

V1.0 (DRAFT)

Smernice za zapis izgovora besed tujega izvora v slovenskem oblikoslovnem leksikonu Sloleks

Datum	Različica	Objavitelj_ica	Opis
2022-11-04	1.0	Jaka Čibej	Objava prve različice.

Besed tujega izvora oz. tistih besed, ki ne sledijo slovenskim pravilom za grafemsko-fonemsko pretvorbo, ne moremo avtomatsko pretvoriti v fonetični zapis na podlagi oblike, zato potrebujemo vmesni ortografski zapis, iz katerega s pretvornikom tvorimo fonetični zapis v mednarodni fonetični abecedi (IPA) oz. v fonetični abecedi SAMPA.

Vmesnih ortografskih zapisov ne pišemo za vsako obliko posebej, temveč samo za osnovno obliko (Johnson - dž=onson). Iz vmesnega ortografskega zapisa osnovne oblike nato s pomočjo slovenskih oblikoslovnih vzorcev generiramo ostale oblike (dž=onsona, dž=onsonu, dž=onsona ...), iz teh pa pridobimo vse fonetične zapise.

1 Ortografski zapis izgovora

1.1 Osnovna načela zapisa

Pri besedah, v katerih izgovor ne sledi slovenskim pravilom za grafemsko-fonemsko pretvorbo (npr. Bordeaux - bord/o, ne bordeauks), izgovor zapišemo fonetizirano s črkami slovenske abecede. Dopuščamo tudi 'w' (glej 2.2 spodaj) in '3', ki nadomešča polglasnik. Naglase zapišemo s posebnimi

simboli, ki jih postavimo pred naglašeno črko: ostrivec je nadomeščen s poševnico (/), krativec z levo poševnico ('), strešica pa z enačajem (=).

Bruce → br/us
Avery → =ejv3ri
Pierre → pj=er
Shakespeare → š=ejkspir
Bydgoszcz → b/idgošč
Zürich → c/irih
Kaandorp → k/andorp
Freud → fr=ojd

1.2 Variantne izgovorjave

Pri besedah, kjer je možnih več izgovorov, zapišemo vse. Pri imenu "Alice" npr. lahko zapišemo angleško izgovorjavo (=elis) ali italijansko (al/iče), v določenih primerih pa je tudi izgovorjav iz istega jezika več (Malcolm - m/alkom in m=elk3m). Po vrstnem redu imajo prednost tisti izgovori, ki jih dojemamo kot pogostejše oz. ki bi jih pričakovali od govorca standardne slovenščine na televiziji. (Končni vrsti red in preference bodo določeni pozneje z empiričnimi preverbami, trenutno delamo po jezikovnem občutku in na podlagi virov, kot je MMC RTVSLO (glej 1.3), da bodo podatki čim bolj koristni za razpoznavo govora.)

Johnny → dž/oni | dž=oni
Malcolm → m/alkom | m=elk3m
Malkovich → m/alkovič | m=elkovič
Manchester → m=enčest3r | m/enčest3r

1.3 Dodatni viri

Fran, YouGlish, MMC RTVSLO, YouTube, Wikipedia, <https://tophonetics.com/>

2 Posebnosti

2.1 Zapis 'v'

Fonem 'v' v položaju, kjer se izgovarja kot [ʋ], zapišemo z grafemom 'v', npr. (a) za samoglasnikom in pred soglasnikom (House - h/avs, ne h/aus); in (b) na koncu besede (Johnnyjev - dž=onijev, ne

dž=onijeu). Pretvornik G2P, ki ga uporabljamo, namreč v teh položajih črko 'v' samodejno prepozna kot [ʊ].

POZOR! Če bi jo zapisali kot 'u', bi jo program neustrezno pretvoril v [u]. Z avtomatskim generiranjem preostalih oblik bi se napaka ponovila še v drugih oblikah, npr. "v/agnerjeu" - "v/agnerjeuega" - "v/agnerjeuemu".

House → h/avs
Wagnerjev → v/agnerjev
Wulfenov → v/ulfenov
Yellowstone → j=elovst=ovn

2.2 Zapis 'w'

V položajih, kjer se 'u' oz. 'u' pojavi med dvema samoglasnikoma, ga zapišemo kot 'w'.

Auer → /awer | /aw3r
Howard → h/aw3rd
Zoe → z/owi

2.3 Široki 'e' in polglasnik

V nekateri primerih (zlasti imena na -er, -en) sta mogoči variantni izgovorjavi s širokim e-jem in s polglasnikom. Navedemo obe. Vrstni red je zaenkrat tak, kot se nam zdi najbolj naraven, a ga bomo empirično določali pozneje.

Baxter → b/akst3r | b/akster | b=ekst3r | b=ekster
Schopenhauer → š/openhaw3r | š/openhawer | š/op3nhawer | š/op3nhaw3r
Pfeifer → f/ajf3r | f/ajfer
Schengen → š/eng3n | š/engen
Rasmussen → r/asmus3n | r/asmusen

2.4 Široki in ozki 'o'

Variantni izgovori so mogoči tudi pri širokih in ozkih o-jih. Navedemo vse, ki jih dojemamo in s pomočjo virov identificiramo kot relevantne, prednost pa imajo bolj relevantni/naravni/pogostejši.

Johnny → dž=oni | dž/oni
Laurie → l=ori | l/ori

2.5 Variantne izgovorjave z 'v' in 'w'

V primerih, ko v besedi nastopa črka "w" oz. je mogoče izgovor z "w", navedemo tako izgovorjavo z "v" kot tudi z "w".

Wolfgang → v/olfgang | w/olfgang
 Netware → n/etver | n=etwer
 software → s=oftver | s=oftwer
 Quentin → kw=entin | kv=entin
 orwellovski → /orvelovski | /orwelovski

2.6 Končni zveneči nezvočniki

Pri končnih zvenelih nezvočnikih ne upoštevamo premene po nezvenečnosti. Pišemo npr. "dž=erald", ne "dž=eralt". Pretvornik končne zveneče nezvočnike samodejno pretvori v nezveneče. Če bi jih že v osnovi zapisali kot nezveneče, bi ob generiranju preostalih oblik propagirali napačen izgovor (npr. "dž=eralt" - "dž=eralta" - "dž=eraltu" namesto "dž=eralt" - "dž=eralda" - "dž=eraldu").

Gerald → g=erald | dž=erald

V drugih položajih premena po nezvenečnosti ne vpliva na končni fonetični zapis (Radcliffe - r=edklif ali r=etklif), saj jih pretvornik v vsakem primeru pretvori v nezveneče. Zavaljo konsistentnosti pa tudi v teh primerih vse zveneče nezvočnike pišemo zveneče (torej r=edklif).

2.7 Polglasnik v osnovni obliki

Ne zapisujemo polglasnika, če je prisoten le v osnovni obliki, v (večini) ostalih pa ne. Npr. "Bayern" zapišemo kot "b/ajern", ne "b/ajer3n". Če zapišemo polglasnik v teh primerih, se polglasnik napačno propagira v druge oblike - b/ajer3n, b/ajer3na, b/ajer3nu. V teh primerih polglasnike dodamo v imenovalniško (morda tudi tožilniško obliko) naknadno, kar je časovno manj potratno kot brisanje polglasnika iz vseh ostalih oblik.

Pearl p/3rl, ne p/3r3l

2.8 Variantne izgovorjave z diftongi ali brez

Dodajamo variantne izgovorjave pri imenih, kjer je izgovor mogoč z diftongom ali brez.

Monroe → monr/o | monr/ov

Oprah → /opra | /ovpra

Poe → p/o | p=ov

Sophie → s/ofi | s=ovfi

Tony → t/oni | t/ovni

2.9 Variantne izgovorjave z 'a' ali 'e'

Pri nekaterih lastnih imenih (zlasti angleških) prihaja do variantne izgovorjave ('a' ali 'e') - navedemo obe.

Patrick → p=etrik | p/atrik

Paltrow → p/altrov | p=eltrov

michiganski → m/išigenski | m/išiganski

2.10 Široki in ozki 'e'

Navajamo variante izgovora s širokim in ozkim 'e'.

losangeleški → los=endželeški | los/endželeški

Manchester → m=enčest3r | m/enčest3r

manhattanski → menh=et3nski | menh/et3nski

Cambridge → k/embridž | k=embridž

2.11 Priimki na -son

Pri priimkih tipa "Jacobson", "Robson" smo v posnetkih zasledili tako izgovorjave z 'o' kot izgovorjave s polglasnikom. Zaenkrat navajamo oboje - empirično določamo pozneje. Dilema je predvsem takrat, ko tvorimo pregibne oblike: ali se tudi v teh polglasnik ohrani ali pa bi morala biti že sama zapisana oblika drugačna, če se polglasnik izpusti (dž=ejk3bs3n, dž=ejk3bsna)?

r=obson, r=obs3n

dž=ejkobson, dž=ejk3bs3n, dž=ejkobs3n, dž=ejk3bson

Operacijo Razvoj slovenščine v digitalnem okolju sofinancirata Republika Slovenija in Evropska unija iz Evropskega sklada za regionalni razvoj. Operacija se izvaja v okviru Operativnega programa za izvajanje evropske kohezijske politike v obdobju 2014-2020.

Revision #2

Created 4 November 2022 12:36:50 by Jaka Čibej

Updated 4 November 2022 12:48:03 by Jaka Čibej