

Tematické okruhy k bakalářské SZk pro obor Český jazyk se specializací počítačová lingvistika

1) fonetika a fonologie češtiny

- fonetika, její náplň a cíle práce – základní fonetické pojmy, základní typy fonetických bádání; fonetická transkripce a její typy
- fonologický aspekt zkoumání procesu komunikace – základní pojmy z fonologie, fonologický přepis; foném jako jednotka systému (inventář fonémů); vztahy mezi fonémy (fonologické opozice); korelace (pár, řada atd.); slabika z fonologického hlediska; delimitativní signály z fonolog. hlediska
- vokální a konsonantické fonémy a jejich alofony – klasifikace vokálů, diftongy; klasifikace konsonantů (konsonanty vlastní/párové, sonory, glide); změny hlásek v souvislé řeči (asimilace a akomodace, změny asimilační, ostatní změny – disimilace, elize, proteze, metateze, hiát atd.)
- zvukové členění výpovědi – prostředky a formy modulace souvislé řeči (přízvuk slovní a větný, intonace, důraz); členění souvislé řeči (promluva, výpověď, výpovědní úsek, takt, slabika); slabika, stavba slabiky, vlastnosti slabiky
- ortoepie, ortofonie – vymezení pojmů, kodifikace spisovné výslovnosti; současná spisovná výslovnost (základní ortoepické příručky, konkrétní ortoepické zásady)

2) morfologie

- slovní druhy, charakteristika podle různých kritérií, značkování kategorie slovního druhu v korpusech ČNK
- přechody mezi slovními druhy, slovnědruhovými homonymiemi, disambiguace slovnědruhově víceznačných jednotek v korpusech ČNK
- pád jmen a role předložek, tokenizace a značkování víceslovných jednotek v korpusech
- rod a životnost, dvojí rod, jména vespolná, disambiguace homonym lišících se kategorií rodu a životnosti
- číslo, anomálie gramatické kategorie čísla, gramatická kategorie čísla v morfologickém značkování českých korpusů z hlediska disambiguace
- uspořádání deklinačního systému
- vývojové tendence v deklinaci substantiv, variantní koncovky substantiv v nové češtině z hlediska vývojového
- hlavní významové skupiny adjektiv, jejich morfologie, jmenné tvary, značkování jmenných tvarů v korpusech

- syntaktické a morfologické vlastnosti zájmen, druhy zájmen z hlediska morfologického značkování korpusů ČNK
- syntaktické a morfologické vlastnosti číslovek, druhy číslovek z hlediska morfologického značkování korpusů ČNK
- finitní a nefinitní tvary sloves, přechodníky, jejich minulost a přítomnost
- osoba a číslo, značkování gramatické kategorie osoby slovesných tvarů v korpusech ČNK
- vid a čas, značkování v korpusech ČNK
- slovesný rod, způsob, značkování v korpusech ČNK
- uspořádání konjugačního systému, slovesný kmen, určování slovesných tříd a vzorů, přechody mezi slovesnými třídami a vzory
- vývojové tendence v konjugaci
- adverbia a předložky, adverbializace a příslovečné spřežky
- přechody mezi neohebnými slovními druhy, disambiguace neohebných slovních duhů v případě slovnědruhovému homonymie
- citoslovce a částice
- morfemická analýza, morfonologické alternace, morfologická stavba substantivních, adjektivních a slovesných tvarů
- morfologické varianty a dublety z hlediska jejich distribuce a kodifikace, značkování variant a dublet a jejich diasambiguace

3) tvoření slov

- typy významových změn při tvoření slov
- formální postupy při derivaci; segmentace slov a slovních tvarů
- vlastnosti slovotvorných prefixů a sufixů
- kompozice, významové změny
- kritéria dělení kompozit, nevlastní kompozita z hlediska tokenizace a lemmatizace
- desubstantivní substantiva, značkování některých typů v korpusech ČNK
- deverbální substantiva
- deadjektivní substantiva
- substantivní názvy okolností
- desubstantivní adjektiva, značkování některých typů v korpusech ČNK
- deadjektivní adjektiva
- deverbální adjektiva, adjektiva slovesná a jmenné tvary adjektiv z hlediska lemmatizace v ČNK, značkování některých typů v korpusech ČNK

- deadverbiální adjektiva
- adverbializace – desubstantiva
- adverbia z ostatních slovních druhů, včetně deadverbií
- slovotvorba a zájmena, číslovky
- slovotvorná třída/kategorie, slovotvorný typ, slovotvorné hnízdo, slovotvorná synonyma
- slovotvorný a lexikální význam, lexikalizace

4) syntax

- valence
- jednotlivé větné členy, výrazová a funkční charakteristika
- hierarchizace větné propozice
- syntaktické vztahy, způsoby a prostředky jejich realizace
- klasifikace souvětí
- aktuální větné členění
- negace
- nominální skupina

5) lexikologie

- lexikologie, předmět zájmu, členění, blízké disciplíny
- jednotky lexikální roviny jazyka
- lexikální sémaziologie (sémantika) a onomaziologie
- pojmenování: proces, rezultat
- autosémantika a synsémantika
- apelativa a propria
- termín, terminologie, terminografie
- přenesená pojmenování
- frazém-idiom, frazeologie a idiomatika, frazeografie
- lexikální význam a jeho složky
- možnosti popisu lexikálního významu
- významová vágnost, prototypová teorie

- významové vztahy strukturující lexikon: synonymie, opozitnost a její typy, hyperonymie – hyponymie, holonymie – meronymie ad.
- polysémie a homonymie
- slovní zásoba, její vrstvy, historické proměny, současné vývojové tendence
- lexikografie a její vztah k lexikologii
- typologie slovníků češtiny s uvedením hlavních reprezentantů od poč. 20. století do současnosti
- výkladový slovník češtiny: charakteristika, historie a současnost v českém prostředí
- počítačová lexikografie, pozitiva a negativa využití korpusových zdrojů pro lexikografii

6) stylistika

- definice stylu – 1) obecně, 2) v jazykových komunikátech
- stratifikace stylistiky
- klasifikace stylů – různé způsoby a kritéria
- funkční styly – obecná charakteristika (způsob vymezování, možný počet f. s., stylové normy jednotlivých f. s.)
- obecné vymezení stylové normy, v porovnání s jazykovou normou a komunikační normou
- stylové faktory (slohotvorní/stylové činitele)
- kompozice textu z hlediska stylistiky
- základní slohové postupy
- intertextovost a styl
- stylová charakteristika výrazových prostředků (typy stylových příznaků, stálá × kontextová stylová hodnota)

7) historie jazyka od praslovanštiny přes starou češtinu k nové češtině

- praslovanština, její časové a teritoriální vymezení
- klasifikace slovanských jazyků
- historické souvislosti vzniku staroslověnštiny, staroslověnské písemné soustavy
- tendence ke stoupavé sonoritě slabiky
- praslovanské palatalizace
- pozdně praslovanské změny, na jejichž základě se vydělila pračeština – kontrakce, zánik a vokalizace jerů, denazalizace

- vývoj pravopisných soustav češtiny
- vývoj fonologického systému češtiny od roku 1000 do roku 1400
- vývoj fonologického systému češtiny od roku 1400 do současnosti
- vývoj a proměny gramatických kategorií jmenných
- vývoj a proměny gramatických kategorií slovesných
- vývojové tendence české deklinace
- vývojové tendence české konjugace

8) obecně teoretické otázky

- jazykověda, její členění, metodologické postupy
- fylogeneze a ontogeneze jazyka, diferenciaci jazyků v čase, jazykové kontakty
- klasifikace jazyků podle genetického a typologického hlediska, zařazení češtiny
- indoevropské jazyky
- teorie jazykového znaku
- jazyk jako hierarchický systém, paradigmatické a syntagmatické vztahy
- komunikační model, jazykové funkce
- teritoriální a sociální stratifikace národního jazyka, nástin českých dialektů, problematika obecné češtiny
- úzus, norma, kodifikace, problematika spisovné češtiny
- historicko-srovnávací jazykověda – základní obrysy
- strukturalismus – základní obrysy
- pragmatický obrat v jazykovědě – základní obrysy

9) znalost významných osobností vývoje české jazykovědy

10) znalost základní literatury oboru český jazyk a nejdůležitějších periodik

11) znalost základní terminologie oboru český jazyk

12) základy matematiky a statistiky

- definice, věty, důkazy
- výroková logika, predikátová logika 1. řádu
- množiny, uspořádané množiny a operace na nich
- funkce a relace, jejich uzávěry
- posloupnosti a řady
- přirozená čísla a základní operace na nich
- kombinatorika a pravděpodobnost, podmíněná pravděpodobnost
- statistický soubor, charakteristiky polohy a variability
- statistika a pravděpodobnost: pravděpodob. rozložení, pravděpodob. prostor
- Bayesův vzorec a jeho aplikace
- Zipfův zákon, zákon velkých čísel

13) základy informatiky

- von Neumannův model počítače
- reprezentace čísel v počítači
- reprezentace znaků v počítači, znakové sady a jejich kódování
- algoritmy a datové struktury (pole, spojový seznam, fronta, zásobník)
- grafy, stromy, vlastnosti grafů, základní grafové algoritmy, prohledávání stromu do hloubky, do šířky
- základní pojmy z algoritmické složitosti (časová a prostorová složitost algoritmu, asymptotická složitost a třídy složitosti, složitost algoritmu a složitost problému)

14) počítačová lingvistika, korpusová lingvistika

- formální jazyky vs. přirozené jazyky
- Chomského hierarchie jazyků, gramatiky, automaty
- regulární výrazy a jejich využití ve zpracování přirozeného jazyka
- automatická morfologická analýza přirozeného jazyka
- automatická syntaktická analýza přirozeného jazyka
- formální gramatiky pro přirozené jazyky
- jazykový model, n-gramy
- využití statistiky ve zpracování přirozeného jazyka
- automatická sémantická analýza přirozeného jazyka

- komunikační agenty (chatbots)
- reprezentace znalostí (ontologie, wordnet, rámce)
- word sense disambiguation
- textové korpusy
- známé korpusy
- uložení korpusu v počítači
- indexování a prohledávání korpusu
- nástroje pro práci s korpusy
- CQL