

Dr inż. Mirosław Gajer

Katedra Informatyki Stosowanej

Streszczenie

Lingwiści przewidują, że do końca bieżącego stulecia bezpowrotnie wymrze ponad połowa z istniejących obecnie na ziemi około 7000 języków. Obecnie proces wymierania języków przybrał bardzo na sile i szacuje się, że średnio co dwa tygodnie gdzieś na świecie wymiera jakiś język. Co gorsze, tego rodzaju procesów nie sposób już powstrzymać, w związku z czym w najbliższych dziesiątkach lat wiele języków o małej liczbie użytkowników jest skazanych bezpowrotnie na całkowite wymarcie.

Wymieranie języków stanowi niepowetowaną stratę dla ogólnoludzkiej kultury, ponieważ wraz ze śmiercią danego języka ginie bezpowrotnie związana z nim swoista kultura oraz specyficzny sposób widzenia i rozumienia świata, określane w językoznawstwie kognitywnym mianem językowego obrazu świata. Wymieranie licznych małych języków to także wielka strata dla językoznawstwa jako dyscypliny nauki, ponieważ tak znaczna utrata potencjalnego materiału badawczego utrudnia późniejszą weryfikację licznych hipotez lingwistycznych, dotarcie do naukowej prawdy i udzielenie satysfakcjonujących odpowiedzi na wiele fundamentalnych pytań, dotyczących przykładowo powstania i ewolucji języków ludzkich czy też istnienia w nich tzw. uniwersaliów językowych.

W związku z powyższym wydaje się rzeczą wysoce pożądaną, aby zachować w stopniu możliwie jak największym wiedzę o językach skazanych bezpowrotnie na wymarcie, tak aby w okresie późniejszym było możliwe przeprowadzenie przez specjalistów rekonstrukcji warstwy składniowej tych języków, a następnie podjęcie próby ich rewitalizacji w gronie entuzjastów tychże języków. Prowadzone dotychczas w tym zakresie działania sprowadzają się głównie do rejestracji warstwy leksykalnej zagrożonych wymarciem języków. Niestety, działania takie nie są wystarczające z punktu widzenia późniejszej rewitalizacji rozważanych języków, ponieważ nie pozwalają na rekonstrukcję ich morfologii i składni.

W przeciwieństwie do tego w referacie zaproponowano podjęcie czynnych działań zmierzających do zachowania wiedzy o wymierających językach. Do tego rodzaju działań autor zalicza tworzenie tzw. języków kontrolowanych na podstawie języków zagrożonych wymarciem, które następnie posłużą do budowy generatorów struktur syntaktycznych rozważanych języków. Z kolei wymienione generatory struktur syntaktycznych stanowiąc będą podstawę do budowy systemów klasy Machine-Aided Human Translation, za pomocą których użytkownik może tworzyć zdania w języku obcym, którego w zasadzie może w ogóle nie znać bądź posiadać na jego temat jedynie pewne szczątkowe informacje.

Obecnie w obszarze zainteresowań autora znajdują się północnogermańskie języki skandynawskie. Aktualnie generatory struktur syntaktycznych oraz opierające się na nich systemy należące do klasy Machine-Aided Human Translation są rozwijane przez autora dla języków takich jak szwedzki, duński i norweski (ten ostatni w wersjach językowych: bokmål, riksmål i nynorsk). Uzyskane na tym polu rezultaty zostaną następnie ekstrapolowane na należące do wymienionej grupy języki potencjalnie zagrożone wymarciem, takie jak islandzki i farski oraz wymarły już od ponad wieku północnogermański język norn, dla którego opracowywany jest aktualnie nowy standard na podstawie analizy licznych silnie zróżnicowanych jego dialektów, określane mianem nowego języka norn – tzw. nynorn.