

LATVIJAS UNIVERSITĀTE
DATORIKAS FAKULTĀTE

Latviešu valodas runas sintēze

MAĢISTRA DARBS

Autore: **Ginta Garkāje**

Studenta apl. nr. gg10040

Darba vadītāja: pētniece M. dat. Lauma Pretkalniņa

RĪGA 2016

ANOTĀCIJA

“Latviešu valodas runas sintēze” darbā pētīta latviešu valodas runas sintēzes esošā situācija kontekstā ar pasaules straujajām attīstības tendencēm runas tehnoloģijās. Autore izstrādājusi jaunu latviešu valodas runas sintēzes rīku, kā arī apkopojusi informāciju par pēdējos gados izmantotajām runas sintēzes metodēm un citiem latviešu valodas runas sintēzes rīkiem. Runas sintezatori atbilstoši vispārpieņemtai pieejai subjektīvi testēti. Pamatojoties uz testēšanas rezultātiem, izdarīti secinājumi par iespējamajiem uzlabojumiem izstrādātajā runas sintēzes rīka darbībā.

Atslēgas vārdi: runas sintēze, latviešu valodas tehnoloģijas.

ABSTRACT

“Speech synthesis of Latvian language” contains a study of the field of Latvian speech synthesis in context of developments in speech technologies for other languages. Author has developed new tool for Latvian text-to-speech and has gathered and compiled information on other speech synthesis methods and tools used in recent years. The speech synthesizers respective to the common practice are subjectively tested. Conclusions are made based on testing results for possible improvements in the tool.

Keywords: speech synthesis, text-to-speech, Latvian language technologies.

AUTOREFERĀTS

Autore darbā vispusīgi aplūkojusi runas sintēzes tematiku latviešu valodai. Darba ietvaros ir izstrādāts gan teorētiskā materiāla apkopojums (pamatojoties uz literatūras sarakstā minētajiem resursiem un tekstā iekļautajām atsaucēm), gan jauns publiski pieejams rīks. Autore darba ietvaros izstrādāto runas sintēzes risinājumu ir izvērtējusi ar praktiskiem eksperimentiem, salīdzinot sintezētas runas piemērus esošajiem rīkiem latviešu valodai (tai skaitā, izmantojot lietotāju atsauksmes) un darbā no jauna izstrādātajam rīkam.

Autores darbs uzskatāms par veiksmīgi pabeigtu, jo tiek apstiprināta, darbu sākot, izteiktā hipotēze, ka esošos latviešu valodas runas sintezators iespējams pārspēt, izmantojot jaunākās metodes un rīkus. Sintezatora izvērtēšanas rezultāti norāda, ka iespējami uzlabojumi nākamajos pētījumos. Praktiskais risinājums šādam mērķim ir pieejams publiski, darbā tas ir aprakstīts pietiekamā līmenī, lai būtu iespēja tā izstrādi turpināt.

Dokumenta saturs ir izstrādāts un noformēts atbilstoši Latvijas Universitātes Datorikas fakultātes Datorzinātņu maģistra studiju programmas “Maģistra darba izstrādes un aizstāvēšanas metodiskie norādījumi” (ar pēdējām izmaiņām veiktām 2015. gada 9. janvārī), dodot iespēju to viegli izmantot kā materiālu nākamajiem mācību darbiem.

Darbs izstrādāts ar pārtraukumiem: vispirms teorētiskā literatūra analizēta 2015. gada jūnijā, savukārt 2015. gada decembrī un 2016. gada janvārī un februārī tā pārskatīta un atjaunota atbilstoši krasajām izmaiņām runas tehnoloģiju nozarē. Straujās izmaiņas pētāmajā jomā vienlaikus norāda gan uz tēmas aktualitāti, gan uz iespējamu darba teorijas un risinājuma novecošanās risku.

SATURA RĀDĪTĀJS

APZĪMĒJUMI UN TERMINI	6
IEVADS	8
1. RUNAS SINTĒZES PAMATI	10
1.1. Pamatjēdzieni	10
1.2. Veidi	11
1.3. Vēsture.....	11
1.4. Pielietojumi	12
2. RUNAS SINTĒZES METODEDES	13
2.1. Runas sintēzes process	13
2.2. Konkatenatīvā sintēze.....	17
2.3. Artikulārā sintēze	18
2.4. Parametriskā runas sintēze	18
2.5. Metožu salīdzinājums.....	20
3. RUNAS SINTĒZES RĪKI.....	21
3.1. Pasaulē izmantotie runas sintezatori.....	21
3.2. Latviešu valodas runas sintezatori.....	21
3.3. Citi rīki	22
4. LATVIEŠU VALODAS RUNAS SINTĒZES RĪKS “RUNĀTĀJS”	23
4.1. Uzbūve un darbības principi.....	23
4.2. Izmantotie latviešu valodas runas resursi	25
4.3. Versiju apraksts	26
5. EKSPERIMENTĀLIE NOVĒRTĒJUMI.....	28
5.1. Eksperimentu konteksts.....	28
5.2. Stratēģija un kritēriji.....	28
5.3. Organizācija.....	29
6. REZULTĀTI UN DISKUSIJA.....	30
6.1. Latviešu valodas runas sintezatoru salīdzinājums.....	30

6.2. <i>Runātājs</i> sintezatora versiju salīdzinājums.....	30
7. SECINĀJUMI.....	33
8. NOBEIGUMS.....	35
PATEICĪBAS.....	36
IZMANTOTĀ LITERATŪRA UN AVOTI.....	37
PIELIKUMI.....	40
1. pielikums – Testpiemēri <i>Runātājs (Laika_zinas_2)</i>	40
2. pielikums – Testpiemēri <i>eSpeak</i>	45
3. pielikums – Testpiemēri <i>Tildes Biroja</i> sintezatoram.....	48
4. pielikums – Testpiemēri <i>runa.ailab.lv</i>	50

APZĪMĒJUMI UN TERMINI

<i>Apzīmējums vai termins</i>	<i>Skaidrojums</i>
1. <i>Allab</i>	Latvijas Universitātes Matemātikas un informātikas institūta Mākslīgā intelekta laboratorija.
2. <i>Apmācības kopa</i>	Mašīnmācīšanās jomā – paraugdatu kopa, kuru lieto statistisko modeļu apmācībai [7].
3. <i>Difons</i>	Divu fonu apvienojums
4. <i>Fonēma</i>	Mazākā skaņu valodas vienība, kas šķir vārdu un morfēmu (vārda vismazākā, nedalāmā gramatiski vai leksiski nozīmīgā daļa) skanējumu [1].
5. <i>Fonētiskās transkripcijas</i>	“Skaņu atspoguļošanas veids, sastopams cilvēciskās valodās neatkarīgi no viņu rakstības. Visplašāk zināma sistēma ir starptautiskais fonētiskais alfabēts (<i>International Phonetic Alphabet</i>).” [30]
6. <i>Fons</i>	Mazākais valodas fonētiskais segments (skaņa) pirms tā identificēšanas ar konkrētu valodas patskani, divskani, līdzskani vai atbilstošo fonēmu [1].
7. <i>Intonācija</i>	Īpašību kopa, “kas izsaka daļu no izteikuma satura” [1]. Intonācijā ir iekļauts skaņas augstums, ilgums un tembra maiņas (“skaņai raksturīga īpatnēja nokrāsa, ar ko atšķiras vienāda augstuma skaņas” [1]).
8. <i>Korpuss</i>	Tekstu un/vai transkribētas runas krājums mašīnlasāmā formā [1].
9. <i>Mašīnmācīšanās</i>	Mākslīgā intelekta nozare, kas saistīta ar sistēmu apmācību [7].
10. <i>N-grammas</i>	Nepārtraukta n simbolu gara virkne.
11. <i>Normalizācija</i>	Teksta pārveidošana vienotā formā.
12. <i>Prosodija</i>	Fonētisko parādību kopums — skaņu frekvences, ilguma, intensitātes, spektra maiņas, kas veido zilbes, vārdus, sintagmas, teikumus u. tml. [1].
13. <i>Raksturiezīme (feature)</i>	Atsevišķa pazīme, kas norāda uz kādu datu īpašību.

<p><i>14. Spektrogramma</i></p>	<p>Grafisks spektra attēls.</p> <p>Dinamiskā spektrogramma val. — spektrogrammas veids, kurā grafiski fiksē runu kā dažāda garuma segmentu spektrālo izmaiņu nepārtrauktu ainu.</p> <p>Spektrogrammu metode val. — akustiskā valodas pētīšanas metode, kuras pamatā ir runas skaņu spektrogrammas akustiskā analīze [1].</p>
<p><i>15. Tilde</i></p>	<p>SIA “Tilde”</p>
<p><i>16. Universālā runas sintēze</i></p>	<p>Saturā neierobežota teksta runas sintēze.</p>
<p><i>17. Zilbe</i></p>	<p>Vārda skaņu vai skaņu kopu, ko izrunā ar vienu elpas izdvesumu [1].</p>

IEVADS

Runas tehnoloģijas kā lietotāja saskarnes sastāvdaļa pēdējo gadu laikā kļuvušas arvien populārākas — tās strauji attīstījušās komerciālos produktos (mobilo telefonu balss saskarnēs, automobiļu navigācijas sistēmās, "gudro māju" tehnoloģijās u.c.). Tādas kompānijas kā *Amazon*, *Google*, *Microsoft*, *Toshiba* ilglaicīgi ieguldījušas līdzekļus runas sintēzes un atpazīšanas tehnoloģijās angļu, spāņu, japāņu, ķīniešu, indiešu un citām valodām.

Runas sintēzes rīkus izstrādā gan tādēļ, lai sintezētu atsevišķas vārdkopas vai to savienojumus, gan arī lai realizētu tā saukto universālo runas sintēzi [10], kas spēj sintezēt visus burtu savienojumu konkrētajai valodai. Dažādās valodās izmantoto skaņu un izrunas īpatnību dēļ universālās runas sintēzes risinājumi ir valodatkarīgi.

Universālā runas sintēze nepieciešama medicīnā (piemēram, cilvēkiem ar redzes vai runas traucējumiem), izglītībā (piemēram, izglītības materiālu ierunāšanā) un dažādu viedtehnoloģiju darbībā. Darbā apskatīta latviešu runas sintēzes tehnoloģiju attīstība, analizējot esošajos risinājumos izmantotās metodes minētajās jomās, kā arī salīdzināta esošo latviešu valodas runas sintēzes rīku darbība.

Latviešu valodas runas sintēze ir svarīga tēma arī latviešu valodas tehnoloģiju pētniecības kontekstā. Cilvēka un datora saskarne bez latviešu valodas integrācijas ilgtermiņā var radīt kaitējumu valodas saglabāšanai un attīstībai, jo atšķirīgās kvalitātes dēļ latvieši kā datora saskarnes valodu izvēlas citu valodu.

Latviešu valodas tehnoloģiju pētnieki pēdējos divos gados veikuši pētījumus un radījuši publiski pieejamus rīkus latviešu valodas runas atpazīšanai (SIA "Tilde" *Runas atpazīnējs* [15], AILab runas atpazīšana [21]), kā arī izstrādājuši latviešu valodas runas resursus turpmākai attīstībai šinī jomā (ir izstrādāts runas korpuss [2] [21]). Savukārt latviešu valodas runas sintēze nav plaši komerciāli izmantota, temats pēdējos desmit gados nav apskatīts publikācijās.

Darba mērķis ir izvērtēt esošo situāciju latviešu valodas runas sintēzei un izstrādāt savu pētījumu par to, kā veidot latviešu valodas runas sintēzi tā, lai tā būtu izmantojama iepriekš minētajos pielietojumos. Darba hipotēze ir, ka esošos latviešu valodas runas sintezators iespējams pārspēt, izmantojot pēdējo gadu laikā izstrādātās metodes un rīkus.

Mērķa sasniegšanai ir izvirzīti šādi darba uzdevumi:

- 1) veikt teorētisko apskatu par mūsdienu runas sintēzes rīkos izmantotajām metodēm;
- 2) eksperimentāli izvērtēt esošo latviešu valodas runas sintēzes rīku kvalitāti;

3) izstrādāt latviešu valodas runas sintēzes rīku un eksperimentāli izvērtēt tā darbību.

Lai izpildītu darba uzdevumus, izanalizēti jaunākie runas tehnoloģiju raksti zinātniskajās konferencēs un žurnālos. To analīze atspoguļota runas sintēzes problemātikas un metožu aprakstos. Tāpat eksperimentāli novērtēti brīvi pieejamie runas sintēzes risinājumi latviešu valodai. Darba praktiskajā daļā ir izstrādāts un eksperimentāli izvērtēts autores runas sintēzes risinājums salīdzinājumā ar citiem pieejamajiem universālajiem runas sintēzes rīkiem latviešu valodai.

Darbam ir septiņas pamata nodaļas, tās strukturētas, vispirms aprakstot runas sintēzes teoriju, tad - runas sintēzes rīkus. Darba ietvaros izstrādātais rīks aprakstīts divās nodaļās, skaidrojot tā uzbūves principus un aprakstot tā izvērtēšanas pieeju un rezultātus. Darba rezultātā izdarītie secinājumi ir lasāmi nobeigumā.

1. RUNAS SINTĒZES PAMATI

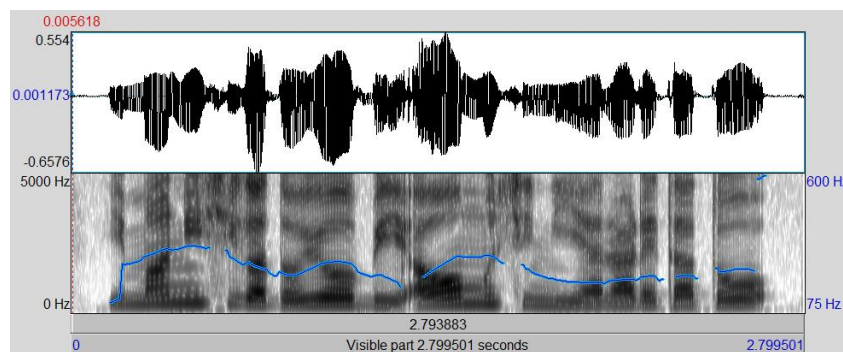
1.1. Pamatjēdzieni

Runa ir cilvēka saziņas pamats, kuru veido liels apjoms dažādu vārdu katrā valodā. Runas pamata vienība ir fonēma, kas ir “mazākā skaņu valodas vienība, kas šķir vārdu un morfēmu (vārda vismazākā, nedalāmā gramatiski vai leksiski nozīmīgā daļa) skanējumu” [1]. Kā vārda pamata vienību izmanto zilbi – “vārda skaņu vai skaņu kopu, ko izrunā ar vienu elpas izdvesumu” [1]. Vienu runu no citas runas atšķir tās intonācija, kas ir īpašību kopa, “kas izsaka daļu no izteikuma saturu” [1]. Intonācijā ir iekļauts skaņas augstums, ilgums un tembra maiņas (“skaņai raksturīga īpatnēja nokrāsa, ar ko atšķiras vienāda augstuma skaņas” [1]).

Fonētikā, kas pētī runas skaņas, runu attēlo, izmantojot:

1. Spektrogrammu metodi – “akustiskā valodas pētīšanas metode, kuras pamatā ir runas skaņu spektrogrammas (skat. 1. attēlu) akustiskā analīze” [1];

1. attēls – “Runas skaņas vilnis un spektrogramma¹”



2. Fonētiskās transkripcijas (skat. 2. attēlu vārdiem ‘Zaļais Glāzušķūnis’) – “skaņu atspoguļošanas veids, sastopams cilvēciskās valodās neatkarīgi no viņu rakstības. Visplašāk zināma sistēma ir starptautiskais fonētiskais alfabēts (*International Phonetic Alphabet*)” [30].

2. attēls – “Fonētiskās transkripcijas piemērs²”

[ˈzɑlɑ̃is ˈɡlɑːzuʃcuːnɪs]

Runas tehnoloģijās visbiežāk tiek izmantota spektrogrammas metode. Izmantojot speciālus rīkus, tiek apstrādāts runas signāls un transkribēti runā izteiktie vārdi un pauzes

¹ <http://media.thedailytouch.com/2015/02/520image.png> [aplūkots 25.01.2016.]

² https://lv.wikipedia.org/wiki/Starptautiskais_fon%C4%93tiskais_alfab%C4%93ts [aplūkots 21.05.2016.]

(skat. nod. “3.3. Citi rīki”). Universāla runas sintēzes sistēma (skat. nod. “1.2. Veidi”) “neliek kopā atsevišķus vārdus vai teikuma daļas, bet automātiski rada runu no fonētiskās transkripcijas” [31].

1.2. Veidi

“Runas sintēzi pēc informācijas reprezentācijas veida iedala ierobežotā (ziņojumapmaiņas, arī — strukturālā) un neierobežotā (teksta–runas, arī — universālā) sintēzē. Šis iedalījums lielā mērā sakrīt ar attiecīgo sistēmu izmantojuma plašumu. Pirmā tiek lietota, piemēram, telekomunikāciju, telefonbanku un citās ierobežota dialoga sistēmās, bet otrā – universāla izmantojuma sistēmās, piemēram, vājredzīgajiem un neredzīgajiem cilvēkiem, datorizētā valodas mācīšanās.

Visātrāk un visvienkāršāk ir panākt to, ka dators var runāt ar ierobežotās runas sintēzes palīdzību, t. i., var izmantot jau iepriekš noteiktam mērķim ieskaņotu vai mākslīgi radītu balsi, liekot konkrētam skaņu segmentam (vārdam, frāzei, teikumam) atskanēt automātiski (piemēram, dažāda veida paziņojumi datorā, laika ziņas pa telefonu u. tml.) vai nospiežot pogu (piemēram, runājošas bērnu rotaļlietas)” [31].

1.3. Vēsture

Cilvēce runu ar iekārtu pirmo reizi mēģinājusi sintezēt pirms vairāk nekā 200 gadiem, kad 1769. gadā tika izstrādāts pirmais mehāniskais runas sintezators – *Wolfgang von Kempelen's Speaking Machine* [23]. Iekārta atgādināja plēšas un bez svarīgās mehāniskās komponentes – mēles – nebija piemērota garāku vārdu vai teikumu atskaņošanai [14].

Pirmais elektroniskais universālais runas sintezators pazīstams ar nosaukumu *Vocoder*. Tas tika izstrādāts 20. gs. trīsdesmitajos gados. Ar to bija iespējams atskaņot zemas kvalitātes universālu angļu valodas runu ar izteikti metālisku skanējumu [22]. Kā lielākais mīnuss minēta tā sarežģītā lietošana – bijis nepieciešams gads, lai apgūtu šī elektroniskā sintezatora izmantošanu [14].

Pirmā runas sintēzes sistēma, kura rada mūsdienu skanējumam līdzīgu skaņu, tika izstrādāta *Massachusetts Institute of Technology* 20. gs. astoņdesmito gadu sākumā, tās nosaukums – *DECtalk*. Iekārta tikusi lietota automātisko atbildētāju funkciju nodrošināšanai un bija izmantojama bez speciālista palīdzības [24].

1.4. Pielietojumi

Mūsdienu universālās runas sintēzes sistēmas veic dažādu līmeņu teksta apstrādi: sākot no teksta sagatavošanas automatizētai fonētiskai transkribēšanai, beidzot ar audiosignāla ģenerēšanu un prosodijas modelēšanu [13] (Prosodija ir “fonētisko parādību kopums — skaņu frekvences, ilguma, intensitātes, spektra maiņas, kas veido zilbes, vārdus, teikumus u. tml.” [1]). Sasniegtā sintezētā runas kvalitāte radījusi virkni pielietojumu tādās nozarēs kā veselība, izglītība, telekomunikācijas un viedās tehnoloģijas.

Veselības nozarē runas sintēzē pielietojumi ir personām ar redzes traucējumiem automātiskai informācijas lasīšanai. Arī cilvēkiem ar disleksiju runas sintēze palīdz uztvert rakstīto tekstu, to nolasot priekšā. Tiek lēsts, ka 6-7% skolas vecuma bērnu piemīt disleksija [28], kas liecina par runas sintēzes būtiskajām pielietojuma iespējām. Kā viens no zināmākajiem pielietojumiem ir sintezēta runa cilvēkiem, kuri ir zaudējuši balsi. Mūsdienās tiek piedāvāta pielāgota sintēzes izstrāde tieši konkrētā cilvēka runas veidam. Izstrāde ir dārgs process, kas prasa lielu datu apjomu, bet rezultāts ir ļoti nozīmīgs konkrētā cilvēka ikdienas dzīvei.

Izglītībā runas sintēze var tikt izmantota valodu apgūvē, mācoties izrunu, un kā mācību līdzeklis rakstiskā teksta apgūšanai, klausoties to nolasām. Mūsdienu tendence ir runas sintēzi apvienot ar mašintulkošanu, piemēram, tērzēšanas programmās, norunātais tiek automātiski atpazīts par tekstu un parādīts tērzēšanas programmā, tad automātiski iztulkots un norunāts otram saziņas biedram. Praktiska realizācija izstrādāta, piemēram, programmā *Skype* [29].

Mūsdienu tendence ir arī aizvien vairāk runas tehnoloģijas izmantot viedajās tehnoloģijās. Runas sintēze tiek plaši izmantota mobilajos telefonos, gudrajās mājās un virtuālajos asistentos. Virtuālie asistenti ir ierīces, kas spēj atpazīt mūsu prasīto un atbildēt. Piemēram, uz jautājumiem par rītdienas temperatūru, meklējumiem internetā, jaunākajām ziņām utt. Kā piemēri ir minami: *Apple Siri*³, *Amazon Echo*⁴, *Microsoft Cortana*⁵ u.c. Iepriekš minētie rīki un lietojumi ir pieejami angļu, spāņu, japāņu un citām valodām, bet nav pieejami latviešu valodai (skat. nod. “3. Runas sintēzes rīki”). Uzņēmumi, kas izstrādājušas šos produktus, ir galvenie zinātnisko rakstu autori par mūsdienu runas sintēzes metodēm.

³ <http://www.apple.com/ios/siri/> [aplūkots 21.01.2016]

⁴ <http://www.amazon.com/Amazon-SK705DI-Echo/dp/B00X4WHP5E> [aplūkots 21.01.2016]

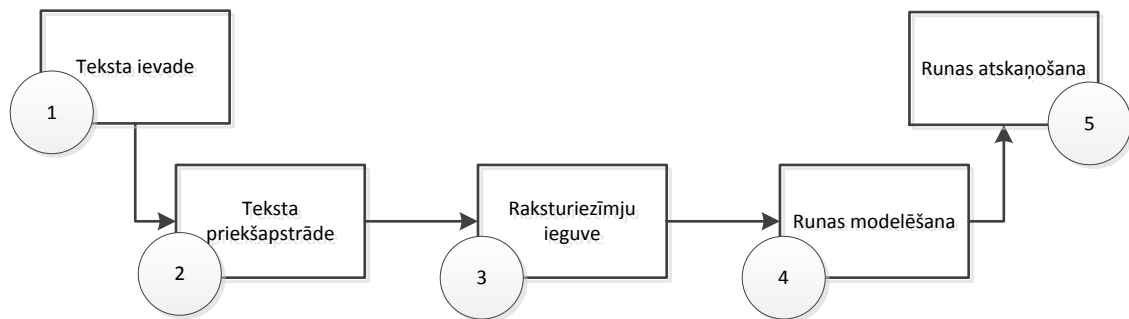
⁵ <https://www.microsoft.com/en-us/mobile/experiences/cortana/> [aplūkots 21.01.2016]

2. RUNAS SINTĒZES METODES

2.1. Runas sintēzes process

Runas sintēzes sastāv no teksta automātiskas pārveides uz runu. 3. attēlā ir redzams shēma runas sintēzes procesam. Attēlā redzamā procesa soļi teksta priekšapstrādei un pārvēršanai par runu ir atkarīgi no konkrētā pielietojuma mērķa un pieejamā resursu daudzuma. Procesu soļi sadalīti pa tajos risināmajiem uzdevumiem. Turpmākajās nodaļās aprakstīta katra soļa nozīme runas sintēzes procesā un metodes, kas tiek izmantotas uzdevumu risinājumiem.

3. attēls – “Runas sintēzes process”



2.1.1. Teksta ievade

Teksta ievades rezultātā tiek iegūts datorlasāms teksts. Atkarībā no sintezatora izmantošanas veida atšķiras teksta ievades saskarne. Visbiežāk tas ir no klaviatūras ievadīts teksts, taču tas var būt arī skenēts un atpazīts vai automātiski tulkots teksts. Šādi iegūtā tekstā bieži ir vairāk kļūdu, kas var apgrūtināt tālāko apstrādi un sintēzi. Cilvēkiem ar kustību traucējumiem populārs veids teksta ievadei ir acu izsekošanai (*eye trackers*), kur iekārta uztver acu zīlīšu pozīciju, kas skatās uz klaviatūrai līdzīgu ekrānu. Sintezatoriem vēl viens integrācijas ziņā atšķirīgs veids ir ekrānlasītāji, kas nolasa visu tekstu izvēlētajā programmas logā. Atšķirībā no ievades saskarnes var mainīties, kā jāveic teksta priekšapstrāde, runas modelēšana un atskaņošana.

2.1.2. Teksta priekšapstrāde

Teksta priekšapstrādē tiek īstenota ieejas teksta pārveidošana uz tādu tekstu, kuru lasot skaļi, nekā nebūtu jāizlasa vai jāpārveido. Ievaddati ir teksts un izvaddati ir runas sintēzes vajadzībām apstrādāts teksts. Soli tiek īstenoti turpmāk lasāmie uzdevumi.

1. Pieturzīmju apstrāde – atkarībā no pielietojuma pieturzīmes tiek pārveidotas par vārdiem (piemēram, ‘;’ aizstāti ar ‘semikols’) vai dzēstas, ja tās izrunāt nevajag.
2. Speciālo simbolu apstrāde – tiek dzēstas liekas atstarpes, pēdiņas un citi simboli, tiek pārveidoti speciālie simboli par vārdiem (piemēram, ‘@’ par ‘at’);
3. Saīsinājumu apstrāde – atsevišķi saīsinājumi tiek pārveidoti lasīšanai pa burtiem (piemēram, ‘ASV’ par ‘ā es vē’);
4. Skaitļu apstrāde – tiek pārveidoti skaitļa vārdi to atbilstošajos locījumos (piemēram, ‘5 aitas’ par ‘piecas aitas’).

Apjoms, kādā veikt šos uzdevumus, ir atkarīgs lietotāju vajadzībām konkrētajā sintezatora pielietojumā. Kā arī tas ir atkarīgs no teksta veida, piemēram, tiešsaistes tekstiem nepieciešama papildus apstrāde arī interneta adrešu un emocijzīmju apstrādē [5].

2.1.3. Raksturiesīmju ieguve

Raksturiesīmju solī no apstrādāts teksts tiek pārveidots uz runas modelēšanai nepieciešamo formātu. Soļa rezultāts ir raksturiesīmju (*features*) vektors, kas satur fonēmas, pauzes, uzsvarus un intonācijas. Latviešu valodas gadījumā šajā solī tiek īstenoti turpmāk lasāmie uzdevumi un risinātas turpmāk lasāmajās nodaļās aprakstītās problēmas latviešu valodas runas sintēzes gadījumā.

1. Fonēmu iegūšana;
2. Intonācijas iegūšana;
3. Paužu iegūšana;
4. Uzsvaru iegūšana.

2.1.3.1. Fonēmu iegūšana

Latviešu valodā, kur vārdu izruna no to rakstības maz atšķiras, fonēmu ģenerēšana ir vieglāks uzdevums kā, piemēram, angļu valodā, kur vārdu izruna atšķiras no rakstības. Šādos gadījumos mēdz sintezējamā teksta burtus izmantot kā fonēmas. Šāda pieeja pielietota esošajos latviešu valodas sintezatoros (skat. nod. “3.2 Latviešu valodas runas sintezatori”). Taču, izmantojot šādu risinājumu, nepieciešams papildus darbs ar turpmāk aprakstītajām, citādi sintezētajai runai rodas nedabisks skanējums [31].

- Šaurā un platā ‘e’ izruna.

Vārdā ‘deguns’ tas ir plats, bet vārdā ‘ezis’ – šaurs. Ir gadījumi, kuros burta izruna mainās atkarībā no locījuma. Piemēram, vārdā ‘lecu’ ‘e’ ir plats, bet vārdā ‘lekt’ – šaurs. Ir arī gadījumi, kuros burta ‘e’ izruna mainās atkarībā no konteksta, lai gan rakstība paliek tā pati. Piemēram, frāzē “es vakar ēdu” vārdā ‘ēdu’ burts ‘e’ tiek izrunāts šauri, bet frāzē “es tagad ēdu” burts ‘ē’ tiek izrunāts plati. Lai gan eksistē

likumsakarības, ar kurām var noteikt, kā burts būs jāizrunā konkrētajā vārdā, eksistē daudz izņēmuma gadījumu. Veids, kā nodrošināt pareizu burta izrunu, ir uzskaitīt visus izņēmuma gadījumus, taču tas ne vienmēr ir iespējams.

- Līdzīgi apsvērumi ir par ‘o’. Tas var būt: īss, garš vai pat divskanis.
Piemēram, vārdā ‘bļoda’ burts ‘o’ ir kā divskanis un vārdā ‘lokomotīve’ burts ‘o’ ir īss, savukārt vārdā ‘zona’ burts ‘o’ ir garšs.
- Divskaņi un to pareizs dalījums zilbēs.
Piemēram, ‘neiet’ izrunā ‘ne iet’, nevis ‘nei et’.

Iepriekš minētās īpatnības var risināt, definējot likumus vai izmantojot dažādas mašīnmācīšanās metodes. Problēma ir līdzīga kā mašīntulkošanā, kur tiek tulkota viena valoda uz otru. Runas sintēzē valoda būtu vārdu tulkošana uz to fonētisko pierakstu, kur teikums ir vārds sadalīts par burtiem un vārds ir viens burts.

2.1.3.2. Uzsvāru iegūšana

Lai sintezētā runa izklausītos dabiski, svarīgi ir arī pareizi vārdu uzsvāri. Vairākās valodās, piemēram, krievu valodā, atkarībā no uzsvērtās zilbes var mainīties vārda nozīmē (čurāt un rakstīt - писать). Latviešu valodā uzsvāri vairumā gadījumu ir salīdzinoši vienkārši, parasti tiek uzsvērtā pirmā zilbe [31]. Izņēmuma gadījumi ir daļa svešvārdu, piemēram, vārdos, kas beidzas *o* (Arno, Bruno) uzsvērtā parasti ir pēdējā zilbe. Līdzīgi kā fonēmu modelēšanā, arī šo var risināt vai nu ar mašīnmācīšanos vai ar likumiem.

2.1.3.3. Paužu iegūšana

Dabiskā runā starp vārdiem pauzes bieži nesakrīt ar rakstītā teksta pieturzīmēm un atstarpēm. Piemēram, izrunājot ‘viss sākas’, nav pauzes starp abiem vārdiem. Ieturot šo pauzi, runas uztveramība tiek pat apgrūtināta – tā šķiet saraustīta. Līdzīgs efekts rodas, lasot tekstu, kur aiz katra vārda ir punkts, piemēram, “šeit. ir. punktu. pārspīlējums., kas. apgrūtina. teksta. uztveri.” Paužu iegūšanā runas sintēzē apskata pauzes starp teikuma daļām. Teikuma daļu robežas nosaka, izmantojot visas pieturzīmes, kas nav aizstātas teksta priekšapstrādes procesā. Atkarībā no pieturzīmes veida, atšķiras nepieciešamās pauzes garums.

2.1.3.4. Intonācijas iegūšana

Atšķirībā no uzsvāru iegūšanas, kas tiek veikta zilbēm un kas ietekmē vārdu izrunu, intonācijas iegūšana tiek veikta izteikumiem. Izteikumu beigās, kas izsaka stāstījumu, intonācija iet uz leju, bet to, kas izsaka jautājumu, intonācija iet uz augšu [31]. Pieturzīmes, kas netika aizstātas teksta priekšapstrādes posmā, jāiekļauj kā intonācijas izmaiņas. Lai gan vairumā gadījumu intonācija iet uz leju, eksistē izņēmuma gadījumi, kuros intonācija atšķiras

un to ir grūti pareizi noteikt. Piemēram, saliktos jautājuma teikumos kā “Vai tu zini, cik ir pulkstenis?”. Lai gan jautājuma zīme ir teikuma beigās, jautājuma intonācija ir virsteikumam un stāstījuma intonācija ir palīgteikumam. Arī tipiskākās jautājuma partikulas, piemēram, ‘vai’, ‘kas’, ‘ko’, ‘kurš’, neizsaka, ka konkrētajai teikuma daļai nepieciešama jautājuma intonācija, jo ar šīm partikulām var sākties arī palīgteikums.

2.1.4. Runas modelēšana

Runas modelēšanā no raksturierzīmju vektora tiek modelēta ir dabiska runa. Viens no vispārīgiem runas modelēšanas veidiem sastāv no šādiem četriem soļiem [11]:

1. Fonēmu sintēze un sapludināšana;
2. Uzsvāru modelēšana.
3. Paužu modelēšana;
4. Intonācijas modelēšana;

Vispirms tiek veikta fonēmu sintēze. Tajā ir svarīgi ņemt vērā kontekstu, jo tas nosaka pārēju no vienas fonēmas uz otru. Konkatenatīvajā sintēzē (skat. nod. “2.2 Konkatenatīvā sintēze”), lai veiktu fonēmu sapludināšanu, kā mazākā vienība tiek izmantota otra puse no pirmās fonēmas un pirmā puse no otrās fonēmas. Arī vārdu robeža tiek uzskatīta kā fonēma, un no tās tiek ņemtas puses vārdu pirmajai un pēdējai fonēmai. Parametriskajā sintēzē (skat. nod. “2.4. Parametriskā runas sintēze”) tiek izmantotas fonēmu n-grammas, lai nodrošinātu fonēmu sapludināšanu.

Pēc viena vārda fonēmu sintēzes tiek veikta uzsvāru modelēšana šim vārdam. Pēc katra vārda sintēzes tas tiek pievienots jau uz sintezētajiem vārdiem, starp tiem ievietojot atbilstoša garuma pauzi, ja tas ir nepieciešams. Kad vairāku sintezēto vārdu kopums veido izteikumu, šim izteikumam tiek veikta intonācijas modelēšana, kas arī veido gala audio failu.

2.1.5. Runas atskaņošana

Runas sintēzes procesa rezultāts ir runa, kas tiek atskaņota lietotājam. Runas sintezatora izstrādes procesā tas ir galvenais rezultāts, kuru izvērtēt. Sintezētu runu raksturo tās dabiskums un kvalitāte. No tehniskiem parametriem, parasti sintezatoriem ir pieejamas izvēles iespējas:

1. Atskaņošanas formātam;
2. Atskaņošanas ātrumam;
3. Burtošanai.

2.2. Konkatenatīvā sintēze

Konkatenatīvās runas sintēzes pamatā ir iepriekš ierakstītu runas audio fragmentu savienošana. Metode tiek izmantota runas modelēšanas solī (skat. nod. “2.1. Runas sintēzes process”).

2.2.1. Domēnspecifiska runas sintēze

Domēna specifiska runas sintēze ietver konkrētu iepriekš ierakstītu vārdu un frāžu savienošānu [9]. Metode ir plaši izmantota gan latviešu, gan citām valodām gadījumos, kuros ir ļoti ierobežota atskaņošana nepieciešamā vārdkopa, piemēram, vilcienu stacijas informācijas paziņojumi par vilcienu pienākšanu un atiešanu, laika ziņas, pulksteņa laika atskaņošana informatīvos paziņojumos u.c. Metode sniedz labus rezultātus konkrētajiem lietojumiem, bet nav vispārināma universālās runas sintēzes gadījumam, jo nespēj atskaņot iepriekš neieskaņotus vārdus un saīsinājumus.

2.2.2. Difonu runas sintēze

Difonu (divu mazākais valodas fonētiskais segmentu (fonu) apvienojums) runas sintēzei atšķirībā no domēna specifiskas sintēzes tiek izmantots nevis vārdu runas korpusu, bet valods specifisku skaņas pāreju (divu skaņas vienību apvienojums) korpusu. Katrai valodai šāds skaņu komplekts, lai gan kvantitatīvi atšķirīgs, ir ierobežots. Difonu runas sintēze ir izmantojama universālai runas sintēzei. Izmantojot šo sistēmu, netiek zaudēta kvalitāte, sintezējot izrunu arī vārdam, kas nav atrodamas vārdnīcā. Tās lietojums pēdējos gados samazinās, jo attīstās citas runas sintēzes metodes ar labākiem kvalitātes rādītājiem [9]. Šīs metodes realizācijas galvenā problēma ir atrast neitrālus difonus, kas izklausās dabiski konkatenācijas vietās.

2.2.3. Vienību izvēles runas sintēze

Vienību izvēles runas sintēzē tiek apvienota gan difonu, gan citu runas vienību konkatenācija. Sintezējot runu, tiek izmantoti garākie pieejamie fragmenti – atsevišķi vārdi, zilbes vai, ja tādi nav pieejami, tad difoni [11]. Vienību izvēles runas sintēze spēj radīt dabisku runu, bet tai ir nepieciešams liels ierakstītas runas korpus (runas datubāze ar desmitiem tūkstošu vienībām), kurā var būt marķētas atsevišķas fonēmas, vārdi, frāzes un teikumi (patentēts realizācijas piemērs ar šīm īpašībām ir *Google* konkatenācijas runas sintezators [8]). Papildus runas korpusā tiek atzīmētas tā vienumu īpašības, piemēram, tonis, ilgums, konteksts. Ierakstīto skaņu savienošānu un parametrizēšanas realizācija sniedz labus rezultātus un līdzīgi kā difonu runas sintēze pielietojama universālai runas sintēzei, ja tās datubāzē ir pieejami atbilstoši audio resursi.

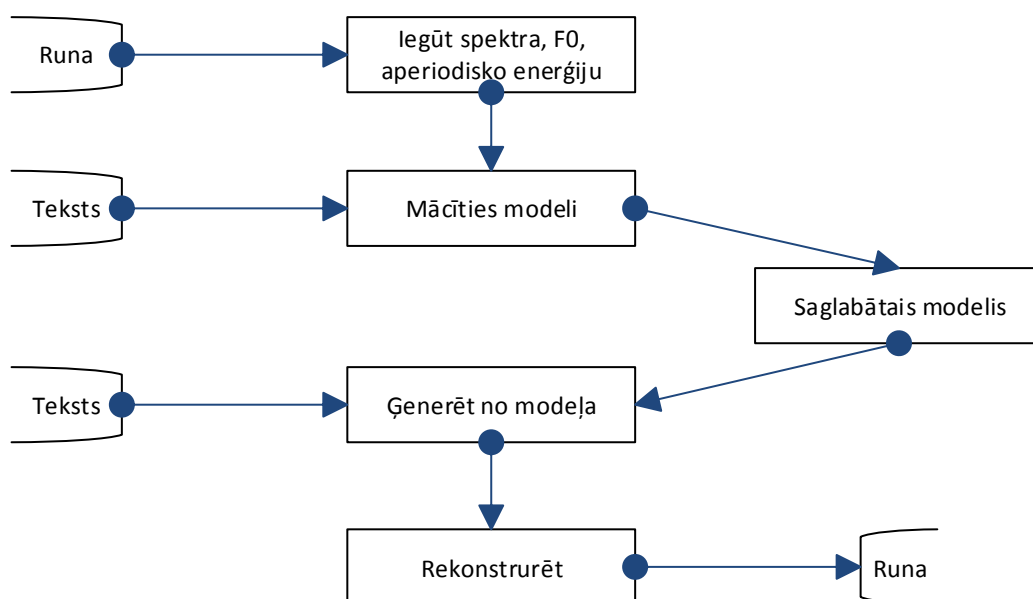
2.3. Artikulārā sintēze

Artikulārā runas sintēze izstrādāta, izmantojot cilvēka balsis trakta darbības modeli. Metode sasaucas ar vēsturiskajiem runas sintēzes mēģinājumiem mehāniski modelēt cilvēka runu (skat. nod. "1.3 Vēsture"), tomēr šīs metodes realizācija ir elektroniska. Risinājums ir izstrādāts pirms vairāk nekā 10 gadiem, pieeja vairs netiek attīstīta mūsdienu pētniecības projektos, jo sniedz salīdzinoši zemus rezultātus. Risinājums ir pieejams *gnuspeech* bibliotēkā [16].

2.4. Parametriskā runas sintēze

Parametriskā runas sintēzes metode (parametriskās runas sintēzes darbības shēmu skat. 4. attēlā.) runu modelē kā vairāku faktoru mijiedarbības rezultātu. Tā tiek izmantota runas modelēšanas solī (skat. nod. "2.1. Runas sintēzes process"). Kā faktorus tipiski izmanto tembru (spektra ietvaru – *spectral envelope*), balsi (pamata frekvenci – *fundamental frequency*), toni (balsis frekvenci – *voice frequency*), fona skaņas (aperiodisko enerģiju – *aperiodic energy*), kā arī ilgumu [11].

4. attēls – "Parametriskās runas sintēzes darbība [11]"



2.4.1. Slēpto Markova modeļu izmantošana runas sintēzē

Lai no iepriekš iegūtajām parametru vērtībām un dotā teksta ģenerētu skaņas vilni, visbiežāk tiek izmantots. Runas ģenerēšanas procesā no modeļa stāvokļiem tiek nolasīti un reprezentēti skaņas viļņa varianti, kas pēc tam tiek pielāgoti, izmantojot modeļa parametrus. Modeļa apmācībai tipiski ir nepieciešamas 2 līdz 5 stundas audio materiāla vienas balsis sintēzei. Apmācībai vajadzīgi viena cilvēka ierunāti labas kvalitātes (bez fona trokšņiem un

traucējumiem) dati. Savas kvalitātes un pielāgojamības dēļ statistiskā parametriskā sintēze ir kļuvusi populārākā no metodēm.

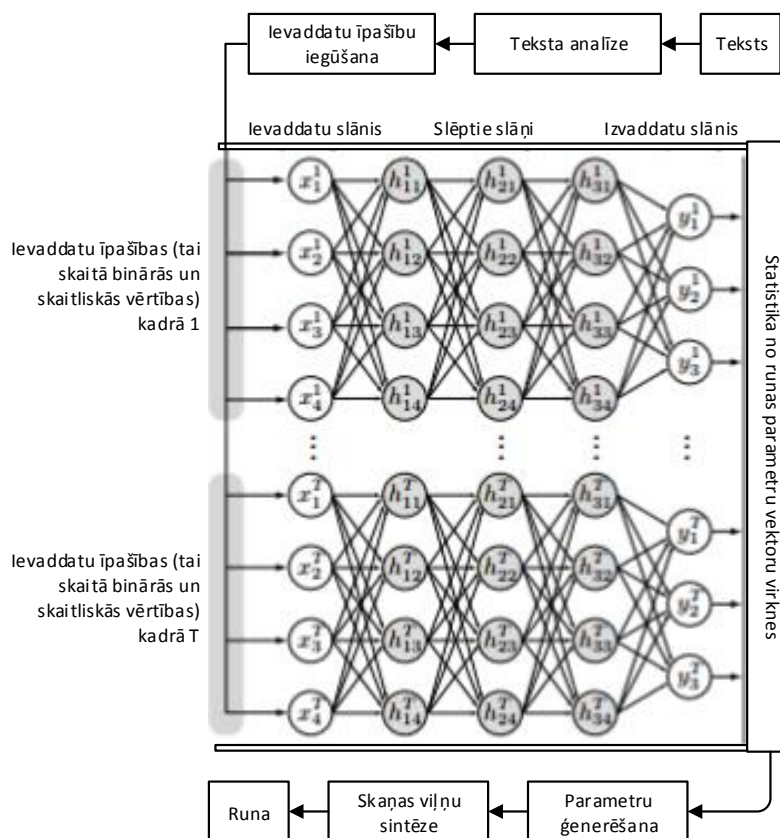
2.4.2. Dziļo neironu tīklu metodes runas sintēzē [20]

Slēpto Markova modeli var aizstāt vai papildināt ar dziļo neironu tīklu modeli [12], tādējādi iegūstot sintēzes sistēmu, kas praksē sniedz labākus rezultātus. Runas sintēzes ietvars, kas balstīts uz dziļajiem neironu tīkliem, ir izstrādāts, pamatojoties uz ideju, ka, “mainot informāciju no lingvistiskā līmeņa uz skaņas viļņiem, cilvēka runai ir slāņaina hierarhiskā struktūra” [20].

5. attēlā redzams runas sintēzes ietvars, kas izstrādāts, izmantojot dziļos neironu tīklus. Dotais teksts sākotnēji tiek pārvērsts ievaddatu pazīmju virknē ($f x t n g$), kur $x t n$ nozīmē n -to īpašību kadrā t . Ievaddatu īpašības iekļauj bināras atbildes uz jautājumiem par lingvistisko kontekstu (piemēram, vai konkrētā fonēma ir ‘aa’?) un skaitliskas vērtības (piemēram, vārdu skaits frāzē, relatīvā pozīcija konkrētajā kadrā konkrētajā fonēmā un konkrētās fonēmas ilgums). Tad ievaddatu īpašības tiek kartētas ar izvaddatu īpašībām $f y t m g$ ar iepriekš apmācītu dziļo neironu tīklu.

Neironu tīkla trenēšanā tiek izmantotas no apmācības kopas iegūtās izvaddatu un ievaddatu īpašības. Līdzīgi kā slēptā Markova modeļa pieejā ir iespējams ģenerēt runas parametrus no iepriekš uz apmācītās kopas. Runas parametru ģenerācijas algoritms var ģenerēt gan statiskas, gan dinamiskas īpašības. Kā pēdējais solis ir skaņas viļņa sintēze atbilstoši dotajiem skaņas runas parametriem. Teksta analīzes, runas parametru ģenerēšanas un skaņas viļņu sintēzes soļi var tikt pildīti, apvienojot slēptā Markova modeļa un dziļo neironu tīkla pieeju.

5. attēls – “Runas sintēzes sistēma, izmantojot dziļos neironu tīklus [20]”



2.5. Metožu salīdzinājums

Konkatenācijas runas sintēzes priekšrocības ir augsta kvalitāte, bet, lai to sasniegtu, ir nepieciešama ļoti liela datu kopa [11], kā arī tā prasa daudz procesora jaudas pie lielajiem apjomiem. Savukārt parametriskā sintēze paļaujas uz mašīnmācīšanos no iepriekš ierakstītas un transkribētas runas. Parametriskā runas sintēze sniedz labākus rezultātus pie nelieliem datu apjomiem (zem stundas). Tā paļaujas uz statistiski iegūtajiem datiem par apmācības datu kopu. Tā pretēji konkatenācijai ir vieglāk pielāgojama dažādām balsu īpašībām. Praktiskajā darba daļā tiek izmantota parametriskā sintēze.

3. RUNAS SINTĒZES RĪKI

3.1. Pasaulē izmantotie runas sintezatori

Konkatenatīvā sintēzes metode un parametriskā sintēzes metode atrodamas virknē rīku. Tomēr ir pielietotas arī citas metodes, kā populārākie runas sintezatori ir:

1. *Festival* – Parametriskā sintēze;
2. *eSpeak* – Konkatenatīvā sintēze;
3. *Ivona TTS* – Vienību izvēles metode;
4. *Cepral* – Parametriskā sintēze (izmanto *Festival*);
5. *Ekho* - Vienību izvēles metode (izmanto *Ivona TTS*);
6. *Praat* - Vienību izvēles metode.

Lielākā daļa no šiem rīkiem (3.-6.) ir maksas pakalpojums vai vienreizēja iegādes programmatūra. Vairāki no rīku izplatītājiem piedāvā pielāgotu runas sintēzes pakalpojumu tiešsaistē kā mākoņpakalpojumu. Kā viens labākajiem sintezatoriem šobrīd tiek uzskatīts *Ivona TTS*⁶, taču no atvērtajiem ietvariem vispopulārākais ir *Festival*⁷ ietvars, uz kura pamata izstrādāti virkne sintezatoru dažādām valodām.

3.2. Latviešu valodas runas sintezatori

Latviešu valodas sintēzei šobrīd pieejamo rīku skaits ir trīs, tie visi izstrādāti, izmantojot konkatenatīvo runas sintēzi.

3.2.1. *eSpeak*

eSpeak ir atvērtā koda runas sintezators (pieejams *OSX*, *Linux* un *Windows* operētājsistēmām) angļu, latviešu un daudzām citām valodām. Rīkā tiek izmantota parametru runas sintēze ar iespēju lietotāja saskarnē mainīt runas ātrumu, tembru un citas īpašības [17].

3.2.2. *Runa.ailab.lv*

Konkatenatīvā aizvērta koda bezmaksas latviešu valodas runas sintēzes sistēma *T2S VI* izstrādāta Latvijas Universitātes Matemātikas un informātikas institūta Mākslīgā intelekta laboratorijā [19] [27]. Sintezators ir pieejams tiešsaistē adresē <http://runa.ailab.lv>.

⁶ <https://www.ivona.com/> [aplūkots 21.01.2016]

⁷ <http://www.cstr.ed.ac.uk/projects/festival/> [aplūkots 21.01.2016]

3.2.3. Tildes Biroja runas sintezators

Tildes Biroja sastāvā ir latviešu teksta izrunāšanas robots Tildes Birojs *Text-to-Speech*. Tas darbojas kā aizvērtā koda programmatūra, kas ir nopērkama kopā ar Tildes Biroju. Sintezators ir pielietojams Tildes Birojā, kā arī citu izstrādātāju programmās, piemēram, *Microsoft Excel*, *ZoomText*, *Jaws for Windows* [25] [26].

3.3. Citi rīki

Populārākais brīvi pieejamais rīks runas transkribēšanai ir *Transcriber*⁸. Tā galvenā funkcija ir ērta teksta rediģēšana un sasaiste ar audio failu. Papildus tiek nodrošināta arī runātāju segmentēšana un citu meta datu norādīšana.

⁸ <http://trans.sourceforge.net/en/presentation.php> [aplūkots 21.05.2016]

4. LATVIEŠU VALODAS RUNAS SINTĒZES RĪKS “RUNĀTĀJS”

Ņemot vērā ierobežoto latviešu valodas runas sintēzes rīku pieejamību, darba ietvaros izstrādāts jauns latviešu valodas runas sintēzes rīks. Autore rīku nosaukusi par *Runātājs*. Izstrādes ideja ir pārbaudīt, vai, izmantojot ar mūsdienu runas sintēzes metodes nelielam datu apjomam, ir iespējams sasniegt tikpat labus vai labākus rezultātus kā esošajiem latviešu valodas runas sintezatoriem. Izstrādei izvēlēta parametriskā runas sintēze.

4.1. Uzbūve un darbības principi

Rīks *Runātājs* darbojas atbilstoši nod. “2.1. Runas sintēzes process” aprakstītajam runas sintēzes procesam un parametriskās sintēzes metodes aprakstam (skat. nod. “2.4. Parametriskā runas sintēze”). Rīka apmācība īstenota atbilstoši *Ossian* dokumentācijai⁹. Rīka versijas (skat. nod. “4.3. Versiju apraksts”) izvērtētas atbilstoši darba eksperimentālo novērtējumu pieejai.

4.1.1. Eksperimentālās vides parametri

Linux Ubuntu 14.04.4

Virtuālā mašīna

Atmiņa 7GB

4 procesori Intel(R) Xeon(R) CPU E5-2673 v3 @ 2.40GHz

4.1.2. Izmantotie rīki

Ossian

Versija 1.3, <http://homepages.inf.ed.ac.uk/owatts/ossian/html/>

Ossian ir Python programmēšanas valodā izstrādāts kods runas sintezatoru izstrādei.

HTS - Hidden Markov Model (HMM)-based Speech Synthesis System

Versija 2.3, <http://hts.sp.nitech.ac.jp/>

HTK - Hidden Markov Model (HMM) Toolkit

Versija 3.4.1, <http://htk.eng.cam.ac.uk/>

C bibliotēka un rīki, pieejami kā C pirmkods

hts_engine

Programmatūra runas viļņu sintēzei no Slēptajiem Markova modeļiem, kas apmācīti ar HMM-based speech synthesis system (HTS).

⁹ <http://homepages.inf.ed.ac.uk/owatts/ossian/html/index.html> [aplūkots 21.05.2016]

Versija 1.10, <http://hts-engine.sourceforge.net/>

C bibliotēka

Festival - The Festival Speech Synthesis System

Versija 2.4, <http://www.cstr.ed.ac.uk/projects/festival/>

C++ pirmkods

Festival piedāvā vispārēju ietvaru (tai skaitā lietojumprogrammu saskarnes, bibliotēkas) runas sintēzes sistēmu. Ietvars iekļauj arī HTS parametriskās runas sintēzes dzini.

GlottHMM vocoder

Fizioloģiski orientēta Slēpto Markova modeļu bāzētu runas sintēzes sistēma.

Versija 1.0.15, iekļauta komplektā ar Ossian

<http://www.helsinki.fi/speechsciences/synthesis/glott.html>

SPTK - The Speech Signal Processing Toolkit (SPTK) ir signālu apstrādes rīkkopa UNIX vidēm *Versija 3.6, <http://sp-tk.sourceforge.net/>*

4.1.3. Sintezatora lietotāja saskarne

Lietotāja saskarne ir publiski pieejama mājas lapa adresē <http://40.69.82.98/>. Tiešsaistē pieejamā *Runātājs* versija ir *Laika_zinas_2* (skat. nod. "4.2. Versiju apraksts"). *Runātājs* lietotāja saskarnē ir iespēja ievadīt savu tekstu (skat. 6. attēlu), to sintezēt un klausīties sintezētos rezultātus ar iebūvēto tīmekļa pārlūka atskaņotāju.

6. attēls – “Runātājs lietotāja saskarne”

Laipni aicināti runātāja mājas lapā!

Raksti teksta logā tekstu un spied pogu "Runāt"!

Runāt

[Rezultāti](#)

Darbības ierobežojumi

- Runātājs darbojas eksperimentālā režīmā.
- Lietojumprogrammā ir izmantots:
 - Linux Ubuntu 14.04.4
 - Ossian 1.3
 - HTS 3.4.1
 - htc_engine 1.10
 - GlottHMM vocoder 1.0.15
 - SPTK 3.6

Izstrādāts studiju darba ietvaros 2016. gada maijā.

Runātāja versija: Laika_zinas_2

Runātājs ir teicis:

Nosaukums	Lejupielādēt
2016-05-22-23-09-58-Laika_zinas_2.wav	Datne
2016-05-22-23-08-34-balss.wav	Datne
2016-05-22-23-02-40-sp1.wav	Datne
2016-05-22-22-59-43-sp1.wav	Datne
2016-05-22-22-46-34-sp1.wav	Datne
2016-05-22-22-44-59-Laika_zinas_2.wav	Datne
2016-05-22-17-11-22-sp1.wav	Datne
2016-05-22-13-00-54-sp1.wav	Datne
2016-05-21-16-53-54-runu.wav	Datne
2016-05-21-16-10-14-runu.wav	Datne
2016-05-21-16-09-58-runu.wav	Datne
2016-05-21-16-09-19-runu.wav	Datne

Izstrādāts studiju darba ietvaros 2016. gada maijā.

Runātāja versija: Laika_zinas_2

4.2. Izmantotie latviešu valodas runas resursi

Rīka pielāgošanā ir izmantoti dažādi runas resursi, tā darbības pārbaudei pie dažādām balsīm un apjomiem. Runas resursi pielāgoti sintezatora vajadzībām, izmantojot *Transcriber* (skat. nod. “3.3. Citi rīki”) rīku, teksta redaktorus un apstrādes rīkus. Iegūtie runas resursi ir audio faili un atbilstošie teksti, kas izmantoti rīka apmācībā.

4.2.1. Latviešu valodas runas korpuss

Latviešu valodas runas korpuss ir 100 stundu ortogrāfiski marķēts runas audio datu un 4 stundas fonētiski marķēta runas audio datu apkopojums. Korpuss ir līdzsvarots un fonētiski bagāts. Izstrāde veikta atbilstoši runas datu kvalitātes un proporcionālā sadalījuma kritērijiem. Lai korpuss reprezentētu dabisku runu, izvēlēti dažādi runātāji (pēc vecuma, dzimuma, profesijas) [2].

4.2.2. Pasakas

Pasakas runas resurss ir nekomerciāliem mērķiem izmantojamas audiogrāmatas, kas satur dažādu autoru priekšā lasītas pasakas bērniem. Pasaku teksts un atbilstošais audio lasījums ir pieejams lejupielādei tiešsaistes vietnē pasakas.lv [32]. Pasakas ierunājuši dažādi autori, izstrādē izmantotas šādas pasakas un audiogrāmatu lasījumi:

1. “Bezbārdis¹⁰”, teicējs: Zane Burnicka;
2. “Runcis Maksims¹¹”, teicējs: Zane Burnicka.

4.2.3. Laika ziņas

Sintezatora testēšanai izmantoti runas resursi no *T2S VI Runa.ailab.lv* (skat. nod. “3.2.2. Runa.ailab.lv”). Valodas resurss sastāv no vārdiem un plašāk izmantotajām frāzēm laika ziņu prognozēs [27]. Dati izmantoti kā runas fragmenti Runātāja apmācīšanai.

4.3. Versiju apraksts

Runātājs izvērtēšanā izmantotas vairākas rīka izstrādes versijas. Rīka versiju uzskaitījums ir redzams 1. tabulā. Versijas maiņu nosaka rīka apmācība ar jaunu datu kopu. Pēc testēšanas rezultātiem (skat. nod. “6.2. *Runātājs* sintezatora versiju salīdzinājums”) labākā runas sintezatora versija ir tiešsaistē pieejamā “Laika_zinas_2”. Tā ir izmantota *Runātājs* salīdzināšanā ar citiem runas sintēzes rīkiem.

1. tabula – “Runātājs versijas”

<i>Versija</i>	<i>Apraksts</i>
<i>Laika_zinas_1</i>	Rīka versijas apmācību datu kopa ir atbilstoša valodas resursam “Laika ziņas” (skat. nod. “4.2.3. Laika ziņas”). Resurss ir izmantots kā izolēti vārdi. <i>Kopējais datu apjoms: 35 minūtes 36 sekundes</i> <i>Vidējais viena faila garums: 1,05 sekundes</i>
<i>Laika_zinas_2</i>	Rīka versijas apmācību datu kopa ir atbilstoša valodas resursam “Laika ziņas” (skat. nod. “4.2.3. Laika ziņas”). Resursa iegūšanai tikuši transkribēti resursā pieejamie runas segmenti (frāzes). <i>Kopējais datu apjoms: 8 minūtes 53 sekundes</i> <i>Vidējais viena faila garums: 4,68 sekundes</i>

¹⁰ http://www.pasakas.net/pasakas/citu_tautu_pasakas/ciganu-pasakas/bezbardis/ [aplūkots 22.05.2016.]

¹¹ http://www.pasakas.net/pasakas/citu_tautu_pasakas/baltkrievu-pasakas/runcis-maksims/ [aplūkots 22.05.2016.]

<i>Versija</i>	<i>Apraksts</i>
<i>Pasakas_1</i>	<p>Rīka versijas apmācību datu kopa ir transkribēta pasaka “Runcis Maksims” (skat. nod. “4.2.2. Pasakas”).</p> <p><i>Kopējais datu apjoms:</i> 17 minūtes 14 sekundes <i>Vidējais viena faila garums:</i> 4,38 sekundes</p>
<i>Pasakas_2</i>	<p>Rīka versijas apmācību datu kopa ir transkribētas pasaka “Runcis Maksims” apvienojumā ar transkribētu pasaku “Bezbārdis” (skat. nod. “4.2.2. Pasakas”).</p> <p><i>Kopējais datu apjoms:</i> 27 minūtes 28 sekundes <i>Vidējais viena faila garums:</i> 4,97 sekundes</p>
<i>Saeima</i>	<p>Rīka versijas apmācību datu kopa ir no Saeimas korpusa (skat. nod. “4.2.1. Latviešu valodas runas korpus”) viena runātāja (S. Āboltiņa) visi ieraksti un transkribētais teksts.</p> <p><i>Kopējais datu apjoms:</i> 42 minūtes <i>Vidējais viena faila garums:</i> 3,41 sekunde</p>

5. EKSPERIMENTĀLIE NOVĒRTĒJUMI

5.1. Eksperimentu konteksts

Runas sintēzes izvērtējums tiek veikts, pārbaudot runas sintēzes sistēmas rezultātu – sintezētās runas atbilstību cilvēka runai. Kvalitāte mērīta atbilstoši cilvēku izvērtējumam. Ideāla runas sintēzes sistēma izpildītu Tjūringa testu – izrunas kvalitāte nebūtu atšķirama no cilvēka runas kvalitātes vai pārspētu to [3].

Literatūrā apskatīti arī citi varianti runas sintēzes izvērtēšanai, tomēr pēdējo gadu konferenču rakstos tiek izmantota tieši subjektīvu vērtējumu pieeja [18], dažādojot testētāju tipus, testa datu kopas un vērtēšanas kritērijus. Šāda pieeja valodas tehnoloģijās tiek izvēlēta arī, piemēram, mašīntulkošanā, emocionālās noskaņas analizē un citur, ar pieņēmumu, ka mašīnmācīšanās rezultātā apmācīto modeļu sistemātiskais izvērtējums nereprezentē cilvēka izvērtējumam līdzvērtīgu – nav cita zelta standarta kā cilvēka dabiskā valoda vai uztvere.

5.2. Stratēģija un kritēriji

Darbā analizēti visi latviešu valodai paredzētie runas sintezatori – *Runātājs*, *eSpeak*, *runa.ailab.lv* un *Tildes Birojs* runas sintezators. Testēšanas stratēģija ir izvērtēt sintezētās runas kvalitāti, tas darīts turpmāk apskatītajos veidos.

Rīku testēšanā izvērtēts sintezētās runas dabiskums. Dabiskuma kritērijs ir subjektīvs, tāpēc tā izvērtēšanai parasti tiek izvēlēta liela testētāju kopa, kas reprezentē iedzīvotāju kopu. Veicot pirmreizējo autores subjektīvo novērtējumu, secināts, ka sajaukt nevienu no četriem sintezatoriem ar cilvēka balsi nevar, tāpēc kritērijam piemērota 5 punktu sistēma ar mērķi starp četriem rīkiem atrast labāko ar nelielu testētāju skaitu. Testētāji dzirdētajiem fragmentiem izlika vērtējumu pēc turpmāk lasāmās subjektīvās skalas:

5 – nevar pamanīt, ka nerunā cilvēks;

4 – atsevišķās vietās var pamanīt metālisku skaņu, bet izklausās pēc dabiskas runas;

3 – bieži var pamanīt metālisku, nedabisku skaņu;

2 – nedabiska skaņa, būtiski traucē, ar šādu sintezatoru būtu ļoti grūti sarunāties;

1 – nedabiska skaņa, pēc noklausīšanās džinkst ausīs, vairs negribu dzirdēt šādu sintezatoru.

Testpiemēri izstrādāti atbilstoši dabiskai cilvēka runai, kas pierakstīta sadzīviskos apstākļos. Lai nodrošinātu iespēju rīkus savstarpēji salīdzināt, katram rīkam izvēlēti vieni un tie paši testpiemēri.

Teksta pirmsapstrādes kvalitātes testēšanai izvēlēts kļūdu skaita kritērijs. Testētāji atzīmē, cik kļūdas pamanītas, klausoties sintezētos tekstus. Visos testpiemēros iekļauti speciālie simboli ‘%’, ‘”’’, ‘,’’, skaitļi, gada skaitļi un decimālskaitļi, kā arī iekļauti īpašvārdi, lai pārbaudītu, ka rīki spēj sintezēt vārdus, kas nav atrodamā latviešu valodas vārdnīcā.

Teksta saprotamībai izvēlēts nedabiskas runas paraugs, izvērtējot, vai tekstu ir iespējams saprast no dzirdētā. Izvērtējums tika izdarīts pēc šādas skalas:

- 4 – nav jāpiepūlas saprast;
- 3 – var saprast koncentrējoties;
- 2 – grūti saprast, pat koncentrējoties;
- 1 – gandrīz nevar saprast, var izšķirt atsevišķus vārdus.

5.3. Organizācija

Eksperimentālie novērtējumi organizēti divās daļās. Vispirms izvērtēts un uzlabots rīks *Runātājs*, tad *Runātāja* labākā versija testējot salīdzināta ar citiem runas sintēzes rīkiem.

Izstrādāto testpiemēru saturs atrodams 1.-4. pielikumā. To saturs sintezēts un saglabāts skaņas failos, kas izmantoti testēšanā.

Testēšanā tika iesaistītas 4 dažādas personas. Katrs testētājs aizpildījis testēšanas protokolu par katru no testpiemēriem atbilstoši to kategorijām un iesniedzis to apkopošanai.

6. REZULTĀTI UN DISKUSIJA

6.1. Latviešu valodas runas sintezatoru salīdzinājums

2. tabulā redzams testēšanas rezultātu apkopojums par runas sintezatoru salīdzināšanu ar darbā izstrādāto rīku *Runātājs*. Tabulā ir redzami summārie rezultāti par 22 izpildītajiem testiem (pilnu testa rezultātu kopu skatīt 1.- 4. pielikumā). Dabiskuma un saprotamības testiem lielāki rezultāti norāda uz augstāku kopējo vērtējumu par sintezētajiem runas paraugiem. Teksta apstrādes rezultāts norāda uz pamanīto kļūdu daudzumu sintezētajā runā (jo mazāks skaits, jo labāk).

Testēšanas rezultāti norāda, ka visdabiskākais un saprotamākais skanējums ir rīkam *Tildes Biroja* runas sintezators. Arī teksta priekšapstrāde viskvalitatīvāk tiek veikta *Tildes Biroja* runas sintezatoram. Sintezatoram *Runātājs* versijā *Laika_zinas_2* priekšapstrādē netiek veikta skaitļu pārveide, tāpēc apstrādāto kļūdu daudzums nav salīdzināms ar citiem sintezatoriem.

2. tabula – “Runas sintēzes rīku salīdzinājums”

<i>Tests</i>	<i>eSpeak</i>	<i>Tildes Biroja runas sintezators</i>	<i>Runa.ailab.lv</i>	<i>Runātājs (Laika_zinas_2)</i>
<i>Dabiskums</i>	25	42.5	15.5	36.6
<i>Saprotamība</i>	7	14	6	13
<i>Teksta apstrāde (pieļauto kļūdu skaits)</i>	56	34	57	-

Sintezatoru savstarpējais izvērtējums kā labāko latviešu valodas runas sintezatoru norāda konkatēnā sintēzes rīku. Tomēr saprotamības ziņā līdzvērtīgs ir arī *Runātājs*. Rezultātu tuvums norāda par iespējām pēc *Runātājs* uzlabošanas uzvarētāju apsteigt. Darba ietvaros izstrādātais *Runātājs* ir tapis ar salīdzinoši maz valodas resursiem (8 minūtes 53 sekundes transkribētas runas), sasniegtie rezultāti liecina par parametriskās runas sintēzes pārsvaru pie maza datu apjoma pār konkatēnā sintēzi.

6.2. *Runātājs* sintezatora versiju salīdzinājums

3. tabulā ir redzams testēšanas rezultātu apkopojums par *Runātājs* versiju testēšanu (versiju aprakstu skat. nod. “4.3. Versiju apraksts”). Tabulā ir redzami summārie rezultāti par 10 izpildītajiem testiem (pilnu testa rezultātu kopu skatīt A. - C. pielikumā). Dabiskuma un

saprotamības testiem lielāki rezultāti norāda uz augstāku kopējo vērtējumu par sintezētajiem runas paraugiem.

Testēšanas rezultāti norāda, ka visdabiskākais un saprotamākais skanējums rīkam *Runātājs* ir ar versiju *Laika_zinas_2*.

3. tabula – “*Runātāja* versiju salīdzinājums”

<i>Tests</i>	<i>Laika_zinas_1</i>	<i>Laika_zinas_2</i>	<i>Pasakas_1</i>	<i>Pasakas_2</i>	<i>Saeima</i>
<i>Dabiskums</i>	9	10.5	4.5	7	4.5
<i>Saprotamība</i>	9.2	9.8	6.7	8.1	8
<i>Summa</i>	18.2	20.3	11.2	15.1	12.5

Ar sākotnējo *Runātāja* versiju *Laika_zinas_1* iegūtā runas sintēze tika vērtēta kā saraustīta. Atbilstoši testētāju komentāriem tika izdarīts pieņēmums, ka trūkst datu kvalitatīva paužu modeļa apmācībai, jo apmācības datus katrā failā bija tikai viens vārds. Rezultāti tika saglabāti un apmācīta jauna rīka versija, izmantojot garākus runas segmentus.

Nākamā versija tika apmācīta uz Saeimas datiem (versija *Saeima*). Iegūtā runa bija daudz plūstošāka kā sākotnējā runa, bet audio kvalitāte bija daudz zemāka. Audio dzirdamie trokšņi norādīja uz zemāku apmācības datu ierakstu kvalitāti. Vairākos segmentos runa nebija arī skaidri saprotama tās ātruma dēļ. Apmācības dati tika novērtēti kā zemākas kvalitātes, jo runātāja nodoms nebija runāt tā, lai ierakstus varētu izmantot runas sintēzē.

Tika izvirzīta hipotēze, ka labākus rezultātus varētu sasniegt ar ierakstiem, kuros runātājs ir ar nodomu centies skaidri ierunāt tekstu. Tika izvēlēta audio pasaka (versija *Pasakas_1*) ar pieejamu rakstītu tekstu, kas atviegloja un paātrināja runas transkribēšanas procesu. Runas temps bija ievērojami labāks, tomēr fonēmu izrunas kvalitāte bija ļoti mainīga, salīdzinot ar Saeimas datiem. Tika pieņemts, ka tas ir mazā datu apjoma dēļ, tāpēc nākamais solis bija palielināt datu apjomu. Iegūtais rezultāts ar versiju *Pasakas_2* bija labāks, bet tas nesniedza pārliecību, ka ar šādā veidā iegūtiem datiem būs iespējams sasniegt tādu pat runas skaidrumu kā ar sākotnējo *Laika_zinas_1* versiju.

Lai salīdzinātu, kādas ir atšķirības starp modeli, kas apmācīts uz pasaku datiem un modeli, kas apmācīts uz līdzvērtīga apjoma un transkripciju veida (frāžu garuma) datiem, no laika ziņām tika no jauna notranskribēti oriģinālie diktora ieraksti skaņu studijā (dati oriģināli izmantoti AIlab konkatēnātajai runas sintēze – skat. nod. “3.2.2. Runa.ailab.lv”). No jauna tika apmācīts rīks un iegūta versija *Laika_zinas_2*. Uz datu apjoma, kas ir līdzīgs versijā *Pasakas_1*, iegūtais modelis bija labāks par *Pasakas_2* versijā iegūto modeli, kas bija apmācīts uz lielāka datu apjoma. No tā tika secināts, ka vissvarīgākais ir ierakstu kvalitāte. Pat pasakām, kas bija speciāli ielasītās kā audio grāmata ļoti labākā kvalitātē, runas sintēzes

kvalitāte bija sliktāka pat uz lielāka datu apjoma. Modeļa apmācība uz garākām frāzēm, nevis uz izolētiem vārdiem ļoti uzlabo sintezētās runas plūdumu. Pat ar mazāku treniņu datu apjomu, sintezētā runa izklausās dabiskāk un ir labāk uztverama.

7. SECINĀJUMI

1. Pēdējos piecos gados trūkst pētījumu par latviešu valodas runas sintēzes tehnoloģijām. Pasaules kontekstā ik gadu runas tehnoloģijas būtiski mainās, savukārt pētījumu trūkums latviešu valodai norāda uz risinājumu attīstības atšķirību.
2. Runas sintēzei latviešu valodā trūkst komerciāla izmantojuma, bet ir vairākas sabiedriski nozīmīgas lietojuma iespējas, it īpaši izglītībā un veselības aprūpē.
3. Literatūras izpēte liecina, ka pēdējā gada laikā ir pieejamas jaunas metodes runas sintēzei, kas nav izmantotas latviešu valodas runas sintēzes rīkos.
4. Parametriskā sintēze ir galvenā izmantotā metode mūsdienu runas sintēzes rīku darbplūsmās. Tās priekšrocības ir mazāks nepieciešamais valodas resursu apjoms (tai skaitā ekspertu līdzdalība šo resursu sagatavošanā).
5. Parametriskajā sintēzē ir nepieciešams izmantot viena cilvēka ierunātu runas resursu. Svarīgākā izmantoto runas resursu īpašība ir tā ierakstu kvalitāte.
6. Latviešu valodas runas sintēzei, darbu sākot, pieejami 3 risinājumi, no kuriem tikai viens (*eSpeak*) ir atvērtā koda risinājums, kas pieejams turpmākai attīstīšanai.
7. Ir pieejams plašs klāsts ar aizvērtā koda risinājumiem runas sintēzē, no kuriem daļa tiek piedāvāti kā pakalpojumi lietotājiem par maksu. Latviešu valodas runas sintēzes rīki (*Tildes birojs* runas sintezators un *runa.ailab.lv*) ir aizvērtā koda programmatūra, *Tildes Birojs* runas sintezators ir maksas rīks.
8. Esošie latviešu valodas runas sintēzes risinājumi sintēzē saprotamu, kaut gan, subjektīvi vērtējot, nedabisku runu.
9. Esošie latviešu valodas risinājumiem (*eSpeak* un *Tildes Biroja* sintezators) tiek veikta teksta priekšapstrāde (normalizācija), tomēr, izrunājot gada skaitļus un decimālskaitļus, novērotas kļūdas, kas būtu labojamas esošajos risinājumos, principāli nemainot to pamata algoritmus.
10. Ir iespējams uztaisīt latviešu valodas runas sintēzes rīku ar atvērtiem bezmaksas rīkiem, iegūstot atbilstošus latviešu valodas runas resursus.
11. Atkārtoti izmantojot runas resursus, kas izmantoti iepriekš izstrādātajos latviešu valodas runas sintēzes rīkos, tiek iegūti labāki rezultāti, ja tiek izmantota parametriskā sintēze iepriekš izmantotās konkatēnatīvās vietās.
12. Modeļa apmācība uz garākām frāzēm, nevis uz izolētiem vārdiem ļoti uzlabo sintezētās runas plūdumu.

13. Atbilstoši latviešu valodas runas sintēzes rīku testētāju atsauksmēm, ilgstoši klausoties latviešu valodas runas sintezatorus, rodas nepatīkami blakusefekti, tādi kā džinkstoņa ausīs, nelielas galvassāpes.
14. Testēšanas rezultāti norāda, ka labākais šī brīža latviešu valodas universālās runas sintezators ir Tildes Biroja runas sintezators. Darba ietvaros izdevies radīt risinājumu, kas pārspēj citus latviešu valodas runas sintēzes rīkus.
15. Testēšanas ietvaros atrastas iespējas latviešu valodas runas sintēzes rīku uzlabošanai. Kā uzlabojumu iespējas *Runātājs* darbībā ir:
 - a. Teksta priekšapstrāde;
 - b. Fonētiskās vārdnīcas pilnveide ar mērķi uzlabot vārdu izrunu;
 - c. Paužu noteikšanas uzlabošana un vārdu dalīšana zilbēs ar mērķi iegūt labāku runas plūdumu;
 - d. Intonāciju modelēšanā ar mērķi iegūt runas dabīgāku skanējumu.

8. NOBEIGUMS

Darba ietvaros ir izpētītas šobrīd aktuālās metodes un rīki runas sintēzei pasaulē un runas sintēzē latviešu valodai. Pēc jaunāko metožu un rīku apzināšanas darba ietvaros tika izstrādāts jauns rīks *Runātājs* latviešu valodas runas sintēzei.

Darba izstrādei tika izvirzīti un izpildīti turpmāk norādītie uzdevumi.

1. Veikt teorētisko apskatu mūsdienas runas sintēzes rīkos izmantotajām metodēm.
2. Eksperimentāli izvērtēt esošo latviešu valodas runas sintēzes rīku kvalitāti.
3. Izstrādāt latviešu valodas runas sintēzes rīku un eksperimentāli izvērtēt tā darbību.

Izpildot darba uzdevumus, tika apstiprināta sākotnēji izteiktā hipotēze, ka esošos latviešu valodas runas sintezators iespējams pārspēt, izmantojot pēdējo gadu laikā izstrādātās metodes un rīkus. Galvenais darba rezultāts ir izstrādāts latviešu valodas runas sintēzes rīks, kas ir brīvi izmantojams un pielietojams visām vajadzībām.

Darba izstrādes laikā tika izdarīti šādi četri galvenie secinājumi:

1. Esošajā situācijā latviešu valodai netiek nodrošināts pilnvērtīgs datoratbalsts runas sintēzei, tas netiek uzturēts un attīstīts reizē ar jauno tehnoloģiju attīstību. Latviešu valodas runas sintēzes rīki to pieejamības un kvalitātes ziņā atpaliek no angļu valodai un citām valodām pieejamajiem runas sintēzes rīkiem, un latviešu valodas runas sintēze pēdējo 5 gadu laikā netiek attīstīta pētnieciskos nolūkos.
2. Runas sintēzei latviešu valodā ir būtiski nekomerciāli pielietojumi konkrētām personu grupām, kā cilvēkiem ar redzes traucējumiem, cilvēkiem, kas slimo ar disleksiju un cilvēkiem, kas kādu iemeslu pēc zaudē balsi.
3. Ir pieejami atvērti un papildināmi risinājumi runas sintezatori izstrādei, kas tiek izmantoti citu valodu runas sintēzes rīku izstrādē ar labiem rezultātiem. Risinājumi ir pielāgojami arī latviešu valodai. Atbilstoši mūsdienu metožu pieejai un pieejamo latviešu valodas resursu apjomam ir pietiekami runas resursi, lai izstrādātu labas kvalitātes latviešu valodas sintezatoru.
4. Darba ietvaros izstrādātais rīks *Runātājs* ir līdzvērtīgs esošajiem latviešu valodas runas sintezatoriem tā sintezētās runas kvalitātē. Ar rīku ir izdevies iegūt labākus rezultātus kā ar iepriekš izstrādātajiem rīkiem, izmantojot tos pašus runas resursus. *Runātājs* ir papildināms nākamajos pētījumos, izmantojot darba aprakstā pieejamo informāciju par rīka uzbūvi un izstrādes procesu atbilstoši darbā izteiktajiem ieteikumiem nākotnes uzlabojumiem.

PATEICĪBAS

Vēlos pateikties Robertam Darģim par atbalstu darba izstrādes laikā un konsultācijām par latviešu valodas runas tehnoloģijām.

IZMANTOTĀ LITERATŪRA UN AVOTI

- [1] *Skaidrojošā vārdnīca*. Pieejams: www.tezaurs.lv [aplūkots 25.01.2016.].
- [2] Pinnis, M., Auziņa, I., & Goba, K. (2014). Designing the Latvian Speech Recognition Corpus. In *Proceedings of the 9th edition of the Language Resources and Evaluation Conference (LREC'14)*. Reykjavik, Iceland: European Language Resources Association (ELRA), 2014 Pieejams: http://www.lrec-conf.org/proceedings/lrec2014/pdf/284_Paper.pdf [aplūkots 25. 01.2016.].
- [3] *Turing test* Pieejams: https://en.wikipedia.org/wiki/Turing_test [aplūkots 25. 01.2016.].
- [4] *Līdzsvarots mūsdienu latviešu valodas korpuss*. Pieejams: <http://www.korpuss.lv/> [atsauce 22.01.2016.].
- [5] Garkāje, G. *Mūsdienu tiešsaistes tekstu transliterācija*. Kursa darbs Latvijas Universitātes Datorikas fakultātē, 2013.
- [7] *Valodniecības pamatterminu skaidrojošā vārdnīca*. Pieejams: termini.lza.lv [aplūkots 25.01.2016.].
- [8] *United States Patent Coorman et. Al. Speech Synthesis Using Concatenation of Speech Waveforms* Pieejams: <http://www.google.com/patents/US7219060> [aplūkots 25. 01.2016.].
- [9] *Speech synthesis* Pieejams: https://en.wikipedia.org/wiki/Speech_synthesis [aplūkots 25. 01.2016.].
- [10] *Latviešu valoda digitālajā laikmetā* Pieejams: <http://www.meta-net.eu/whitepapers/e-book/latvian.pdf> [aplūkots 25.01.2016.].
- [11] *Prof. Simon King - Using Speech Synthesis to give Everyone their own Voice* <https://www.youtube.com/watch?v=xzL-pxcpo-E> [aplūkots 25.01.2016.].
- [12] *DEEP NEURAL NETWORK (DNN) FOR TTS SYNTHESIS* Pieejams: <http://research.microsoft.com/en-us/projects/dnntts/> [aplūkots 25.01.2016.].
- [13] *EECS 225d Guest Lecture: Speech Synthesis* Pieejams: <https://www.icsi.berkeley.edu/icsi/events/2012/04/kim-silverman-talk> [aplūkots 25.01.2016.].
- [14] *Klatt's Last Tapes: A History of Speech Synthesizers* Pieejams: <http://communicationaids.info/history-speech-synthesizers/> [aplūkots 25.01.2016.].
- [15] *Runas sintezators* Pieejams: http://dl.tilde.lv/TildesBirojs/Help/Online/1.0/index.html?runas_sintezators.htm [aplūkots 25. 01.2016.].

- [16] *Gnuspeech* Pieejams: <http://www.gnu.org/software/gnuspeech/> [aplūkots 25. 01.2016.].
- [17] *eSpeak text to speech* Pieejams: <http://espeak.sourceforge.net/> [aplūkots 25. 01.2016.].
- [18] Maria K. Wolters, Karl B. Isaac, Steve Renals *Evaluating Speech Synthesis Intelligibility using Amazon Mechanical Turk* Pieejams: <http://www.cstr.ed.ac.uk/downloads/publications/2010/wolters-ssw2010.pdf> [aplūkots 25. 01.2016.].
- [19] *Par runas sintēzes sistēmu* Pieejams: <http://runa.ailab.lv/tts2/> [aplūkots 25. 01.2016.].
- [20] Heiga Zen, Andrew Senior, Mike Schuster *STATISTICAL PARAMETRIC SPEECH SYNTHESIS USING DEEP NEURAL NETWORKS* <http://static.googleusercontent.com/media/research.google.com/lv//pubs/archive/40837.pdf> [aplūkots 25. 01.2016.].
- [21] *Runas projekts* Pieejams: <http://runa.korpuss.lv/> [aplūkots 25. 01.2016.].
- [22] *Vocoder* Pieejams: <https://en.wikipedia.org/wiki/Vocoder> [aplūkots 25. 01.2016.].
- [23] *Wolfgang von Kempelen's Speaking Machine* Pieejams: https://en.wikipedia.org/wiki/Wolfgang_von_Kempelen%27s_Speaking_Machine [aplūkots 25. 01.2016.].
- [24] *DECTalk* Pieejams: <https://en.wikipedia.org/wiki/DECTalk> [aplūkots 25. 01.2016.].
- [25] *Tildes Birojs* Pieejams: <http://www.tilde.lv/tildes-birojs-2016> [aplūkots 25. 01.2016.].
- [26] K.Goba, A.Vasiļjevs. "Development of Text-To-Speech System for Latvian", Joakim Nivre, Heiki-Jaan Kaalep, Kadri Muischnek and Mare Koit (Eds.) *NODALIDA 2007 Conference Proceedings*, pp. 67-72, Tartu, May 2007. ISBN 978-9985-4-0514-7 1. http://www.valoda.lv/downloadDoc_648/mid_622
- [27] M.Pinnis, I.Auziņa. "Latvian Text-to-Speech Synthesizer", I. Skadiņa, A. Vasiļjevs (Eds.) *Proceedings of the 2010 conference on Human Language Technologies -- The Baltic Perspective: Proceedings of the Fourth International Conference Baltic HLT 2010*, pp. 69-72, IOS Press, Amsterdam, 2014. <http://dl.acm.org/citation.cfm?id=1860937>
- [28] *Dyslexia Basics* /<http://eida.org/dyslexia-basics/> [aplūkots 25. 05.2016.].
- [29] *Skype Translator* <https://www.skype.com/en/features/skype-translator> [aplūkots 25. 05.2016.].
- [30] *Fonētika, skaņu inventārs* <http://lingvo.info/lv/babylon/phonetics> [aplūkots 25. 05.2016.].

[31] I. Auziņa “Latviešu valodas izrunas datormodelēšana” Promocijas darbs filoloģijas doktora grāda iegūšanai valodniecības zinātņu nozares latviešu sinhronās valodniecības apakšnozarē, 2005.

[32] *Pasakas* <http://www.pasakas.net/jaunumi/> [aplūkots 25. 05.2016.].

PIELIKUMI

1. pielikums – Testpiemēri *Runātājs (Laika_zinas_2)*

Nr	Saturs	Tips	Komentārs	Rezultāti			
1.	Vakar vakarā stundas laikā sasaucām visu iesaistīto pušu ārkārtas sanākumi un atkal izdiskutējām galvenos strīdīgos punktus.	Dabiskums	Saturam izvēlēts teksts no ziņu portāla ar dabisku runu (citātu).	3	3.5	4	3.6
2.	Grozījumu projektā paredzēts precizēt, ka izglītības iestādēs, sociālās aprūpes, sociālās rehabilitācijas un ārstniecības iestādēs produktu iepirkumā būs atļauts iegādāties arī tādus produktus, kas satur dabīgos aromatizētājus un dabīgās krāsvielas. Tāpat būs atļauts uzturā iekļaut arī apstrādātus zvejniecības produktus, ievērojot papildu kvalitātes prasības, piemēram,	Saprotamība	Saturam izvēlēts ziņu teksts, kas nesatur dabisku runu.	3.5	2.8	3.5	3.2

	nesatur garšas pastiprinātājus, krāsvielas un tamlīdzīgi.						
3.	Robert, mums ir laiks iet gulēt. Ir jau pāri pulksten divpadsmitiem, un rīt arī ir diena. Es uzlikšu modinātāju uz pieciem trīsdesmit, man ir aizdomas, ka no rīta vajadzēs tīrīt sniegu.	Dabiskums	Saturam izvēlēts fragments no dabiskās runas (citēta ikdienišķa saruna).	3	3.2	3.1	3
4.	Pastāsti man par to, ko mēs ēdīsim vakariņās. Un tu to pierakstīsi? Jā. Mhm. Nu? Es nezinu, ko mēs ēdīsim. Nu, vakariņās mēs ēdīsim zupu ar siermaizēm. Saldajā ēdīsim medusmaizes ar tēju, piparmētru.	Dabiskums	Saturam izvēlēts fragments no dabiskās runas (citēta ikdienišķa saruna).	2	2.7	2.5	3

A. pielikums – Testpiemēri *Runātājs* versiju tests 1

<i>Nr</i>	<i>Saturs</i>	<i>Tips</i>	<i>Komentārs</i>	<i>Laika_zinas_1</i>	<i>Laika_zinas_2</i>	<i>Pasakas_1</i>	<i>Pasakas_2</i>	<i>Saeima</i>
1.	Vakar vakarā stundas laikā sasaucām visu iesaistīto pušu ārkārtas sanākumi un atkal izdiskutējām galvenos strīdīgos	Dabiskums	Saturam izvēlēts teksts no ziņu portāla ar dabisku runu (citātu).	2	3	1	2	2

<i>Nr</i>	<i>Saturs</i>	<i>Tips</i>	<i>Komentārs</i>	<i>Laika_zinas_1</i>	<i>Laika_zinas_2</i>	<i>Pasakas_1</i>	<i>Pasakas_2</i>	<i>Saeima</i>
	punktus.							
2.	Grozījumu projektā paredzēts precizēt, ka izglītības iestādēs, sociālās aprūpes, sociālās rehabilitācijas un ārstniecības iestādēs produktu iepirkumā būs atļauts iegādāties arī tādus produktus, kas satur dabīgos aromatizētājus un dabīgās krāsvielas. Tāpat būs atļauts uzturā iekļaut arī apstrādātus zvejniecības produktus, ievērojot papildu kvalitātes prasības, piemēram, nesatur garšas pastiprinātājus, krāsvielas un tamlīdzīgi.	Saprotamība	Saturam izvēlēts ziņu teksts, kas nesatur dabisku runu.	2	3.5	2	3	2

B. pielikums – Testpiemēri Runātājs versiju tests 2

<i>Nr</i>	<i>Saturs</i>	<i>Tips</i>	<i>Komentārs</i>	<i>Laika_zinas_1</i>	<i>Laika_zinas_2</i>	<i>Pasakas_1</i>	<i>Pasakas_2</i>	<i>Saeima</i>
1.	Vakar vakarā stundas laikā sasaucām visu iesaistīto pušu ārkārtas sanākumi un atkal izdiskutējām galvenos strīdīgos punktus.	Dabiskums	Saturam izvēlēts teksts no ziņu portāla ar dabisku runu (citātu).	3	3.5	1.5	2	2.5
2.	Grozījumu projektā paredzēts precizēt, ka izglītības iestādēs, sociālās aprūpes, sociālās rehabilitācijas un ārstniecības iestādēs produktu iepirkumā būs atļauts iegādāties arī tādus produktus, kas satur dabīgos aromatizētājus un dabīgās krāsvielas. Tāpat būs atļauts uzturā iekļaut arī apstrādātus zvejniecības produktus, ievērojot papildu kvalitātes prasības, piemēram, nesatur garšas pastiprinātājus, krāsvielas un tamlīdzīgi.	Saprotamība	Saturam izvēlēts ziņu teksts, kas nesatur dabisku runu.	3.2	2.8	1.7	2.1	3

C. pielikums – Testpiemēri Runātājs versiju tests 3

<i>Nr</i>	<i>Saturs</i>	<i>Tips</i>	<i>Komentārs</i>	<i>Laika_zinas_1</i>	<i>Laika_zinas_2</i>	<i>Pasakas_1</i>	<i>Pasakas_2</i>	<i>Saeima</i>
1.	Vakar vakarā stundas laikā sasaucām visu iesaistīto pušu ārkārtas sanākumi un atkal izdiskutējām galvenos strīdīgos punktus.	Dabiskums	Saturam izvēlēts teksts no ziņu portāla ar dabisku runu (citātu).	4	4	2	3	4.5
2.	Grozījumu projektā paredzēts precizēt, ka izglītības iestādēs, sociālās aprūpes, sociālās rehabilitācijas un ārstniecības iestādēs produktu iepirkumā būs atļauts iegādāties arī tādus produktus, kas satur dabīgos aromatizētājus un dabīgās krāsvielas. Tāpat būs atļauts uzturā iekļaut arī apstrādātus zvejniecības produktus, ievērojot papildu kvalitātes prasības, piemēram, nesatur garšas pastiprinātājus, krāsvielas un tamlīdzīgi.	Dabiskums	Saturam izvēlēts ziņu teksts, kas nesatur dabisku runu.	4	3.5	3	3	3

2. pielikums – Testpiemēri *eSpeak*

<i>Nr</i>	<i>Saturs</i>	<i>Tips</i>	<i>Komentārs</i>	<i>Rezultāti</i>			
1.	Vakar vakarā stundas laikā sasaucām visu iesaistīto pušu ārkārtas sanākumi un atkal izdiskutējām galvenos strīdīgos punktus.	Dabiskums	Saturam izvēlēts teksts no ziņu portāla ar dabisku runu (citātu).	2	2	2	2
2.	Grozījumu projektā paredzēts precizēt, ka izglītības iestādēs, sociālās aprūpes, sociālās rehabilitācijas un ārstniecības iestādēs produktu iepirkumā būs atļauts iegādāties arī tādus produktus, kas satur dabīgos aromatizētājus un dabīgās krāsvielas. Tāpat būs atļauts uzturā iekļaut arī apstrādātus zvejniecības produktus, ievērojot papildu kvalitātes prasības, piemēram, nesatur garšas pastiprinātājus, krāsvielas un tamlīdzīgi.	Saprotamība	Saturam izvēlēts ziņu teksts, kas nesatur dabisku runu.	1	2	2	2

3.	Robert, mums ir laiks iet gulēt. Ir jau pāri pulksten divpadsmitiem, un rīt arī ir diena. Es uzlikšu modinātāju uz pieciem trīsdesmit, man ir aizdomas, ka no rīta vajadzēs tīrīt sniegu.	Dabiskums	Saturam izvēlēts fragments no dabiskās runas (citēta ikdienišķa saruna).	2	2	3	2
4.	Pastāsti man par to, ko mēs ēdīsim vakariņās. Un tu to pierakstīsi? Jā. Mhm. Nu? Es nezinu, ko mēs ēdīsim. Nu, vakariņās mēs ēdīsim zupu ar siermaizēm. Saldajā ēdīsim medusmaizes ar tēju, piparmētru.	Dabiskums	Saturam izvēlēts fragments no dabiskās runas (citēta ikdienišķa saruna).	2	2	2	2
5.	2015. gadā salīdzinājumā ar 2014. gadu vidējais ražotāju cenu līmenis Latvijas rūpniecībā samazinājās par 1,0 %, liecina Centrālās statistikas pārvaldes (CSP) dati. Vietējā tirgū realizētajai produkcijai cenas samazinājās par 1,4 %, eksportētajai produkcijai –	Teksta apstrāde	Saturam izvēlēts teksts no Centrālās statistikas biroja raksta.	12	3	13	12

	par 0,5 %, t. sk. eksportam uz eirozonas valstīm cenas samazinājās par 0,4 %, eksportam uz ārpus eirozonas valstīm – par 0,5 %.						
6.	"Klasikas" vilnī - koncerts, kas 2015. gada 13. novembrī izskanēja Lielajā ģildē. Tur muzicēja pianists Nikolass Angeličs (Nicholas Angelich, ASV/Francija), Latvijas Nacionālais simfoniskais orķestris un diriģents Andris Poga. Skan Roberta Šūmaņa Koncerts klavierēm un orķestrim laminorā op. 54 (1846) un Johanneses Brāmsa Ceturtā simfonija mīnīorā op. 98 (1885).	Teksta apstrāde	Saturam izvēlēts teksts no radio "Klasika" ziņu rakstiem.	4	2	4	6

3. pielikums – Testpiemēri *Tildes Biroja* sintezatoram

Nr	Saturs	Tips	Komentārs	Rezultāti			
1.	Būsim reālisti - šādu iestāžu vadības maiņa Latvijā nav jauna prakse, un tam biju gatavs.	Dabiskums	Saturam izvēlēts teksts no ziņu portāla ar dabisku runu (citātu).	4	3.5	3	4
2.	Viņa amata pildīšanas laikā veiksmīgi ieviests vienotais uzraudzības un vienotais noregulējuma mehānisms, FKTK pievienojoties Eiropas līmeņa banku uzraugu saimei. Tāpat šo gadu laikā izdevies stiprināt nerezidentu banku uzraudzību un pilnveidot noziedzīgi iegūtu līdzekļu legalizēšanas un terorisma finansēšanas novēršanas jomu, ieviešot speciālu pasākumu kopumu banku segmenta efektīvai pārraudzībai.	Saprotamība	Saturam izvēlēts ziņu teksts, kas nesatur dabisku runu.	3	3	4	4
3.	Robert, mums ir laiks iet gulēt. Ir jau pāri pulksten divpadsmitiem,	Dabiskums	Saturam izvēlēts fragments no	4	3	4	4

	un rīt arī ir diena. Es uzlikšu modinātāju uz pieciem trīsdesmit, man ir aizdomas, ka no rīta vajadzēs tīrīt sniegu.		dabiskās runas (citēta ikdienišķa saruna).				
4.	Pastāsti man par to, ko mēs ēdīsim vakariņās. Un tu to pierakstīsi? Jā. Mhm. Nu? Es nezinu, ko mēs ēdīsim. Nu, vakariņās mēs ēdīsim zupu ar siermaizēm. Saldajā ēdīsim medusmaizes ar tēju, piparmētru.	Dabiskums	Saturam izvēlēts fragments no dabiskās runas (citēta ikdienišķa saruna).	3	3	3	4
5.	2015. gada decembrī salīdzinājumā ar novembri ražotāju cenu līmenis Latvijas rūpniecībā samazinājās par 0,2 %. Vietējā tirgū realizētai produkcijai cenu līmenis saruka par 0,4 %, eksportētajai produkcijai – par 0,1 %. Eksportam uz eirozonas valstīm cenu līmenis samazinājās par 0,3 %, bet eksportam uz ārpus eirozonas	Teksta apstrāde	Saturam izvēlēts teksts no Centrālā statistikas biroja raksta.	10	1	10	8

	valstīm pieauga par 0,1 %.						
6.	27. janvārī 18.00 Radisson Blu Hotel Latvija konferenču zālē notiks pēc vēsturiskiem aprakstiem restaurētā krieviņu jeb votu tautastērpa prezentācija, kuru organizē biedrība "Zemgales mantojums" ar Zemgales plānošanas reģiona un VKKF atbalstu sadarbībā ar Latvijas folkloras biedrību.	Teksta apstrāde	Saturam izvēlēts teksts no radio "Klasika" ziņu rakstiem.	1	1	1	2

4. pielikums – Testpiemēri *runa.ailab.lv*

<i>Nr</i>	<i>Saturs</i>	<i>Tips</i>	<i>Komentārs</i>	<i>Rezultāti</i>			
1.	KM nešaubās, ka šādi būvnieka izvēlētie indeksi tam ir labvēlīgāki nekā līgumā paredzētie. Tomēr KM jau vairākkārt ir norādījusi, ka tai nav tiesību grozīt līgumā paredzētos indeksus, ja tie palielina par LNB celtniecību maksājamo līguma cenu.	Dabiskums	Saturam izvēlēts teksts no ziņu portāla ar dabisku runu (citātu).	0.5	2	1	1

2.	FKTK padome var veikt visas tai likumā noteiktās funkcijas arī četru padomes locekļu sastāvā, līdz ar to, kamēr Saeima apstiprinās jaunu iestādes vadītāju, FKTK komanda var pilnvērtīgi turpināt sāktos darbus.	Saprotamība	Saturam izvēlēts ziņu teksts, kas nesatur dabisku runu.	2	2	1	1
3.	Robert, mums ir laiks iet gulēt. Ir jau pāri pulksten divpadsmitiem, un rīt arī ir diena. Es uzlikšu modinātāju uz pieciem trīsdesmit, man ir aizdomas, ka no rīta vajadzēs tīrīt sniegu.	Dabiskums	Saturam izvēlēts fragments no dabiskās runas (citēta ikdienišķa saruna).	1	2	2	1
4.	Pastāsti man par to, ko mēs ēdīsim vakariņās. Un tu to pierakstīsi? Jā. Mhm. Nu? Es nezinu, ko mēs ēdīsim. Nu, vakariņās mēs ēdīsim zupu ar siermaizēm. Saldajā ēdīsim medusmaizes ar tēju, piparmētru.	Dabiskums	Saturam izvēlēts fragments no dabiskās runas (citēta ikdienišķa saruna).	1	2	1	1

5.	2015. gada decembrī, salīdzinot ar 2014. gada decembri, vidējais ražotāju cenu līmenis Latvijas rūpniecībā samazinājās par 1,6 %. Vietējā tirgū realizētajai produkcijai cenas saruka par 2,6 %, eksportētajai produkcijai - par 0,5 %. Eksportam uz eirozonas valstīm cenas samazinājās par 1,2 %, eksportam uz ārpus eirozonas valstīm vidējais ražotāju cenu līmenis nemainījās.	Teksta apstrāde	Saturam izvēlēts teksts no Centrālā statistikas biroja raksta.	13	2	10	10
6.	29. janvārī Liepājas koncertzālē "Lielais dzintars", 30. janvārī Vidzemes koncertzālē "Cēsis" un 31. janvārī Rēzeknē, Latgales vēstniecībā "Gors" ar koncertu "Mūžam jaunais" Liepājas Simfoniskais orķestris ieskandinās Imanta Kalniņa 75. jubilejas gadu.	Teksta apstrāde	Saturam izvēlēts teksts no radio "Klasika" ziņu rakstiem.	6	8	4	4

Maģistra darbs: **Latviešu valodas runas sintēze**

Ar savu parakstu apliecinu, ka pētījums veikts patstāvīgi, izmantoti tikai tajā norādītie informācijas avoti un iesniegtā darba elektroniskā kopija atbilst izdrukai.

Autors: _____

(Autora paraksts)

Ar savu parakstu apliecinu, ka esmu lasījis augstāk minēto maģistra darbu un atzīstu to par **p i e m ē r o t u / n e p i e m ē r o t u** (nevajadzīgo svītrot) aizstāvēšanai Latvijas Universitātes datorzinātņu maģistrantūrā.

Darba vadītājs: _____

(Vadītāja paraksts)

Darbs iesniegts **maģistrantūras sekretariātā** _____

(Iesniegšanas datums)

Ar šo es apliecinu, ka darba elektroniskā versija ir augšupielādēta LU informatīvajā sistēmā.

Studiju metodiķe: _____

(Metodiķes paraksts)

Recenzents: _____

(Akad.amats, zin.grāds, vārds, uzvārds)

Darbs aizstāvēts maģistra gala pārbaudījuma komisijas sēdē

_____ prot. Nr. _____

(Darba aizstāvēšanas datums)

Komisijas sekretārs: _____

(Sekretāra paraksts)