

УДК 004.934.5

**Я. С. Зяноўка**

Аб'яднаны інстытут праблем інфарматыкі Нацыянальнай акадэміі навук Беларусі

**СУЧАСНЫЯ ТЭХНАЛОГІІ Ў ПЕРАЎТВАРЭННІ ТЭКСТАВАЙ ІНФАРМАЦЫІ  
Ў МАЎЛЕННЕ: СІНТЭЗАТАРЫ МАЎЛЕННЯ ПА ТЭКСЦЕ**

У дадзеным артыкуле разглядаюцца сістэмы сінтэзу маўлення (ССМТ) як праграмны прадукт для агучвання тэкставай інфармацыі. Праведзена падрабязная характарыстыка выкарыстання з апісаннем структуры, відаў і сферы прымянення ў сучасным свеце. Апісаны магчымыя метады ацэнкі якасці сінтэзаванага маўлення і прыведзены прыклады найбольш вядомых сістэм сінтэзу маўлення для розных моў.

Сістэма сінтэзу маўлення – гэта тэхналогія, якая дазваляе пераўтвараць уваходную тэкставую інфармацыю ў агучанае маўленне. Галоўная задача дадзеных сістэм – канвертацыя тэксту ў маўленне, якое не павінна адрознівацца ад натуральнага. Гісторыя распрацоўкі сістэм сінтэзу маўлення пачалася яшчэ ў XVIII ст., але большая частка даследаванняў датычыцца XX ст. у сувязі са стварэннем і развіццём вылічальнай тэхнікі.

Механізм ССМТ складаецца з двух этапаў: гэта канвертацыя “графема – фанема” і генерацыя гукавога файла. Навукоўцы выдзяляюць тры падыходы да стварэння ССМТ у залежнасці ад апрацоўкі гукавога сігналу: гэта фармантны, артыкуляцыйны і канкатэнатыўны. Асноўнымі крытэрыямі ацэнкі натуральнасці сінтэзаванага маўлення лічацца разборлівасць і натуральнасць маўлення, ацэнка працы асобных модуляў сістэмы і ацэнка пазнавальнасці, разумення сэнсу, выразнасці і эмацыйнасці. На сённяшні дзень ССМТ разглядаецца як асобны напрамак даследаванняў і таксама мае шырокі патэнцыял прымянення ў такіх галінах, як лінгвадыдактыка, вакальны маніторынг, тэлекамунікацыйныя паслугі, мультымедыя і г. д.

**Ключавыя словы:** сістэмы сінтэзу маўлення па тэксце, камп'ютарныя тэхналогіі, ацэнка якасці, сінтэз, тэкставая і электронная інфармацыя, канвертацыя.

**Ya. S. Zianouka**

The United Institute of Informatics Problems of National Academy of Sciences of Belarus

**COMPUTER TECHNOLOGIES IN CONVERTING TEXT DATA INTO SPEECH:  
TEXT-TO-SPEECH SYNTHESIZERS**

The article describes text-to-speech systems (TTS) as a software which converts language text into speech with its structure, types and quality evaluation. It also depicts its sphere of application and the most popular examples.

The main task of text-to-speech systems is to convert text information into speech, which should not be different from the natural one. The history of the development began in the XVIII century. But most of the research concerns the twentieth century in connection with the creation and development of computer technology.

TTS mechanism consists of two stages: the conversion of “grapheme-to-phoneme” and generation of a sound file. Scientists distinguish three approaches to creating TTS depending on the audio processing. These are formant, articulation and concatenation synthesis. The main criteria for evaluation of the naturalness of the synthesized speech are intelligibility and naturalness of speech, the assessment of individual system's modules and evaluation of awareness, understanding of the meaning, clarity and emotion. TTS is treated as a separate field of research and therefore has broad potential as a part of other areas such as linguistics, vocal monitoring, telecommunications, multimedia, etc.

**Key words:** text-to-speech systems, computer technologies, quality evaluation, synthesis, text and electronic information, conversion.

**Уводзіны.** У апошнія гады спецыялісты самых розных абласцей праяўляюць усё большую зацікаўленасць да вывучэння тэхналогій у розных сферах дзейнасці, уключаючы фанетыку сучаснай беларускай мовы. Варта падкрэсліць, што назіраецца інтэнсіўнае развіццё ў галіне тэарэтычнага і прыкладнага даследавання маўлення, якое ахоплівае такія напрамкі, як мадэляван-

не артыкуляцыйных і перцэпцыйна-слыхавых працэсаў, аўдыёвізуалізацыі маўлення, фарміраванне баз маўленчых даных і маўленчых фондаў, гаворка ў шумах і перашкодах, маўленчыя камп'ютарныя дыялогавыя сістэмы, фанетычныя аспекты дыялога з камп'ютарам, аўтаматызаваныя і паўаўтаматызаваныя сістэмы ідэнтыфікацыі і верыфікацыі, якія агучваюць

інфармацыю, маўленчыя экспертныя сістэмы, маўленчыя навучальныя сістэмы, аналіз і сінтэз маўлення на базе новых тэхналогій.

**Асноўная частка.** Сёння камп'ютарныя тэхналогіі займаюць вядучую ролю ў інфармацыйным асяроддзі. Адным з важных накірункаў іх прымянення з'яўляецца распрацоўка сістэм сінтэзу маўлення, якія аўтаматычна апрацоўваюць электронныя тэксты ў маўленне. Сістэмы сінтэзу і распазнавання маўлення (ССМТ) – гэта сістэмы, здольныя агучваць апрацаваны тэкст, пераўтвараць маўленне ў тэкст, выконваць пераклад з адной мовы на іншую, ажыццяўляць камунікацыю паміж чалавекам і камп'ютарам з дапамогай толькі голасу і г. д. [1]. Сярод гэтых сістэм асобае месца займае сінтэзатар маўлення па тэксце (СМТ). Гэта камп'ютарная сістэма, якая пераўтварае адвольны тэкст, незалежна ад таго, якім чынам ён уведзены, у гукавы файл праз выкарыстанне спецыяльных праграмных сродкаў [9]. Высакаякасныя ССМТ маюць шырокі патэнцыял прымянення [9]. Сёння сінтэзатар маўлення знаходзіць самае шырокае прымяненне ў разнастайных сферах народнай гаспадаркі, навукі, культуры, медыцыны, адукацыі і інш. Узгадаем найбольш вядомыя.

- Тэлекамунікацыйныя паслугі. ССМТ дазваляюць перадаваць тэкставую інфармацыю праз тэлефонную лінію. Больш за 70% тэлефонных размоў патрабуюць нізкі ўзровень камунікацыі, што дазваляе выкарыстоўваць сінтэзатары маўлення як сродак агучкі неабходных паведамленняў. Гэта датычыцца агучкі электроннай чаргі, разнастайных анонсаў, інфакіёскаў і г. д.

- Лінгвадыдактыка. Высакаякасны сінтэз можа спалучацца з сістэмамі камп'ютарнага навучання і служыць карысным інструментам для вывучэння мовы. Аўдыявізуалізацыя пашырае магчымасці ўспрымання новага матэрыялу як на слых, так і візуальна.

- Дапамога людзям з абмежаванымі здольнасцямі. Машыны могуць быць неацэнным сродкам перадачы маўлення з дапамогай спецыяльна распрацаванай клавіятуры і хуткай праграмы зборкі сказаў. Сістэмы сінтэзу маўлення ў спалучэнні з аптычнымі сістэмамі распазнавання прадастаўляюць доступ да пісьмовай інфармацыі. Рынак праграм для сляпых карыстальнікаў у хуткім часе папоўніцца сінтэзатарамі маўлення ў камплекце з гукавымі картамі. На сённяшні дзень ужо вядомыя гукавыя карты DECTalk апошняй версіі SoundBlaster (NTM), хоць яшчэ не ў фармаце, карысным для сляпых людзей.

- Аўдыякнігі і гаворачыя цацкі. Сёння ўсё большая колькасць людзей аддае перавагу агучаным кнігам замест друкаваных ці электронных крыніц. Таму з'явілася неабходнасць запісу матэрыялаў. Прыцягненне новых дыктараў, на жаль, цалкам не вырашае пытанне па агучцы вялікіх аб'ёмаў

тэксту, таму што гэта даволі доўгі і складаны працэс. Ад усіх дыктараў патрабуецца высокі ўзровень прафесійнасці і прад'яўляюцца такія крытэрыі, як добра пастаўлены голас, падтрыманне адзінага тэмпу і хуткасці агучвання, магчымасць запісу вялікіх аб'ёмаў тэксту і г. д. Таксама неабходна ўлічваць умовы, у якіх адбываецца запіс. Усе пералічаныя пытанні здымае сінтэзаванае маўленне, якое дазваляе эканоміць час, скарачаць затраты на дыктараў і абсталяванне.

- Вакальны маніторынг. У некаторых выпадках вусная інфармацыя з'яўляецца больш эфектыўнай, чым пісьмовыя паведамленні. Гэта дазваляе засяродзіць увагу як на візуальных, так і на гукавых сродках інфармацыі. Адсюль ідэя ўключэння сінтэзатараў у вымяральныя або кіраўнічыя сістэмы.

- Мультымедыя, дыялог паміж чалавекам і машынай. Распрацоўка высакаякасных СМТ з'яўляецца неабходным крокам (як і ўдасканаленне сістэм распазнавання маўлення) у напрамку больш канчатковых сродкаў камунікацыі паміж чалавекам і камп'ютарам. Мультымедыя – гэта першы, але шматбацьцальны крок у гэтым кірунку.

- Фундаментальныя і прыкладныя даследаванні. Сінтэзатары маўлення з'яўляюцца выдатнымі лабараторнымі інструментамі для лінгвістаў. Яны дазваляюць вывучаць і даследаваць розныя аспекты тэарэтычнага і практычнага працэсу маўлення, уключаючы эфектыўнасць інтанацыйных і рытмічных мадэляў. Напрыклад, фармантны тып сістэм шырока ўжываецца фанетыстамі для вывучэння маўлення ў рамках акустычных правілаў.

Калі ж звярнуцца да гісторыі пытання, то трэба ў першую чаргу прыгадаць, што ў канцы XVIII ст. дацкі вучоны Хрысціян Карцэнштайн стварыў мадэль маўленняга тракту чалавека, якая магла прамаўляць некалькі галосных гукаў. Далей з канца XVIII – пачатку XIX ст. шэраг прац В. Кампелена, Ч. Уітстана і А. Бэла дазволілі атрымаць сінтэзаванне галосных і амаль усіх зычных гукаў, а Д. Фабер рэалізаваў спробу сінтэзавання спеваў. Развіццё ў XX ст. электронікі і вылічальнай тэхнікі дазволіла атрымаць якасныя прыклады сінтэзаванага маўлення [4]. У 1930-х гг. супрацоўнікі лабараторыі Бэла (BellLabs) распрацавалі вакодэр – электронны сінтэзатар маўлення для тэлефоннай сеткі. У 1960-х гг. Г. Фант распрацаваў фармантны сінтэзатар, з дапамогай якога ў будучым былі ўпершыню атрыманы прыклады сінтэзу рускага маўлення высокай якасці Б. М. Лабанавым і інш. [2]. На працягу 70–90-х гадоў XX ст. усталяваліся асноўныя метады сінтэзу маўлення: фармантны, мікрахвалевы і кампіляцыйны. У 1990-х гг. былі пабудаваны першыя мадэлі даступных для карыстальнікаў камерцыйных сінтэзатараў па тэксце,

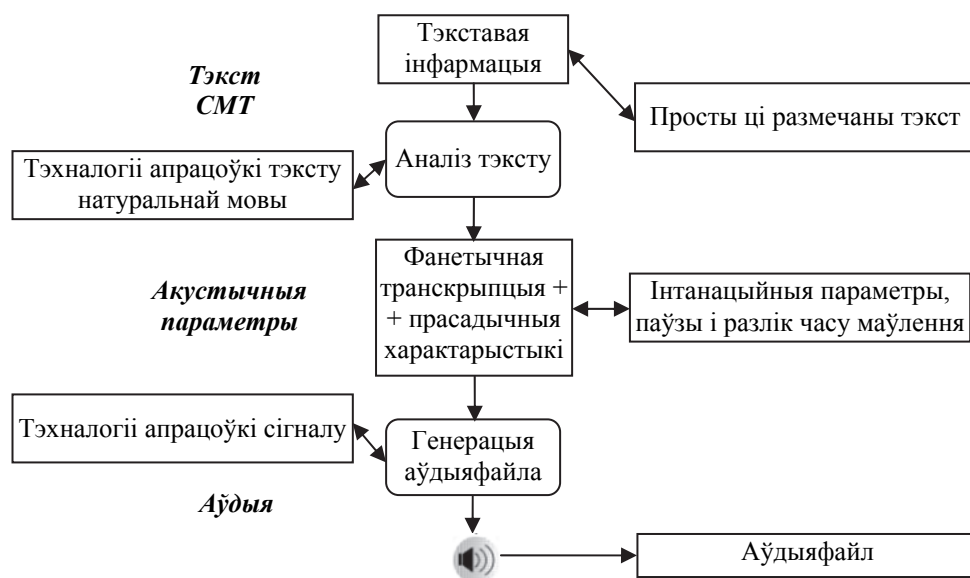
напрыклад DIGALO (www.digalo.com). У 1999 г. у Інстытуце тэхнічнай кібернетыкі (зараз АПП НАН Беларусі) быў рэалізаваны шматгаласавы сінтэз рускага маўлення па тэксце, а таксама сінтэз беларускага, польскага і ўкраінскага маўлення, а на працягу 2000–2007 гг. Удакладняліся навуковыя алгарытмы распрацоўкі сінтэзатараў маўлення па тэксце для славянскіх моў [6].

Пачатак сучаснай гісторыі маўленчых даследаванняў у СССР датуецца сярэдзінай 60-х гг. XX ст., калі ўпершыню пачала працаваць Усесаюзная школа-семинар па аўтаматычным распазнаванні слышавых выяў (АРСВ), якая збірала ў лепшыя гады да 300 удзельнікаў. Да гэтага ж часу адносіцца і пачатковы этап развіцця маўленчых даследаванняў у Беларусі. У 1965 г. у навуковай лабараторыі кафедры радыёпрыёмных прылад Мінскага радыётэхнічнага інстытута пад кіраўніцтвам Б. М. Лабанава была арганізавана група па даследаванні маўленчых сігналаў. У той час у яе ўваходзілі М. П. Дзегцяроў, Б. В. Панчанка, М. К. Фацееў і інш., якія яшчэ доўгі час, а некаторыя з іх і зараз працуюць у гэтым накірунку.

У нашай краіне практычная рэалізацыя сістэмы сінтэзу маўлення звязана з распрацоўкай прыстасавання “Фанемафон-1” у 1971 г. З кожнай новай версіяй якасць сінтэзаванага тэксту паляпшалася. Так, у 1975 г. з’явіўся “Фанемафон-2”, а ўжо “Фанемафон-3” быў прадстаўлены на выставе “Telesom” у Жэневе. Ён атрымаў вельмі высокія водгукі з боку іншых краін-удзельніц: Германіі, Францыі, Італіі, ЗША і інш. [6]. На Сусветным кангрэсе фанетычных навук у 1987 г. свет пабачыў “Фанемафон-4”, праца якога заснавана на фанемна-алафонным метады. Таксама ў ім прадугледжвалася сінтэзаванне не толькі мужчынскім, але і жаночым голасам.

Беларускімі распрацоўшчыкамі сінтэзаванага маўлення ў межах выканання праекта “Сінтэз выразнага маўлення на базе глыбокага сінтаксічнага аналізу чытаемага тэксту” (грант БРФФД – РФФД на 2009–2011 гг.) упершыню створана двухмоўная сістэма сінтэзу выразнага маўлення на беларускай і рускай мовах, якая дэманстравалася на шэрагу навукова-тэхнічных выстаў. У гэтыя ж гады праведзены вялікі аб’ём даследаванняў, накіраваных на распрацоўку беларускамоўнай мадэлі сінтэзу маўлення. Асноўную ролю ў стварэнні сінтэзатара беларускага маўлення адыгралі даследаванні і распрацоўкі Ю. С. Гецэвіча, які зараз з’яўляецца загадчыкам лабараторыі распазнавання і сінтэзу маўлення АПП НАН Беларусі [6]. Пачынаючы з 2011 г. і па сённяшні дзень, лабараторыя распазнавання і сінтэзу маўлення праводзіць тэарытычныя і практычныя даследаванні ў дадзенай вобласці [7]. Ужо распрацаваны беларускамоўны сінтэзатар для стацыянарных платформаў “Мультыфон”, сінтэзатар маўлення для мабільных платформаў і інтэрнэт-версія сінтэзатара [8].

Працэдура канвертацыі тэксту ў маўленне складаецца з двух асноўных этапаў. Першы ўяўляе сабой аналіз уваходнай інфармацыі, дзе тэкст транскрыбіруецца ў фанетычнае прадстаўленне, а другі – генерацыя маўленчых сігналаў, дзе адбываецца вывад апрацаванай інфармацыі ў выглядзе гукавога файла (гл. схему ніжэй) [11]. Першапачаткова паслядоўнасць сімвалаў аналізуецца і апрацоўваецца ў іх фанетычнае прадстаўленне, якое звычайна адлюстравана ў выглядзе паслядоўнасці фанем з дадатковай інфармацыяй для вызначэння правільнай інтанацыі. Дадзены феномен называецца пераўтварэннямі “тэкст – фанема” і “графема – фанема”.



Агульная структура сістэм сінтэзу маўлення [12]

Другі этап складаецца з генерацыі фанетычнага прадстаўлення тэксту ў лічбавы сігнал з выкарыстаннем алгарытмаў сінтэзу. Кожная фанема пераўтвараецца ў асобную адзінку гукавой базы даных. Такім чынам, на гэтым кроку блок канкатэнацыі счытвае фанемы адну за адной з радка вываду фанем. Для кожнай такой фанемы праграма выбірае адпаведную гукавую адзінку ў гукавой базе даных, потым усе апрацаваныя фанемы злучаюцца адна з адной і адбываецца вывад маўленчага сігнала.

Навукоўцы выдзяляюць тры асноўныя падыходы да апрацоўкі маўленчага сігнала СМТ [3]:

1) артыкуляцыйны сінтэз, які заснаваны на фізічнай мадэлі апісання фізіялогіі мова- і гукаўтварэння ў маўленчым апарате з улікам звестак аб артыкуляцыі, што выкарыстоўваюцца для колькаснай ацэнкі формы маўленчага апарату, яго рэзанансных уласцівасцей і характарыстык гукавых крыніц. Асноўныя параметры – гэта пазіцыя і кінематыка артыкулятара. Вымаўлены гук вылічваецца згодна фізічнаму ўраўненню. Дадзены сінтэзатар выкарыстоўваецца даволі рэдка ў сувязі з вялікімі выдаткамі, таму што камп'ютарныя, тэарытычныя і практычныя задачы яшчэ цалкам не вырашаны;

2) фармантны сінтэз, які ўяўляе сабой апісальны, акустычна-фанетычны падыход да сінтэзаванага маўлення. Пры фармантным сінтэзе мадэлююцца акустычныя характарыстыкі маўленчай хвалі замест фізіялагічных працэсаў моваўтварэння. У дадзеным выпадку мадэлююцца галоўныя акустычныя асаблівасці маўленчага сігнала. Пад фармантай разумеюцца частотныя рэзанансы маўленчай акустычнай сістэмы. Параметры фармантаў (частата, шырыня, узровень) вызначаюцца акустычнымі ўласцівасцямі сістэмы. Найбольш важны параметр – частата фарманты, якая цесна звязана з геаметрычнай канфігурацыяй маўленчага апарату. Паколькі ў працэсе гаворкі канфігурацыя маўленчага апарату змяняецца, то адпаведна змяняюцца фармантныя частоты [3];

3) канкатэнацыйны сінтэз мадэлюецца па маўленчым сігнале, які апрацаваны маўленчай базай даных. Сегментаваная база даных створана для адлюстравання галоўных фаналагічных характарыстык мовы. У канкатэнацыйным сінтэзе выкарыстоўваюцца элементарныя адрэзкі натуральнай маўленчай хвалі. У дадзеным выпадку адсутнічае неабходнасць мадэлявання складаных акустычных працэсаў моваўтварэння. Скампіляваны сігнал, складзены з сегментаў натуральнага маўлення, дадаткова апрацоўваецца, каб згладзіць пераходы паміж аб'яднанымі адрэзкамі і змяніць прасадывныя характарыстыкі ў адпаведнасці са значэннямі, якія атрымаліся на этапе прасадывнай апрацоўкі.

Сінтэз маўлення можа разглядацца як праграмны прадукт, да якога прад'яўляюцца пэўныя патрабаванні ў пэўных умовах і для пэўнай задачы. Задача ацэнкі якасці сінтэзаванага маўлення цесна пераплятаецца з тэсціраваннем сінтэзу маўлення як камп'ютарнай праграмы [5]. Сярод задач, для вырашэння якіх можа прымяняцца сістэма ацэнкі якасці сінтэзаванага маўлення, вылучаюцца наступныя[4].

- Тэсціраванне сістэмы сінтэзу ў працэсе яе распрацоўкі.

Галоўная задача такога тэсціравання – наступнае паляпшэнне параметраў сістэмы. Сістэма павінна быць аўтаматычнай, гэта значыць функцыянаваць без удзелу чалавека; мець дастаткова высокую хуткадзейнасць; мець магчымасць ацэньваць сінтэзаванае маўленне як адпаведнасць голасу канкрэтнага дыктара, так і адпаведнасць агульных параметраў маўленчага сігнала. Для ацэначнага аналізу даступныя вынікі ўсіх этапаў сінтэзу, і праверка можа ажыццяўляцца з выкарыстаннем прамежкавай інфармацыі, згенераванай сістэмай у відавочным выглядзе.

- Ацэнка ўласнай сістэмы сінтэзу маўлення ў параўнанні з канкурэнтамі.

Пры ацэнцы ўласнай сістэмы адбываецца параўнанне розных сістэм сінтэзу на розных галасы. Для гэтага можа прымяняцца як аўтаматычная дыктаранезалежная адзнака, так і ацэнка экспертаў. У дадзеным тэсціраванні можа быць ускладнены доступ да вынікаў сінтэзу: для камерцыйных праграм звычайна даступныя толькі інтэрактыўныя дэма-версіі, пры дапамозе якіх можна атрымаць узоры сінтэзаванага маўлення даволі нізкай якасці, з фонавай музыкай або іншымі абмежаванымі ў мэтах абароны ад камерцыйнага выкарыстання, ці даступныя толькі загадзя падрыхтаваныя прыклады гучання. Для карэктнага параўнання вынікаў працы сінтэзатараў неабходна выкарыстоўваць іх поўнафункцыянальныя версіі.

- Удзел у конкурсах, якія праводзяцца незалежнымі кампаніямі.

У такіх умовах сістэма ацэнкі можа быць не аўтаматычнай, але аўтаматызаванай. Прыцягваюцца вялікія чалавечыя рэсурсы (напрыклад, зацікаўленыя карыстальнікі інтэрнэту). Пры гэтым унутраная структура сістэм сінтэзу і застанеца закрытай, але маецца магчымасць атрымання прамежкавых вынікаў працы сістэмы ва ўніфікаваным выглядзе, пры зацікаўленасці ў гэтым саміх удзельнікаў конкурсу. ССМТ могуць тэсціравацца па адной і той жа маўленчай базе, з выкарыстаннем якой будзе сінтэзаваны голас. Можа прымяняцца таксама эталонная адзнака пры параўнанні з дыктарам-донарам.

Асноўнымі агульнапрынятымі мерамі праверкі якасці сінтэзаванага маўлення з'яўляюцца [4]:

- 1) разборлівасць маўлення;
- 2) натуральнасць маўлення;
- 3) ацэнка працы асобных модуляў сістэмы;
- 4) ацэнка пазнавальнасці, разумення сэнсу, выразнасці, эмацыйнасці.

Якасць маўлення – велічыня, якая характарызуе суб'ектыўную ацэнку гучання маўлення падчас тэсціравання. Разборлівасць маўлення – гэта адносная колькасць (у працэнтах) правільна прынятых элементаў (складоў, слоў, фраз) артыкуляцыйных табліц. На аснове гэтага можна лічыць, што пад «якасцю» сінтэзаванага маўлення разумеецца яе натуральнасць, гэта значыць велічыня, якая характарызуе суб'ектыўную ацэнку гучання маўлення ў параўнанні з гучаннем натуральнай гаворкі.

Акрамя вышэйпералічаных магчымасцей ацэнкі якасці сінтэзаванага маўлення неабходна таксама ўлічваць мноства фактараў, якія ўплываюць на яго ўспрыманне. Гэта:

- 1) умовы, у якіх адбываецца праслухоўванне сінтэзаванага маўлення;
- 2) сістэмныя характарыстыкі камп'ютарнага абсталявання;
- 3) вопыт і трэнероўка слухача;
- 4) лінгвістычная структура паведамлення, складанасць тэксту;
- 5) якасць запісу маўленчага сігнала і ўмовы ўспрымання (гучнасць, шум, іншыя перашкоды і г. д.).

Як бачна, ацэнка якасці сінтэзатараў маўлення – гэта складаны працэс, які мае суб'ектыўны характар, таму што цалкам залежыць ад тэсціроўшчыка і ўмоў правядзення эксперыменту па праслухоўванні згенераванага гукавога файла. На сённяшні дзень адсутнічаюць адзіныя правілы па тэсціраванні дадзеных сістэм, таму кожны навуковец выбірае свой асобны падыход для вырашэння дадзенай задачы.

**Заклучэнне.** Такім чынам, аўтаматычны сінтэз маўлення – гэта тэхналогія, якая дазваляе пераўтвараць уваходную тэкставую інфармацыю ў агучнае маўленне. Дадзеныя тэхналогіі

з'яўляюцца даволі папулярнымі, таму што сфера іх прымянення даволі шырокая. Асобны прывілей падобных сістэм – гэта прастата выкарыстання: дастаткова ўвесці тэкст у дыялогавае акно, і машына агучыць згенераваны файл. Такого роду ўзаемадзеянне не патрабуе спецыяльнай падрыхтоўкі і значна павялічвае колькасць карыстальнікаў такіх сістэм. Таксама гэта магчымасць павялічыць хуткасць падчас выкарыстання камп'ютараў, пазбаўляючы ад неабходнасці карыстацца клавіятурай пры ўводзе тэксту. Сінтэз маўлення можа ўжывацца ў сістэмах падтрымкі беспяпярковых тэхналогій (тэкставых файлаў), у маўленчых інтэрфейсах для інвалідаў па слыху і зроку, у сістэмах камп'ютарнай тэлефоніі (даведачныя службы, моўная электронная пошта), у сістэмах маўленчага кіравання (навігацыйныя сістэмы, дыспетчарскія службы), у сістэмах абароны доступу да баз даных (галасавы ключ), у сістэмах для крыміналістычнай экспертызы на аснове голасу і мовы, у навучанні замежным мовам і г. д. [3].

На сённяшні дзень налічваецца вялікая колькасць сінтэзатараў, якія апрацоўваюць розныя мовы. Сярод іх варта адзначыць такія сістэмы, як *AT&T, Nuance, NeoSpeech, Microsoft, IVONA Software, CereProc, Cepstral, Acapela-Group*. Такія сістэмы сінтэзу маўлення па тэксце, як *Acapela, Vokalizer, RHVoice, ESpeak, Festival*, з'яўляюцца якаснымі сродкамі апрацоўкі рускай мовы. Для беларускай мовы кампанія *Sakrament* распрацавала СМТ беларускага маўлення для камерцыйных мэт на аснове рускага СМТ. Лабараторыя распазнавання і сінтэзу маўлення АПП НАН Беларусі на працягу апошніх 55 гадоў займаецца распрацоўкамі сінтэзу маўлення па тэксце [7], якія ўключаюць сінтэзатар для стацыянарных платформаў «Мультифон-4», сінтэзатар маўлення для мабільных платформаў і анлайн сінтэзатар [9]. Галоўная асаблівасць дадзеных распрацовак, у прыватнасці інтэрнэт-сінтэзатара, – адкрыты доступ, гэта значыць кожны бясплатна можа выкарыстаць яго, прайшоўшы па спасылцы <http://corpus.by/TextToSpeechSynthesizer/>.

### Літаратура

1. Гецэвіч Ю. С. Алгарытмы лінгвістычнай апрацоўкі тэкстаў для сінтэзу маўлення на беларускай і рускай мовах: дыс. ... канд. тэхн. навук. Мінск, 2012. 198 с.
2. Лобанов Б. М. Синтезатор персонализированной речи по тексту «ЛобаноФон-2000» // Тр. Междунар. конф., посвящ. 10-летию рос. эксперим. фонетики, 1–4 февр. 2001 г. / С.-Петербург. ун-г. СПб., 2001. С. 147–153.
3. Лобанов Б. М. Компьютерный синтез и клонирование речи. Минск: Беларус. наука, 2008. 357 с.
4. Соломенник А. И. Оценка качества селективного синтеза речи: методы и результаты: дис. ... канд. фил. наук. М., 2016. 127 с.
5. Русанов О. А. Современные технологии синтеза устной речи [Электронный ресурс]. URL: <http://docplayer.ru/38896137-Sovremennye-tehnologii-sinteza-ustnoy-rechi.html> (дата обращения: 29.07.2017).

6. Лабанаў Б. М. Развіццё маўленчых тэхналогій у Беларусі // Лабараторыя распазнавання і сінтэзу маўлення [Электронны рэсурс]. 2016. URL: <https://ssrlab.by/b-m-labanau-razvicio-mauliencych-technologij-u-bielarusi> (дата звароту: 02.01.2017).
7. Лабараторыя распазнавання і сінтэзу маўлення [Электронны рэсурс]. 2017. URL: <http://ssrlab.by/> (дата звароту: 11.07.2017).
8. Сінтэзатар маўлення па тэксце [Электронны рэсурс]. 2017. URL: <http://corpus.by/TextToSpeechSynthesizer/> (дата доступу: 01.07.2017).
9. Dutoit T. An Introduction to text-to-speech synthesis. Norwell, MA, USA: Kluwer Academic Publishers, 1997. 420 p.
10. Ramteke Surendra P. Development of TTS for Marathi Speech Signal Based on Prosody and Concatenation Approach // International Journal of Engineering Research & Technology (IJERT). 2012. Vol. 1, issue 10. P. 82–89.
11. Flanagan J. L., Rabiner L. R. Speech synthesis. Stroudsburg, Dowden, Hutchinson & Ross, 1973. 173 p.
12. Cole R. Survey of the State of the Art in Human Language Technology. Cambridge: University Press and Giardini, 1997. 166 p.
13. Lambert T., Breen A. A Database design for a TTS synthesis system using lexical diphones // Speech Communication and Technology: proceedings of the 9-th European conference InterSpeech'2004, Jeju Island, Korea, October 4–8, 2004. Jeju Island, 2004. P. 1381–1384.
14. d'Alessandro C., Liénard J.-S. Synthetic speech generation // Survey of the State of the Art in Human Language Technology. Cambridge: University Press, 1996. Chapter 5. P. 155–185.
15. Захарьев В. А. Мультиголосовой синтез речи по тексту для построения естественно-языковых интерфейсов интеллектуальных систем // Открытые семантические технологии проектирования интеллектуальных систем = Open Semantic Technologies for Intelligent Systems: материалы Междунар. науч.-техн. конф. Минск, 16–18 февр. 2017 г.: в 2 ч. / Белорус. гос. ун-т информатики и радиоэлектрики. Минск, 2017. Ч. 1. С. 167–170.
16. Taylor P. Analysis and synthesis of intonation using the tilt model // J. Acoust. Soc. America. 2000. Vol. 107, no. 3. P. 1967–1714.

#### References

1. Hetsevich Yu. S. *Algorytmy lingvistychнай apratsouki tekstau dlya sintezu maulennyya na belaruskay i ruskay movakh: dys. kand. tekhn. navuk* [Algorithms of linguistic processing for speech synthesis in Belarusian and Russian languages. Cand. Diss.]. Minsk, 2012. 198 p.
2. Lobanov B. M. Synthesizer of personalized text-to-speech "LobanoFon 2000". *Trudy Mezhdunarodnoy konferentsii, posvyashchennoy 100-letiyu rossiyskoy eksperimental'noy fonetiki* [Proceedings of the International Conference dedicated to the 100th anniversary of the Russian experimental phonetics], St. Petersburg, 2001, pp. 147–153 (In Russian).
3. Lobanov B. M. *Komp'yuternyy sintez i klonirovanie rechi* [Computer synthesis and speech cloning]. Minsk, Belorusskaya nauka Publ., 2008. 357 p.
4. Solomennik A. I. *Otsenka kachestva selektivnogo sinteza rechi: metody i rezul'taty*: dis. kand. filol. nauk [Evaluation of the selective speech synthesis quality: methods and results. Cand. Diss.]. Moscow, 2016. 127 p.
5. Rusanova O. A. *Sovremennyye tekhnologii sinteza ustnoy rechi* [Modern technologies of oral speech synthesis]. Available at: <http://docplayer.ru/38896137-Sovremennyye-tehnologii-sintezo-ustnoy-rechi.html> (accessed 29.07.2017).
6. Labanau B. M. *Razvitstsyо maulenchykh tekhnologiy u Belarusi* [The development of speech technologies in Belarus]. Available at: <https://ssrlab.by/b-m-labanau-razvicio-mauliencych-technologij-u-bielarusi> (accessed 02.01.2017).
7. *Лабараторыя распазнавання і сінтэзу маўлення* [Speech synthesis and recognition laboratory]. Available at: <http://ssrlab.by/> (accessed 11.07.2017).
8. *Сінтэзатар маўлення па тэксце* [Text-to-speech synthesizer]. Available at: <http://corpus.by/TextToSpeechSynthesizer/> (accessed 01.07.2017).
9. Dutoit T. An Introduction to text-to-speech synthesis. Norwell, MA, USA: Kluwer Academic Publishers, 1997. 420 p.
10. Ramteke Surendra P. Development of TTS for Marathi Speech Signal Based on Prosody and Concatenation Approach. *International Journal of Engineering Research & Technology (IJERT)*, 2012, vol. 1, issue 10, pp. 82–89.
11. Flanagan J. L., Rabiner L. R. Speech synthesis. Stroudsburg, Dowden, Hutchinson & Ross, 1973. 173 p.

12. Cole R. Survey of the State of the Art in Human Language Technology. Cambridge, Cambridge University Press and Giardini, 1997. 166 p.
13. Lambert T., Breen A. A Database design for a TTS synthesis system using lexical diphones. *Proceedings of the 9-th European conference InterSpeech'2004 "Speech Communication and Technology"*. Jeju Island, 2004, pp.1381–1384.
14. d'Alessandro C., Liénard J. -S. Synthetic speech generation. *Survey of the State of the Art in Human Language Technology*. Cambridge, Cambridge University Press, 1996, chapter 5, pp. 155–185.
15. Zakhar'ev V. A. Multi-voice text-to-speech synthesis for the construction of natural language interfaces of intelligent systems. *Materialy Mezhdunarodnoy nauchno-tekhnicheskoy konferentsii "Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh sistem"* [Proceedings of the International Scientific and Technical Conference "Open Semantic Technologies for Intelligent Systems"]. Minsk, 2017, part 1, pp. 167–170 (In Russian).
16. Taylor P. Analysis and synthesis of intonation using the tiltmodel. *J. Acoust. Soc. America*, 2000, vol. 107, no. 3, pp. 1697–1714.

### **Інфармацыя пра аўтара**

**Зяноўка Яўгенія Сяргееўна** – аспірант аддзела сучаснай беларускай мовы Цэнтра даследаванняў беларускай культуры, мовы і літаратуры, малодшы навуковы супрацоўнік лабараторыі распазнавання і сінтэзу маўлення. Аб'яднаны інстытут праблем інфарматыкі Нацыянальнай акадэміі навук Беларусі. Нацыянальная акадэмія навук Беларусі (220072, г. Мінск, вул. Сурганава, д. 1, корп. 2, Рэспубліка Беларусь). E-mail: evgeniakacan@gmail.com

### **Information about the author**

**Zianouka Yauhenia Siarheevna** – PhD student, the Department of Modern Belarusian Language, Junior Researcher of Speech Synthesis and Recognition Laboratory. The United Institute of Informatics Problems of National Academy of Sciences of Belarus (1, housing 2, Surganova str., 220072, Minsk, Republic of Belarus). E-mail: evgeniakacan@gmail.com

*Пастыніў 10.01.2018*