

Posudek bakalářské práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce Kristýna Neumannová
Název práce Identification and analysis of Czech equivalents of German compounds
Rok odevzdání 2021
Studijní program Informatika **Studijní obor** Obecná informatika

Autor posudku Mgr. Magda Ševčíková, Ph.D. **Role** Vedoucí
Pracoviště Ústav formální a aplikované lingvistiky

Prosím vyplňte hodnocení křížkem u každého kritéria. Hodnocení *OK* označuje práci, která kritérium vhodným způsobem splňuje. Hodnocení *lepší* a *horší* označují splnění nad a pod rámec obvyklý pro bakalářskou práci, hodnocení *nevyhovuje* označuje práci, která by neměla být obhájena. Hodnocení v případě potřeby doplňte komentářem. Komentář prosím doplňte všude, kde je hodnocení jiné než *OK*.

K celé práci

	lepší	OK	horší	nevyhovuje
Obtížnost zadání	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Splnění zadání	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Rozsah práce ... <i>textová i implementační část, zohlednění náročnosti</i>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Komentář Cílem předkládané práce bylo navrhnout a implementovat postup, kterým se pro německá kompozita z existujících zdrojů na základě dat z paralelních česko-německých korpusů stanoví české protějšky různé komplexity. Kol. Neumannová pracovala se seznamem kompozit z německé databáze GermaNet, z paralelního korpusu získala ekvivalenty pro více než 50 tisíc německých kompozit. Automatická extrakce byla následována lingvistickou analýzou jednotlivých typů českých ekvivalentů. Kol. Neumannová víceslovné ekvivalenty třídila podle slovních druhů jednotlivých slov, u jednoslovných ekvivalentů pak kombinací dostupných nástrojů a vlastních metod zjišťovala jejich vnitřní strukturu a popisovala, jak morfémy v českém slově korespondují s částmi německého protějšku. Zadání práce tím bylo beze zbytku splněno.

Bakalářská práce je příspěvkem k aktuální datově založené diskuzi o konkurenci ve slovo tvorných systémech jednotlivých jazyků a mezi nimi. Navržený postup a vyhodnocení dat přináší poznatky pro zvolený jazykový pár, ale je relevantní i z obecnějšího, metodologického hlediska.

Textová část práce

	lepší	OK	horší	nevyhovuje
Formální úprava ... <i>jazyková úroveň, typografická úroveň, citace</i>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Struktura textu ... <i>kontext, cíle, analýza, návrh, vyhodnocení, úroveň detailu</i>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Analýza	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vývojová dokumentace	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Uživatelská dokumentace	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Komentář Práce má standardní strukturu. Dílčí kroky automatické identifikace českých ekvivalentů německých kompozit jsou popsány detailně, podrobná je i následující lingvistická analýza jednotlivých typů českých ekvivalentů. Text je napsán základní odbornou angličtinou – jde o první větší text, který studentka psala v angličtině.

Implementační část práce

	lepší	OK	horší	nevyhovuje
Kvalita návrhu ... <i>architektura, struktury a algoritmy, použité technologie</i>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Kvalita zpracování ... <i>jmenné konvence, formátování, komentáře, testování</i>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Stabilita implementace	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Komentář Implementace procedury pro identifikování českých ekvivalentů je rozvržena logicky. Oceňuji vlastní metodu pro lingvistický rozbor jednoslovných ekvivalentů (jejich rozlišení na kompozita a deriváty), jejíž výsledky kol. Neumannová kombinovala s výstupy dostupných, zatím pouze pilotních nástrojů. Získané výstupy byly zpracovány tak, že ústí v koherentní analýzu.				

Celkové hodnocení Výborně

Práci navrhuji na zvláštní ocenění Ne

Datum 21. června 2021

Podpis