



Інструкцыя па апісанні
сэрвісаў партала
www.corpus.by



Гэтая інструкцыя – стандарт падрыхтоўкі і афармлення дакументаў апісання – даведка да сэрвісаў праекта corpus.by

Тут змяшчаецца інфармацыя па наступных аспектах:

- патрабаванні да **структуры** даведкі
- патрабаванні да **афармлення** даведкі
- **прыклады** рэкамендаванага апісання сэрвісаў



Правілы, вызначаныя ў патрабаваннях, з'яўляюцца абавязковымі для выканання ўсімі распрацоўшчыкамі апісанняў сэрвісаў. Гэты дакумент з'яўляецца адкрытым, таму самімі карыстальнікамі і распрацоўшчыкамі даведак могуць быць унесены прапановы па паляпшэнні стандарту, якія будуць разгледжаны і ўзгоднены з удзельнікамі суполкі "Семінар па Інтэрнэт-тэхналогіях".

Патрабаванні да структуры даведкі

Кожны дакумент даведкі павінны ўключаць наступныя вылучаныя асобнымі блокам часткі:

- **Загалавак**
- **Апісанне сэрвісу**
- **Асноўныя тэрміны і паняцці**
- **Практычная каштоўнасць**
- **Асаблівасці сэрвісу**
- **Апісанне інтэрфейсу карыстальніка**
- **Карыстальніцкі(я) сцэнар(ы) працы з сэрвісам**
- **Мадэль, алгарытм**
- **Спасылкі на крыніцы**
- **Перакрыжаваныя спасылкі**

Патрабаванні да афармлення даведкі

Кожная з пералічаных частак з'яўляецца асобным раздзелам дакумента, які абмежаваны абзацам і загаловак другога узроўню.

Рэкамендавана пазбягаць абаротаў, якія ўключаюць займеннікі, напрыклад, “я прапанаваў”, “ён апрацоўвае” і г. д.



Прыклад

Апісанне сэрвіса “Графічнае
адлюстраванне контуру асноўнага тону”

Загалолак і апісанне сэрвісу

✓ Графічнае адлюстраванне контуру асноўнага тону

Сэрвіс [“Графічнае адлюстраванне контуру асноўнага тону”](#) дазваляе карыстальніку атрымаць графічнае адлюстраванне контуру асноўнага тону маўленчай фразы анлайн. На ўваход сэрвісу даецца фанаграма фразы працягласцю не больш за хвіліну. Фанаграма можа быць загрузана на сэрвіс з цвёрдага дыску камп’ютара або па спасылцы з сеткі Інтэрнэт, павінна быць у фармаце .wav ці запісана праз магчымасці аўдыязапісу сэрвісу. На выхадзе атрымліваюцца малюнкi асцылаграмы, спектраграмы і контура асноўнага тону запампаванай фразы.

Асноўныя тэрміны і паняцці

Асноўныя тэрміны і паняцці

Частата асноўнага тону (ЧАТ) з’яўляецца адным з важнейшых фізічных параметраў маўленчага сігнала. З пункту гледжання акустыкі яна ўяўляе сабой першую гармоніку танальнага складніка сігнала, які зазвычай нясе максімальную энергію. Наяўнасць частаты асноўнага тону ўласцівая вакалізаваным гукам (усім галосным і частцы зычных, напрыклад, звонкім), і наадварот, яе адсутнасць уласцівая глухім зычным. Дынаміка контуру ЧАТ мае прамую карэляцыйную сувязь з прасадычнымі характарыстыкамі маўлення, такімі як:

- мелодыка (рух частаты асноўнага тону);
- рытміка (бягучая змена працягласці гукаў і паўз);
- энергетыка (бягучае змяненне сілы гука).

EXAMPLE

EXAMPLE

Практычная каштоўнасць

Практычная каштоўнасць

Дакладная ідэнтыфікацыя і аналіз гэтых параметраў з'яўляецца важным этапам у даследаваннях спецыялістамі (фанетыстамі, фонаскапістамі і інш.) гутарковага маўлення з мэтай вызначэння розных прасадых і інтанацыйных з'яў вуснага маўлення. Сэрвіс можа выкарыстоўвацца ў праграмах для вывучэння інтанацый замежных моў, аплікацый фонаскапічнай экспертызы.

Асаблівасці сэрвісу

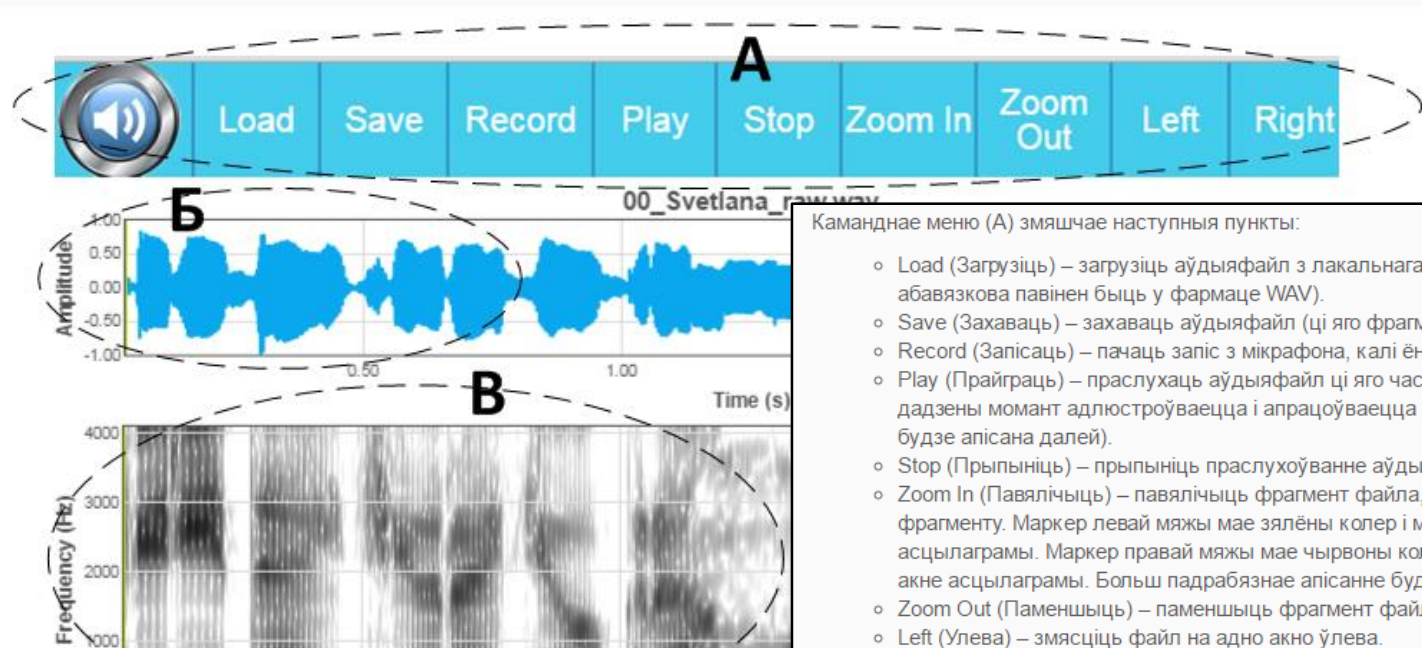
Асаблівасці

Сэрвіс выкарыстоўвае перцэптыўны алгарытм вылічэння ЧАТ – [SWIPEP](#) [1], заснаваны на аўдыторнай мадэлі ўспрымання ЧАТ чалавекам. Дадзены алгарытм з'яўляецца [адным з самых якасных і распаўсюджаных алгарытмаў вызначэння ЧАТ на сённяшні дзень](#) [2].

Апісанне інтэрфейсу карыстальніка

Апісанне інтэрфейсу

Графічны інтэрфейс сэрвісу ўключае наступныя асноўныя часткі, прадстаўленыя на малюнку 1.



Каманднае меню (А) змяшчае наступныя пункты:

- Load (Загрузіць) – загрузіць аўдыяфайл з лакальнага дыску камп'ютара ці з іншага месца па URL (аўдыяфайл абавязкова павінен быць у фармаце WAV).
- Save (Захаваць) – захаваць аўдыяфайл (ці яго фрагмент) на лакальны дыск камп'ютара карыстальніка.
- Record (Запісаць) – пачаць запіс з мікрафона, калі ён ёсць і падключаны да камп'ютара карыстальніка.
- Play (Прайграць) – праслухаць аўдыяфайл ці яго частку (у залежнасці ад таго, увесь файл ці толькі яго фрагмент у дадзены момант адлюстроўваецца і апрацоўваецца ў акне асцылаграмы. Больш падрабязна праца з фрагментам будзе апісана далей).
- Stop (Прыпыніць) – прыпыніць праслухоўванне аўдыяфайла.
- Zoom In (Павялічыць) – павялічыць фрагмент файла, выдзелены з дапамогай маркераў левай і правай мяжы фрагменту. Маркер левай мяжы мае зялёны колер і можа быць выстаўлены клікам левай кнопкі мышы ў акне асцылаграмы. Маркер правай мяжы мае чырвоны колер і можа быць выстаўлены клікам правай кнопкі мышы ў акне асцылаграмы. Больш падрабязнае апісанне будзе прыведзена далей.
- Zoom Out (Паменшыць) – паменшыць фрагмент файла, вярнуцца да прагляду ўсяго файла цалкам.
- Left (Улева) – змясціць файл на адно акно ўлева.
- Right (Управа) – змясціць файл на адно акно ўправа.

Асцылаграма сігналу (Б) – двухмернае часовае адлюстраванне сігналу. Больш дакладна, гэтая залежнасць ад часу і амплітуды сігналу ад часу: вось абсцыс (x) – час, вось ардынат (y) – амплітуда сігналу. Наватым графікам дазваляе дэталізаваць частоты сігналу і змяненне магутнасці ўсіх яго кампанентаў разам (танальнай, шумавой).

Спектраграма сігналу (B) – трохмернае частотна-часовае адлюстраванне сігналу. Больш дакладна, гэтая залежнасць магутнасці сігналу ад часу і частаты адначасова: вось абсцыс (x) – час, вось ардынат (y) – частата, вось аплікат (z) – магутнасць сігналу (у дадзеным выпадку замест асобнай графічнай восі выкарыстоўваецца колер: чым большая інтэнсіўнасць колеру, тым большую магутнасць мае сігнал на гэтай частаце ў дадзены момант часу). У адрозненні ад

Карыстальніцкі(я) сцэнар(ы) працы з сэрвісам

Карыстальніцкія сцэнары працы з сэрвісам

Сцэнар 1. Адлюстраванне контуру ЧАТ аўдыяфайла, захаванага на лакальным дыску камп'ютара

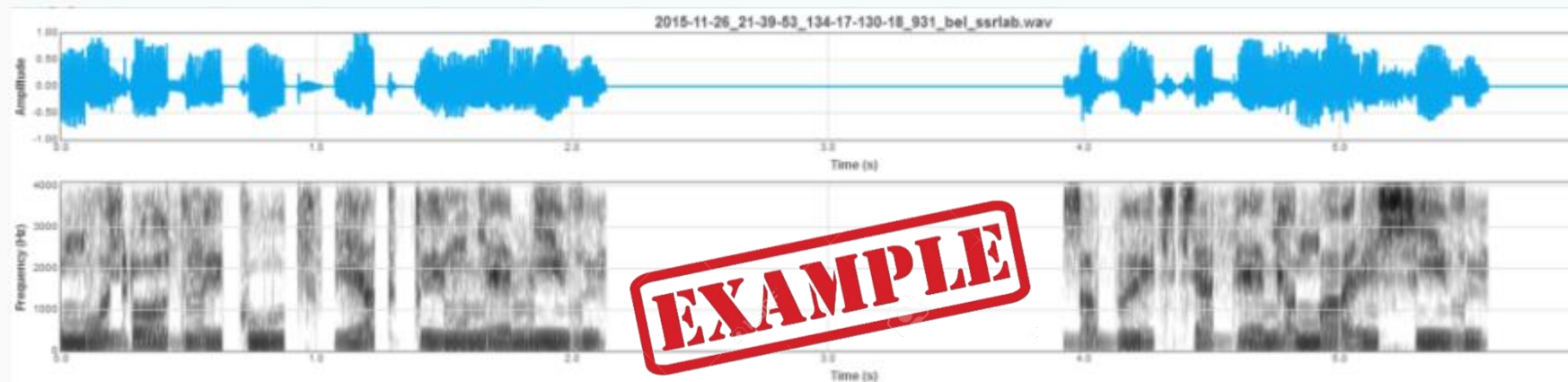
1.1. Зайсці на старонку сэрвісу "[Графічнае адлюстраванне частаты асноўнага тону](#)".

1.2. Націснуць кнопку Load.

1.3-а. У акне выбару файла націснуць "Выбраць файл". Адкрыецца стандартнае дыялогавае акно аперацыйнай сістэмы для выбару файла на дыску. Выбраць файл з лакальнага дыску камп'ютара.

1.3-б. У акне ўвесці URL да файла, які знаходзіцца ў сетцы, напрыклад: http://corpus.by/tts3/cache/out/2015-11-26_21-39-53_134-17-130-18_931_bel_sslab.wav

1.4. Дачакацца адлюстравання асцылаграмы, разліку спектраграмы і контуру ЧАТ (малюнак 2).



Мадэль

Мадэль, алгарытм

З матэматычнага пункту гледжання працэс апрацоўкі інфармацыі ў сэрвісе "Графічнае адлюстраванне контуру асноўнага тону (Pitch Plotter)" можна прадставіць з дапамогай наступнага выразу:

$$(\mathbf{o}, \mathbf{S}, \mathbf{p}) = \mathcal{P}(\mathbf{x}),$$

дзе $\mathbf{x} = \{x(n) \mid x \in R, n = \overline{0, 1, 2, \dots, N-1}\}$ – вектар адлікаў уваходнага маўленчага сігналу. Сігнал быў атрыманы з wav-файла, запампаванага ў сэрвіс пры дапамозе каманды "Load", або з мікрафона карыстальніка пры дапамозе каманды "Record", n – нумар адліку ў сігнале.

дзе $\mathbf{o} = \{o(n) \mid o \in R, n = \overline{0, 1, 2, \dots, N-1}\}$ – асцылаграма сігналу (вектар значэнняў кропак графіка асцылаграмы), n – нумар адліку ў сігнале.

дзе $\mathbf{S} = \{s(k, m) \mid s \in R, m = \overline{0, 1, 2, \dots, M-1}, k = \overline{0, 1, 2, \dots, K-1}\}$ – спектраграма сігналу (матрыца значэнняў кропак графіку спектраграмы), k – нумар гармонікі сігналу, m – нумар фрэіму сігналу.

дзе $\mathbf{p} = \{p(m) \mid p \in R, m = \overline{0, 1, 2, \dots, M-1}\}$ – графік частаты асноўнага тону (вектар значэнняў кропак графіка), m – нумар фрэіму сігналу.

Такім чынам, можна сцвярджаць, што функцыя \mathcal{P} , якая апісвае працу сэрвісу, складаецца з суперпазіцыі трох пераўтварэнняў:

$$\left\{ \begin{array}{l} \mathbf{x} \xrightarrow{\mathcal{P}_o} \mathbf{o}, \\ \mathbf{x} \xrightarrow{\mathcal{P}_s} \mathbf{S}, \\ (\mathbf{x}, \mathbf{S}) \xrightarrow{\mathcal{P}_{x,s}} \mathbf{p} \end{array} \right\}$$

EXAMPLE

Алгарытм прыклад алгарытму генерацыі транскрыпцыі

Алгарытм генерацыі кірылічнай транскрыпцыі па тэксце ў алафонным выглядзе прадстаўлены на малюнку 1.

На ўваход алгарытму падаецца алафонны тэкст T . Далей алгарытм выконвае наступныя крокі:

Крок 1. Фарміраванне правілаў замены. Адбываецца зварот да БД, якая змяшчае мноства адпаведнасцей “алафон – сімвал транскрыпцыі”, і фарміраванне шэрагу правілаў $R = \langle R_1, \dots, R_n \rangle$, дзе $R_i = \langle a_i, tr_i \rangle$, a_i – код алафона, tr_i – адпаведны алафону a_i сімвал транскрыпцыі, $i = 1 \dots n$, n – колькасць правілаў.

Крок 2. Вылучэнне алафонаў. У тэксце T адбываецца паслядоўнае вылучэнне алафонаў згодна з шаблонам іх дапушчальнага выгляду P_t . Шаблон P_t можа быць фармалізаваны на мове рэгулярных выразаў наступным чынам:

$$P_t = [A-Z]\{1,3\}[']\{0,1\}[d]\{3\}, \quad (2.1)$$

дзе $[A-Z]\{1,3\}$ – адна, дзве альбо тры вялікія літары лацінскага алфавіту, $[']\{0,1\}$ – магчымы знак апострафа, $[d]\{3\}$ – тры дзесятковыя лічбы (0, ..., 9). На гэтым кроку здзяйсняецца пошук ва ўваходным тэксце сімвалных паслядоўнасцей, якія адпавядаюць шаблону P_t (2.1), і захаванне іх у спіс алафонаў L_a . Такім чынам з тэксту вымаюцца ўсе сімвалы, якія прадстаўляюць сабой алафоны, а ўсе астатнія сімвалы ігнаруюцца. З тэксту T атрымліваецца паслядоўны спіс алафонаў L_a .

Крок 3. Замена алафонаў кірылічнымі адпаведнікамі. Для кожнага алафона a_i са спісу L_a знаходзіцца адпаведнік – сімвал альбо паслядоўнасць сімвалаў кірылічнай транскрыпцыі – tr_i згодна са сфарміраваным раней шэрагам правілаў R . Адбываецца збор інфармацыі аб адсутнасці ў спісе правілаў R тых ці іншых адпаведнасцей. У выніку паслядоўнай апрацоўкі ўсяго спісу L_a атрымліваем спіс сімвалаў кірылічнай транскрыпцыі L_r .

Крок 4. Генерацыя выніковага тэксту. Спіс сімвалаў транскрыпцыі L_r паслядоўна склейваецца ў выніковы тэкст T_r .

Крок 5. Падрыхтоўка захавання даных у архіве. У працэсе генерацыі транскрыпцыі адначасова адбываецца збор інфармацыі аб уваходных (алафонны тэкст T), выніковых (выніковая транскрыпцыя T_r) і аналітычных (памер уваходнага тэксту, IP-адрас карыстальніка, час запыту і т.д.) даных для аператыўнага выпраўлення памылак алгарытму, статыстычнага аналізу даных. Таксама збіраецца інфармацыя аб адсутнасці тых ці іншых элементаў у базе “алафон – кірылічная транскрыпцыя”.

Крок 6. Рассылка апавяшчэнняў распрацоўшчыкам. Пералічаная вышэй інфармацыя складаецца ў электронны ліст і накіроўваецца распрацоўшчыкам – эксперту-лінгвісту і інжынеру-праграмісту – для вызначэння тыпу праблемнай сітуацыі (памылкі ў лінгвістычных пытаннях, памылкі ў алгарытмах і праграмных кодах) і шляхоў яе вырашэння.

Крок 6. Канец алгарытму.

У выніку апісанага вышэй алгарытму з алафоннага тэксту генеруецца тэкст у кірылічнай транскрыпцыі. Напрыклад, калі падаць на ўваход тэкст

M004,O113,J013./,R032,O022,D001,N004,Y322./,K001,U032,T000./,#C3,
J002,A142,K004./,T002,Y121./,M001,N004,E143./,M002,I041,L004,Y310./,#

E2,

то выніковы тэкст будзе мець наступны выгляд:

[м'ой] [р'одны] [к'ут]
[й'ак] [т'ы] [мн'э] [м'ілы]



Малюнак 1 – Агульная схема алгарытму генерацыі кірылічнай транскрыпцыі па тэксце ў алафонным выглядзе

EXAMPLE

Спасылкі на крыніцы

Спасылкі на крыніцы

Відзаўрок па працы з гэтым сэрвісам: [Pitch Plotter tutorial – Corpus.by](#)

Старонка сэрвісу: <http://corpus.by/PitchPlotter/?lang=be>

Перакрыжаваныя спасылкі

Перакрыжаваныя спасылкі

1. [SWIPEP](#)
2. [Instantaneous Pitch Estimation based on Rapt Framework](#)

Сэрвіс выкарыстоўвае перцэптыўны алгарытм вылічэння ЧАТ – [SWIPEP](#) [1], заснаваны на аўдыторнай мадэлі ўспрымання ЧАТ чалавекам. Дадзены алгарытм з’яўляецца [адным з самых якасных і распаўсюджаных алгарытмаў вызначэння ЧАТ на сённяшні дзень](#) [2].