

Přesný morfologický popis slovních tvarů je prvním předpokladem pro úspěšné automatické zpracování jazykových dat.

Systém kategorií a jejich hodnot, které se k popisu používají, jsou náplní první části práce.

Základním principem je tzv. Zlaté pravidlo morfologie, které říká, že každý slovní tvar by měl být v systému popsán jednoznačně. Existence variant na úrovni slovních tvarů i celých paradigmat však splnění tohoto pravidla komplikuje. Koncept variant rozložíme na tzv. mutace, mezi které řadíme i jiné množiny slovních tvarů se stejným popisem (např. různé tvary osobních zájmen). Mutace dělíme na globální pro popis na úrovni paradigmat a flektivní pro popis jednotlivých slovních tvarů. Toto rozdělení nám umožňuje postihnout jejich různé kombinace. Upoutáme od dělení variant (mutací) podle stylového příznaku jako neobjektivního kritéria. Při následném využívání hodnot kategorií Flektivní mutace a Globální mutace zůstane Zlaté pravidlo morfologie vždy splněno.

V kapitole o lemmatizaci zavádíme vícenásobné lemma pro popis variantních lemmat.

Podrobně se zabýváme popisem tzv. složenin, tedy slovních tvarů typu za, pro, koupilas, koliks. Pro jejich lemmatizaci rovnou využíváme konceptu vícenásobného lemmatu. Podle slovních druhů jejich slooek je dělíme na několik typů. Zabýváme se tímto problémem jejich vyhledávání v jazykových korpusech.

Druhá část práce popisuje systém vzorů pro popis slovních tvarů jednotlivých slovních druhů. U každého vzoru uvádíme sadu parametrů, které umožní postihnout velkou variabilitu v tvorbě konkrétních paradigmat. Vnucujeme se i pravidelnému odvozování příbuzných slov pomocí sufixů.