

V této práci popisujeme poloautomatickou metodu trénování pravděpodobnostního překladového slovníku z rozsáhlých automaticky anotovaných paralelních korpusů. Na základě studia překladových chyb a funkce slovníku v rámci překladového systému TectoMT obecně byly navrženy modely různé složitosti. Tyto základní modely byly zkombinovány do hierarchických modelů, jejichž účel je snížit dopad problému řídkých dat. Slovník byl doplněn o rozšíření, která jsou navržena tak, aby odstraňovala časté problémy lexikálního charakteru. Slovník spolu s rozšířeními byl na testovacích datech porovnán s původním slovníkem a výsledky ukazují, že došlo k zvýšení kvality překladu.