



# Spóznawanje rěče

## Wuhódnoćenje kooperacije z Fraunhofer IKTS a BTU Choćebuz

Daniel Zoba, 15.06.2022



# Dźěłowe pakćiki 2021

- akustiski model (nahrawana řeč → rječaz fonemow):
  - přechod na ryzy hornjoserbski model
  - přiměrjenje na řečnika a akustiske wuměnenja
- słownik (fonemy → rječaz słowow):
  - zľěšjenja wurjekowanskich prawidłow
- řečny model (słowa → dowolene słowne skupiny):
  - zľěšjenje technologije za natwar wjetšich modelow
- nahrawanja:
  - wobšěrny nawod za wuzwolenje sadow a wuměnenja za nahrawanje



# NAHRAWANJA



# Wuzwolenje sadow a wuměnjnja za nahrawanje (1)

- wobstaranje a wobdźěłanje tekstoweho korpusa
- wuzwolenje fonetisce wuwažowanych a bohatych sadow
- nahrawanje z wjacorymi rěčnikami
- [https://github.com/ZalozbaDev/speech\\_recognition\\_corpus\\_creation](https://github.com/ZalozbaDev/speech_recognition_corpus_creation)
  - wobšěrna jendźelska rozprawa
  - němskorěčne zjeće



# Wuzwolenje sadow a wuměńjenja za nahrawanje (2)

- wobdźěłany tekstowy korpus so wužiwa za:
  - fonetiski słownik
  - řečný model
- přikłady za wudźěłanje fonetiskeho leksikona z tekstoweho korpusa pomocy prawidłow namakaja so na samsnym městnje w Github.



# SŁOWNIK



# Fonetiski słownik – prawidła (1)

- kóždemu pismikej słowa so přiřaduje jednotliwy zwuk („kanoniske wurjekowanje“)
- přidatne prawidła wopisuja wotchilenja wot tutoho přiřadowanja
- zasada při spóznawanju řeče:
  - telko zwukowych modelow kaž trěbne, tak mało kaž móžne
  - słowa z rozdźělnym woznamom dyrbja so z rozdźělnymi zwukami modelěrować (Čěsce: „byt“ – bydlenje, „být“ – być)



# Fonetiski słownik – prawidła (2)

- přikłady „kanoniske wurjekowanje“:
  - C            ts
  - Č            tS
  - Ć            tS
  - D            d
  - E            E
  - CH          x
- přikłady „wuwzaće“:
  - \_D\_K        t            zředka, dopředka, srědki, přehladka, wuslědkow ...
  - \_E\_DŽ      e            chcedža, srjedž, njeskedžbliwa, swjedženskim ...
  - E\_J\_I        \*            stejimy, meji, čejim
  - I\_CH\_        C            serbskich, młódšich, nchtó, jich ...
  - prawidło „E → e“ so na př. přepruwuje zmysłapótnosće dla (njezměni „zmysł wuprajenja“)





# Fonetiski słownik – prawidła (3)

- wuslědk:
  - ANTIKSKICH    a n t i k s k i C
  - ANTIKSKICH    a n t i k s k i x
  - ANTIKSKICH    Q a n t i k s k i C
  - ANTIKSKICH    Q a n t i k s k i x
- wužiwanje za:
  - fonetiske popisowanje nahrawanjow (bjeze změnow)
  - spóznawanje (ručne korigowane, trjeba dialekty dodać)
- přikłady: [https://github.com/ZalozbaDev/speech\\_recognition\\_corpus\\_creation](https://github.com/ZalozbaDev/speech_recognition_corpus_creation)



# Fonetiski słownik – statistika

- Jeli steji dosć wulki fonetiski słownik hižo k dispoziciji, móže so pomocu statistiskich gratow rozšěrić:
  - `phonetisaurus predict --model=models/g2p.fst --nbest=5` **ĆOPŁOBĚŁU**
    - ĆOPŁOBĚŁU tS O p w O b l w u
    - ĆOPŁOBĚŁU tS O p w O b l u
    - ĆOPŁOBĚŁU tS O p w O b e u
    - ĆOPŁOBĚŁU tS O p w O b e i u
    - ĆOPŁOBĚŁU tS U p w O b l w u
- [https://github.com/ZalozbaDev/speech\\_recognition\\_statistical\\_g2p\\_modeling](https://github.com/ZalozbaDev/speech_recognition_statistical_g2p_modeling)



# AKUSTISKI MODEL



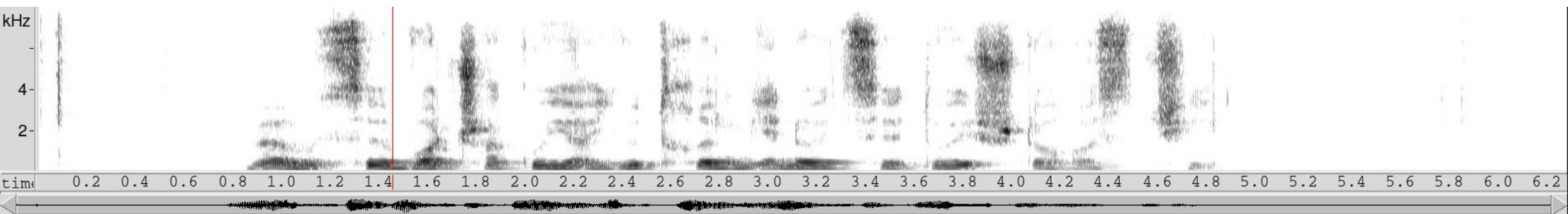
# Fonetiske popisowanje (1)

- trenowanje akustiskich modelow trjeba fonetisce popisowane nahrawanja rěče
- ručne wudźěłanje: jara naročne (a z tym njerealistisce)
- awtomatiske wudźěłanje: jenož mało zmylkow dowolene, dokelž so jenož zwjeršnje kontrolować móže

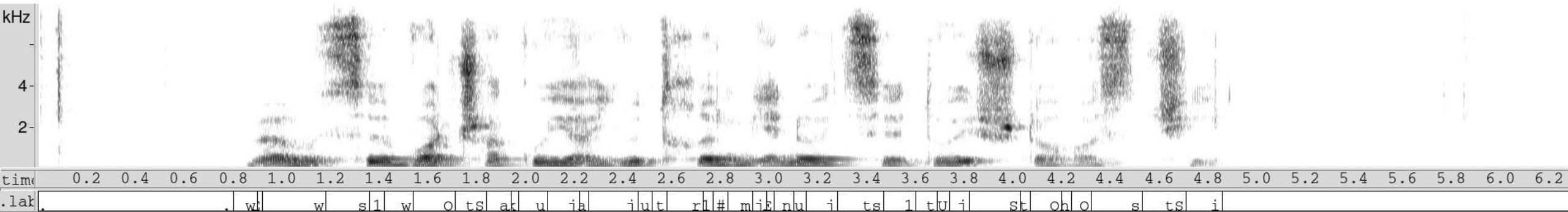


# Fonetiske popisowanje (2)

WE WSY WOČAKUJA JUTRY MJENUJCY TÓJŠTO HOSĆI



požadany rezultat

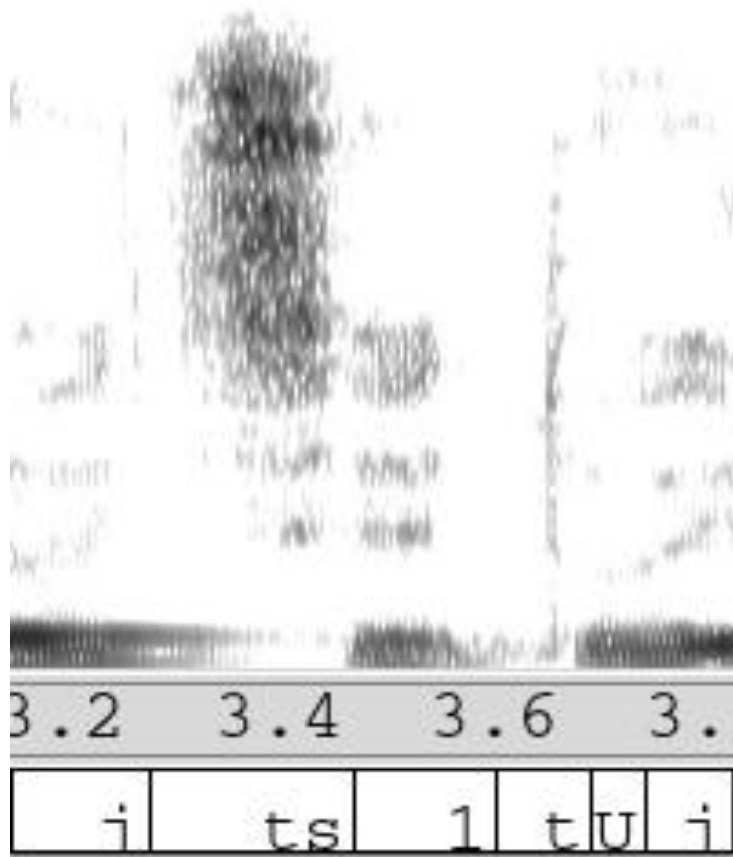




# Fonetiske popisowanje (3)

MJENUJCY TÓJŠTO

Zoom in:





# Fonetiske popisowanje (4)

- pomoc „pře-etiketěrowaneho“ němskeho akustiskeho modela
  - 'Z': ['z', 'S']
  - 'dZ': ['d', 'z', 'S']
  - 'e': ['e:']
  - 'ji': ['j', 'i:']
  - 'jn': ['j', 'n']
  - 's': ['s']
  - 't': ['t']
  - 'tS': ['t', 'S']
  - 'ts': ['t', 's']



## Fonetiske popisowanje (5)

- akustiski model + (znaty) tekst + słownik  
→ awtomatiske fonetiske popisowanje
- njewobmjezowane na jednu řeč
- ale: fonemy dyrbja so pomocu němskeho modela rozeznawać dać (nic nuznje „spóznawać“)





# Trening

- wokomiknje k dispoziciji stejace modele móža so z hač do 50..60 hodžinami nahrawanej řeču wjacorych řečnikow w studijowej kwaliće trenować
- technologiske hranicy wužiwaných monofoných modelow:
  - jenož nahrawanja wysokej kwality (studijo)
  - modele su po 60 hodžinach řečneho materiala „połne“
- trening dobreho zakładneho modela móžne, ale žadyn powšitkowny model



# Adapcija

- adapcija zakładneho modela na řečnika a mikrofon
- nahrawanje na př. 40 sadow pod „wužiwanskimi wuměnjemi“
- jasne zlepšíenje akustiskeho spóznawanja
- wužiwanje za na př. diktowanskeje funkcije abo swójskeho asistenta
- [https://github.com/ZalozbaDev/speech\\_recognition\\_acoustic\\_model\\_training](https://github.com/ZalozbaDev/speech_recognition_acoustic_model_training)



# RĚČNY MODEL



# Word classes (1)

- rěčny model wobmjezuje słowoskład na „dowolene wurjekowanja“
- čim wjace wobmjezowanjow, čim lěpše spóznawanje
- ale: kóžde wurjekowanje dyrbi so do modela zapisać
- wuchadźišćo: wobdźěłany tekstowy korpus



# Word classes (2)

- problem: kompleksnosť rosće
  - GRM.1: <PERCENT> :\_005\_ PJEĆ:PJEĆ
  - GRM.1: <PERCENT> :\_010\_ DŽESAĆ:010
  - GRM.1: <PERCENT> :\_015\_ PJATNAĆ:015
  - GRM.1: <PERCENT> :\_015\_ PJATNAĆE:015
  - GRM.1: <PERCENT> :\_020\_ DWACEĆI:020
  - GRM.1: <PERCENT> :\_020\_ DWACCÍ:020
  - GRM.1: <PERCENT> :\_020\_ DWACCI:020
  - GRM.1: <PERCENT> :\_025\_ PJEĆADWACEĆI:025
  - GRM.1: <PERCENT> :\_025\_ PJEĆADWACCÍ:025
  - GRM.1: <PERCENT> :\_025\_ PJEĆADWACCI:025
  - ...
- ličby, datum, čas atd we wšitkich wariantach napisać njeje wjace realistisce



## Word classes (3)

- rozrisanje: wudźěłanje (kompleksnych) „word classes“
- gramatika wužiwa prosće „zastupowacy symbol“
- GRM: <BRIGHTNESS> <LAMP> <NUM1-99> PROCENTOW



## Word classes (4)

- NUM1-99.txt:
  - 0 1 jedyn +1
  - 0 1 dwaj +2
  - 0 1 tři +3
  - 0 1 štyri +4
  - 0 1 NUM5-99 <eps>
  - 1



## Word classes (5)

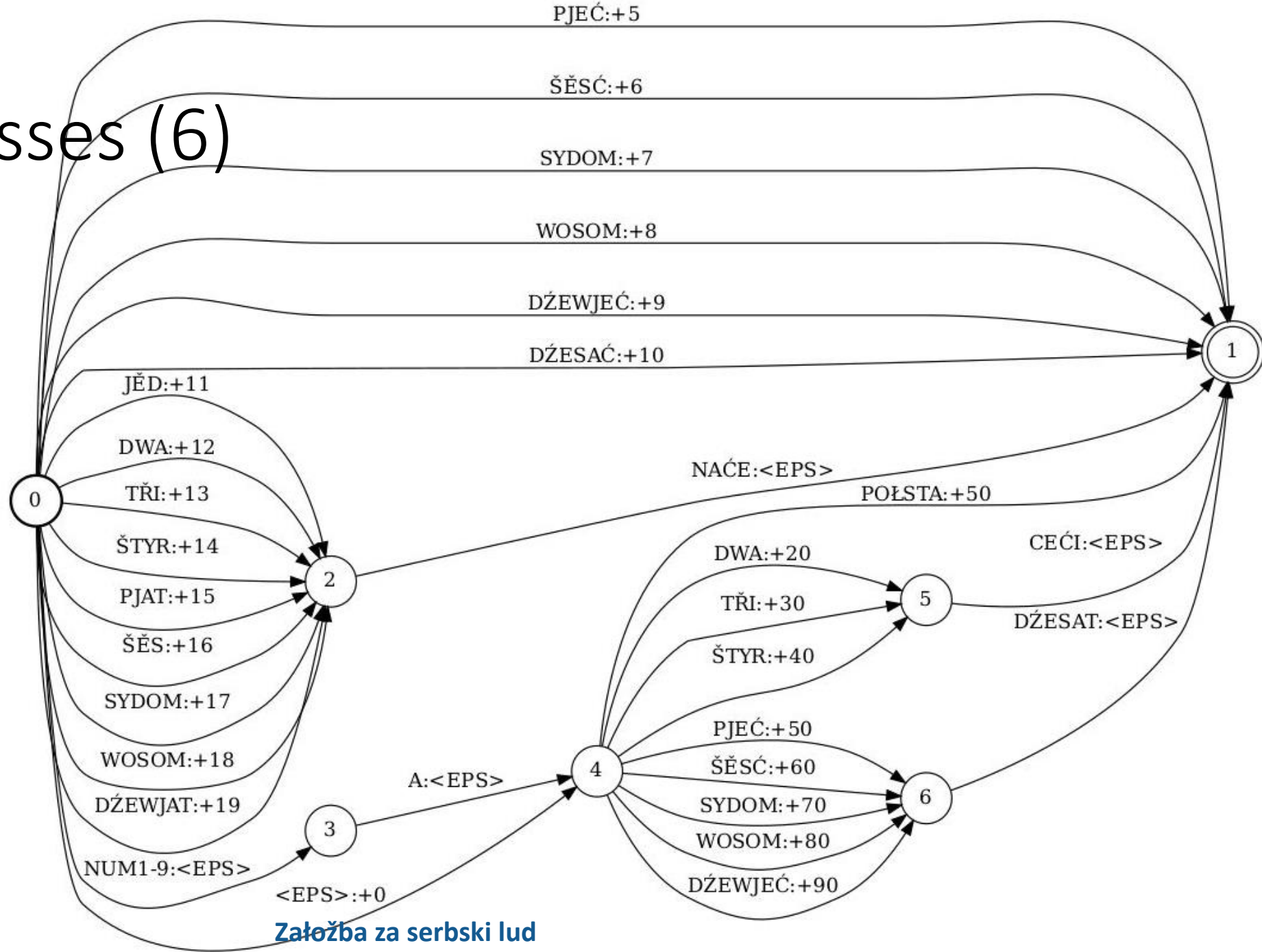
- NUM5-99.txt:
    - 0 6 pjeć +5
    - ...
    - 0 6 dźesać +10
    - 0 1 jěd +11
    - 0 1 dwa +12
    - ...
    - 0 1 dźewjat +19
  - 1 6 naće <eps>
  - 0 2 NUM1-9 <eps>
  - 2 3 a <eps>
  - 0 3 <eps> +0
  - 3 4 dwa +20
  - 3 4 tři +30
  - ...
  - 4 6 ceći <eps>
- 3 5 pjeć +50
  - ...
  - 5 6 dźesat <eps>
  - 3 6 połsta +50
  - 6





# Word classes (6)

- NUM5-99.txt:





## Word classes (7)

- NUM1-9.txt:
  - 0 1 jedyn +1
  - 0 1 dwaj +2
  - 0 1 tři +3
  - 0 1 štyri +4
  - 0 1 pjeć +5
  - 0 1 šěsc +6
  - 0 1 sydom +7
  - 0 1 wosom +8
  - 0 1 dźewjeć +9
  - 1



# Statistiski rěčny model (1)

- problem wulkich rěčnych modelow:
  - njemóža so manuelnje wudźěłać
  - trjebaja přeco wjace wobličenskeho časa
- kompromis:
  - prawdže njepodobne wurjekowanja so wotstronja
- postupowanje:
  - wudźěłanje statistiskeho „n-gram“ rěčneho modela z wobdźěłaneho tekstoweho korpusa
  - tekstowy korpus móže so přidatnje ručnje za wužiwanje „word classes“ wobdźěłać



# Statistiski rěčny model (2)

- małe tekstowe korpusy: Bigramy
  - rěčny model „znaje“ jenož wuzwolene „naslědniki“ jednoho słowa:
    - dobre
      - ranje
      - polěpšenje
      - wjedro
      - ~~smykaće~~
- wulke tekstowe korpusy : Trigramy
  - rěčny model „znaje“ jenož wuzwolene „naslědniki“ poslednjeju słowow:
    - swěca so
      - hasnje
      - swěći
      - ~~polěpši~~



## Statistiski rěčny model (3)

- móžne wužiwanje: diktowanska funkcija
- wobmjezowanje: falowace přiřadowanje „akcijow“, kaž na př. „smart home“
  - móže so zdžěla z pomocu „word classes“ narunać
- [https://github.com/ZalozbaDev/speech\\_recognition\\_language\\_modeling](https://github.com/ZalozbaDev/speech_recognition_language_modeling)



# WUHLAD



# Wuhlrad

- ~~Wozjewjenje wuslědkow z 2021~~
- delnjoserbščina
- planowane kooperacije
  - Serbski Institut: fonetika / linguistika za hornjo- a delnjoserbščinu
  - RCW: kombinacija spóznawanja ze sotra.app
  - FHG a BTU: nałoženja z wulkim rěčnym modelom
- pomoc k wosobinskemu wužiwanju a přiměrjenju prototypa „digidom“
- dalše nahrawanja



Kónc / prašenja