

Měření podobnosti překladů básně Havran

Jaroslav Marek, Jan Šlahora

Univerzita Pardubice, Fakulta elektrotechniky a informatiky, náměstí Čs. legií 565, 530 02
Pardubice, email: jaroslav.marek@upce.cz

Klíčová slova: matematická lingvistika, fónické jevy, agregace, aliterace, podobnost textů, shluková analýza, metoda hlavních komponent, konfidenční oblast, E. A. Poe, překlady básně Havran

Abstrakt:

V článku upozorníme na možnosti využití statistických metod a matematické lingvistiky pro měření podobnosti textů. Pomocí lingvistických charakteristik budeme měřit podobnost básní a pokusíme se o rozpoznávání autorství. K měření podobnosti budou použity metody shlukové analýzy a metoda hlavních komponent, viz [6]. Výpočty budou realizovány na několika překladech básně E. A. Poea Havran.

Analýza textů zahrnuje velké množství jevů, které je možné sledovat, a to na různých textových útvarech. Základní stavební kámen každého textu je písmeno. U písmen můžeme sledovat například jejich frekvenci výskytu v textu. Tato frekvence může být využita v dalších výpočtech u jiných jevů. Dále se zejména u poezie sledují fónické jevy, které zkoumají tzv. libozvuk – eufonii. Ten je dán uspořádáním a následným opakováním hlásek (především samohlásek) ve verši. V této práci budeme z fónických jevů měřit agregaci a aliteraci, viz [2, 3]. Dále se používá také asonance, se kterou ale pracovat nebudeme.

Můžeme si položit otázku: "Lze pomocí lingvistických jevů rozlišit autorství textu, resp. rozpoznat, jakým jazykem je text napsaný?"

Podobnost básní je diskutována v části fiktivního rozhovoru antických filosofů Sokrata a Platóna v knize Dialogy o matematice Alfreda Rényiho.

Sokrates: „ ... žádní dva básníci, kteří se neznají a navzájem o sobě nic nevědí, nenapišou nikdy – ani o téže věci – stejnou báseň? “

Hippokrates: „ Nevím. To nedokážu vysvětlit. ... že by dva básníci napsali stejnou báseň, to jsem opravdu ještě nikdy neslyšel. “

Odpověď se pokusíme najít pomocí metod mnohorozměrné statistické analýzy aplikovaných na vypočtených lingvistických charakteristikách různých překladů básně Havran od E.A. Poea.

Nejprve pomocí vytvořené počítačové aplikace určíme u zvolených textů délku věty, frekvenci slov o zvolené délce, hodnoty agregací, hodnotu aliterace.

Měření a shlukování provedeme pro následující české překlady: 1) Šembera V.K., 2) Vrchlický J. 1890, 3) Mužik A.E., 4) Lutinov K.D., 5) Nezval V., 6) Babler O.F., 7) Taufer J., 8) Stoklas E., 9) Wagnerová D., 10) Havel R., 11) Čapek J. B., 12) Resler K., 13) Černý R., 14) Slavík I., 15) Kadlec S., 16) Bejblík A.

Charakteristiky budou určeny pro celé překlady a také pro 18 strof, ze kterých se skládají.

U získaných lingvistických charakteristik ze všech 18 strof bude pro všechny překlady provedeno hierarchické shlukování. Výsledek shlukování, kdy budou jako nejpodobnější vyhodnoceny překlady J. Vrchlického a E.A. Mužika, není překvapivý. Alois Bejblík v předmluvě v [1] na str. 23 píše o E.A. Mužikově překladu: „*Jeho verš a vyjadřovací způsoby vyšly z Vrchlického básnické školy.*“ Od ostatních překladů se v dendrogramu nejvíce odlišoval Stoklasův překlad. I tento výsledek mohl být očekáván. O jeho překladu A. Bejblík napsal: „*Stoklas cítil příbuzenství Havrana s Máchou a Erbenovým vodníkem.*“ Baladický charakter básně činí překlad odlišným.

Na měření u vybraných strof bude dále aplikována metoda hlavních komponent, viz [6]. Pro jednotlivé překlady určíme z vybraných strof konfidenční elipsy, viz [4, 5] pro jejich první dvě hlavní komponenty. Potom určíme a znázorníme hodnoty hlavních komponent pro strofy, které nebyly použity k výpočtu. Pro tyto strofy budeme sledovat, zda se jejich hlavní komponenty zobrazily do konfidenčních elips správných autorů a zda se nám podařilo úspěšně přiřadit autorství.

Literatura:

- [1] POE, E.A.: Havran, šestnáct českých překladů, ODEON, Praha, 226 s. 1990
- [2] WIMMER, G., ALTMANN, G., HŘEBÍČEK, L., ONDREJOVIČ, S., WIMMEROVÁ, S., Úvod do analýzy textov, VEDA, Bratislava, 345 s. 2003.
- [3] ANDRES, J., KUBÁČEK, L., MACHALOVÁ, J., TUČKOVÁ, M.: Optimization of parameters in the Menzerath–Altmann law, Acta Univ. Palacki. Olomuc., Fac. rer. nat., Mathematica 51, 1 (2012), p. 5-27
- [4] WIMMER, G., PALENČÁR, R., WITKOVSKÝ, V. Spracovanie a vyhodnocovanie meraní. Bratislava: VEDA, 2002. 187 s. ISBN 80-224-0734-8
- [5] KUBÁČEK, L., KUBÁČKOVÁ, L. Statistika a metrologie. Vydavatelství Univerzity Palackého v Olomouci, Olomouc 2000
- [6] ŘEZANKOVÁ, H., HÚSEK, D., SNÁŠEL, V. Shluková analýza dat, PBtisk Příbram, 2007. 196 s. ISBN 978-80-86946-26-9