophy (PhD) on technical sciences was registered at the Supreme Attestation Commission at the Cabinet of Ministers of the Republic of Uzbekistan under number B2023.2.PhD/T3546

The dissertation has been prepared at Tashkent University of Information Technologies.  
The abstract of the dissertation is posted in three languages (Uzbek, Russian, English (resume)) on the website ([www.tuit.uz](http://www.tuit.uz)) and on the website of «Ziyonet» Information and educational portal ([www.ziyonet.uz](http://www.ziyonet.uz).)

**Scientific adviser:** Musaev Mukhammadjon Makhmudovich  
Doctor of technical sciences, professor

**Official opponents:** Saidov Abdusobir Abdurahmonovich  
Doctor of technical sciences, professor

Elov Jamshid Bekmurodovich  
PhD in technical sciences, docent

**Leading organization:** Samarkand state university named after Sharof Rashidov

The defense of dissertation will take place « 23 » June 2023 at 14<sup>00</sup> at the meeting of Scientific Council No. DSc.13/30.12.2019.T.07.01 at Tashkent University of Information Technologies (Address: 100084, Tashkent city, Amir Temur street, 108. Ph.: (+99871) 238-64-43, e-mail: [iktuit@tuit.uz](mailto:iktuit@tuit.uz)).

The dissertation can be reviewed at the Information Resource Centre of Tashkent University of Information Technologies (is registered under No. 278 ). (Address: 100084, Tashkent city, Amir Temur street, 108. Ph.: (+99871) 238-64-70, e-mail: [iktuit@tuit.uz](mailto:iktuit@tuit.uz)).

Abstract of the dissertation sent out on « 10 » June 2023 y.  
(mailing report No. 17 on « 06 » June 2023 y.).



**M.A.Rakhmatullaev**  
Deputy Chairman of the scientific council  
awarding scientific degrees,  
Doctor of Technical Sciences, Professor

**N.O.Rakhimov**  
Scientific secretary of scientific council  
awarding scientific degrees,  
Doctor of Technical Sciences, Docent

**U.R. Xamdamov**  
Chairman of the academic  
seminar under the scientific council  
awarding scientific degrees,  
Doctor of Technical Sciences, Docent

## INTRODUCTION (abstract of PhD dissertation)

**The aim of the research work** is to develop the information system for rehabilitation patients with hearing impairment based on algorithms analyzing and synthesizing Uzbek speech using deep neural networks.

**The object of the research work** is the information system for the rehabilitation process of hearing and speech skills of children after cochlear implantation..

**The scientific novelty of the research work** is as follows:

software components of information system for rehabilitation patients with hearing impairment based on MVT technology were developed;

a speech interface based on speech analysis and synthesis models was developed, an IDEF0 diagram of the developed TTS system of Uzbek language based on deep neural networks was constructed;

algorithms for the process of synthesizing Uzbek speech based on the speech database, and created a database of reference speech units in a specialized vocabulary for patient rehabilitation were developed;

algorithm for the analysis process of phonemes, syllables and words of the process of rehabilitation of the patient's auditory and verbal skills were developed.

**Implementation of the research results.** The developed information system based on methods, algorithms and software for speech analysis and synthesis using artificial intelligence algorithms was implemented:

in the Republican Specialized Scientific and Practical Medical Center of Pediatrics (RSNPMC Pediatrics), in the Department of congenital and acquired diseases of ENT organs for hearing and speech rehabilitation of patients after cochlear implantation (act of implementation dated April 17, 2023). The use of the simulator allowed accelerating the process of rehabilitation of children by 1,5-2 times in comparison with the traditional method of rehabilitation under the guidance of a surdopedagogue. The effect of the scientific work was achieved through the use of the developed speech technologies in the rehabilitation process;

in "Nova research and resolution" Ltd. for elimination of speech defects in children with speech disorders (act of introduction dated April 27, 2023). As a result, the developed information system based on the proposed algorithms made it possible to reduce the number of lessons with a speech therapist, thus accelerating the process of rehabilitation of children with speech defects. The main advantage of the information system is the possibility of online rehabilitation.

**The structure and volume of the dissertation.** The structure of the dissertation consists of an introduction, four chapters, conclusions, references and appendices. The volume of the dissertation is 116 pages.

**ЭЪЛОН ҚИЛИНГАН ИШЛАР РЎЙХАТИ**  
**СПИСОК ОПУБЛИКОВАННЫХ РАБОТ**  
**LIST OF PUBLISHED WORKS**

**I бўлим (1 часть; part 1)**

1. Абдуллаева М.И., Очилов М.М., Ибрагимова С.Н. Усовершенствованный метод извлечение признаков для идентификации диктора с использованием алгоритма классификации // ТАТУ хабарлари. №2(58)/2021. -Б. 41-53. (05.00.00; №31).

2. Musayev M., Abdullaeva M., Khujayorov I., Ochilov M. Recognition of uzbek words by processing the spectrogram image // International scientific and technical journal “CHEMICAL TECHNOLOGY. CONTROL AND MANAGEMENT”. № 1(103)/2022. -P.84-90 (05.00.00, №12).

3. Abdullaeva M. Definition of interphoneme transitions of uzbek words by discrete wavelet transforms // ACTA OF TURIN POLYTECHNIC UNIVERSITY IN TASHKENT. EDITION 2/2022. -P.59-64 (05.00.00, №25).

4. Модель оценки частоты основного тона с использованием вейвлет преобразования для узбекской речи // «ИННОВАЦИОН ТЕХНОЛОГИЯЛАР» илмий-техник журнали. 2021/4(44)-сон. -С.27-32. (05.00.00, №38).

5. M. Abdullaeva, I. Khujayorov and M. Ochilov, “Formant set as a main parameter for recognizing vowels of the Uzbek language,” 2021 International Conference on Information Science and Communications Technologies (ICISCT), 2021, pp. 1-5. Раёсат қарори №525 (30.10.2021 й.) билан ОАК илмий нашрлар рўйхатига киритилган журналларга тенглаштирилган (Scopus).

6. Abdullaeva, D.B.Juraev, M.M.Ochilov, M.F.Rakhimov. Uzbek Speech Synthesis Using Deep Learning Algorithms // 14th International Conference on Intelligent Human Computer Interaction (IHCI-2022). LNCS 13741, 2023. -12p. Tashkent-2022. (11; Springer; IF = 1.49).

**II бўлим (2 часть; part 2)**

7. Musaev M., Abdullaeva M., Turaev B. Image Approach to Uzbek Speech Recognition // Proceedings of 2022 IEEE 22<sup>nd</sup> International Conference on Communication Technology (22<sup>nd</sup> IEEE ICCT). Nanjing, China-2022. -P.1201-1206 (Scopus).

8. Musaev M., Abdullaeva M., Ochilov M. Advanced Feature Extraction Method for Speaker Identification Using a Classification Algorithm // II International Scientific Forum on Computer and Energy Sciences (WFCES-II 2021). AIP Conference Proceedings 2656, 020022 (2022). -7p. (Scopus).

9. Абдуллаева М. Сегментация речевого сигнала на фонемные единицы // Scientific collection “Interconf”. Proceedings of the 2nd International Scientific and Practical Conference Global and regional aspects of sustainable development. №43, February, 2021. Copenhagen, Denmark-2021. -P.653-660.

10. Абдуллаева М. Технология эффективной реабилитации детей послекохлеарной операции с помощью алгоритмов речевого распознавания // Scientific collection "Interconf". Proceedings of the 7th International Scientific and Practical Conference International forum: problems and scientific solutions. №53, April, 2021. Melbourne, Australia-2021. -P.712-717.

11. Абдуллаева М. Эффективная детекция границ речевого сигнала с помощью вейвлет преобразования // Scientific collection "Interconf". Proceedings of the 5th International Scientific and Practical Conference Scientific community: interdisciplinary research. №71, August, 2021. Hamburg, Germany-2021. -P.443-449.

12. Абдуллаева М., Тураев Б. Применение алгоритмов искусственного интеллекта в распознавании узбекской речи // Scientific collection "Interconf". Proceedings of the 10th International Scientific and Practical Conference, Scientific research in XXI century. №85, November, 2021. Ottawa, Canada-2021. -P.425-429.

13. Абдуллаева М., Пулатова С. Application of modern speech recognition technologies by creating a linguistic simulator for improving language competence // Scientific collection "Interconf". Proceedings of the 1st International Scientific and Practical Conference Scientific paradigm in the context of technologies and society development. №86, November, 2021. Geneva, Switzerland-2021. -P.584-589.

14. Абдуллаева М. Управление системой с помощью человеческой речи // "Zamonaviy axborot texnologiyalari va AT-ta'lim tatbiqu muammolari" mavzusidagi Respublika ilmiy-amaliy anjumani ma'ruzalar to'plami. II tom. Samarqand-2021. -B.120-122.

15. Абдуллаева М. Возможности вейвлет преобразования в распознавании человеческой речи // Materials of the international online distance conference on "modern informatics and its teaching methods" (MITM 2020). part-2. Andijan-2020. -p.263-265.

16. Abdullayeva M., Po'latova S., Zamonaviy intellektual interfeys // zamonaviy axborot texnologiyalari va AT-ta'lim tatbiqu muammolari" mavzusidagi respublika ilmiy-amaliy anjumani ma'ruzalar to'plami. II tom. Samarqand-2021. -B.125-127.

17. Musaev M.M., Abdullayeva M.I., Ochilov M.M., Berdanov U.A. O'zbek nutqli diktorni tanish // O'zbekiston respublikasi adliya vazirligi huzuridagi intellektual mulk agentligi. elektron hisoblash mashinalari uchun yaratilgan dasturning rasmiy ro'yxatdan o'tkazilganligi to'g'risidagi guvohnoma № DGU 13024. 14.11.2021.

18. Musaev M.M., Turayev B.Sh. Maktabgacha bo'lgan bolalarning nutqini tanib olish uchun dastur // O'zbekiston respublikasi adliya vazirligi huzuridagi intellektual mulk agentligi. elektron hisoblash mashinalari uchun yaratilgan dasturning rasmiy ro'yxatdan o'tkazilganligi to'g'risidagi guvohnoma № DGU 13812. 21.12.2021.

19. Musaev M.M., Ochilov M.M., Raximov M.F., Jurayev D.B. O‘zbek tili nutqini sintezlovchi “Matn-nutq” dasturi // O‘zbekiston respublikasi adliya vazirligi elektron hisoblash mashinalari uchun yaratilgan dasturning rasmiy ro‘yxatdan o‘tkazilganligi to‘g‘risidagi guvohnoma № DGU 17273. 01.07.2022.

20. Inoyatova F.I., Abdukayumov A.A., Musayev M.M., Najmitdinova N.Sh., Ochilov M.M., Aliyeva M.U., Abdullayeva M.I., To‘rayev B.Sh. Bolalar nutqi to‘g‘ri tallaffuzini baholash dasturi // O‘zbekiston respublikasi adliya vazirligi elektron hisoblash mashinalari uchun yaratilgan dasturning rasmiy ro‘yxatdan o‘tkazilganligi to‘g‘risidagi guvohnoma № DGU 24229. 25.03.2023.

«Муҳаммад ал-Хоразмий авлодлари» илмий журнали таҳририятида таҳрирдан ўтказилди ва ўзбек, рус ва инглиз тилларидаги матнларини мослиги текширилди.