

A negatív emotív elemek vizsgálata a nemek közötti nyelvhasználati különbségek szempontjából

Bevezetés

Az előadás a negatív emotív szemantikai tartalmú elemeket vizsgálja egy magyar nyelvű twitter-korpuszban, a női és férfi nyelvhasználati különbségekre fókuszálva.

A vizsgálat tárgyát azok az elemek képezik, amelyek lexikai szinten negatív emotív szemantikai tartalommal rendelkeznek ugyan, azonban fokozó funkcióban képesek azt részben vagy egészben elveszíteni (pl. *mocskos nagy*, *rohadt jól*), sőt bizonyos esetekben arra is, hogy negatív tartalmuk ellenére pozitív értékelést fejezzenek ki (pl. *brutális volt a hétfégi koncert*) (vö. Szabó 2015). Az előbbi esetben a vizsgált elemek más, értékelő kifejezések jelentését fokozzák (ezt a jelenséget nevezzük *értékvesztésnek*), az utóbbi esetben pedig maguk töltik be értékelő kifejezés (szentimentkifejezés) funkcióját (ezt a jelenséget nevezzük *lexikai szintű értékváltásnak*) (vö. Szabó et al.).

A vizsgált jelenség mind elméleti, mind alkalmazott nyelvészeti, különösen nyelvtechnológiai szempontból egyaránt figyelemre méltó. Kimerítő elméleti nyelvészeti leírása még nem történt meg, hatékony nyelvtechnológiai kezelése pedig egyelőre nem megoldott (vö. pl. Andor 2011; Kugler 2014; Szabó 2015; Drávucz et al.). Ezekre az elemekre ugyanis a szótáralapú automatikus szentiment- és emócióelemzés – a lexikai szintű polaritásuk miatt – téves elemzési eredményt ad.

A jelen előadásban a következőket tárgyaljuk részletesen: 1) bemutatjuk a vizsgált nyelvi jelenséget, és az irodalom eddigi vonatkozó megállapításait; 2) ismertetjük a vizsgálati korpusz létrehozásának a menetét, módszerét és eszközét; 3) bemutatjuk az elemzés részletes, kvalitatív és kvantitatív eredményeit, a korpuszból vett példákkal szemléltetve a legfontosabb kutatási tapasztalatokat.

A kutatás módszere

A nyers korpusz, amelyből a kutatáshoz szükséges anyagot gyűjtöttük, összesen 37818 magyar nyelvű twitter-bejegyzésből áll, ebből 19580 tweet női, 18238 pedig férfi szerzőségű. A korpusz azon nyelvi adataira volt szükségünk, amelyekben szerepelt legalább egy, lexikai szintű értékváltásra vagy értékvesztésre képes elem. Ehhez a feladathoz kézzel összeállítottunk egy, összesen 109 elemből álló listát, amely lexikai szintű értékváltásra vagy értékvesztésre képes elemeket tartalmaz (vö. Szabó et al.). Ezt követően, a lista segítségével kiszűrtünk minden olyan tweetet, amely tartalmazta a listán szereplő elemek valamelyikét. Az így létrehozott korpusz összesen 610 tweetből áll.

A 610 tweetből álló korpuszt a vizsgálat céljainak megfelelően, szemantikai–pragmatikai szempontokat figyelembe véve annotáltuk (vö. Szabó et al.). A feldolgozó munka során bejelöltük azokat az elemeket, amelyek definíciónk szerint lexikai szintű értékváltást vagy értékvesztést mutattak. Az értékvesztésre vonatkozóan további tageket is bevittünk, az annotált elemek szemantikai–pragmatikai viselkedése alapján. Mindemellett

feldolgoztuk a vizsgált elemek által módosított szentimentkifejezéseket, illetve targeteket (az értékelések tárgyát, célpontját) is.

Végül, a létrehozott korpusz alapján részletes vizsgálat tárgyává tettük az adatokat a nemek közötti nyelvhasználati eltérések szempontjából.

A vizsgálat legfontosabb hipotézisei

A vizsgálat alapvető hipotézise az volt, hogy a negatív emotív elemek használatában jelentős nemek közötti eltérést fogunk tapasztalni, és azt feltételeztük, hogy ez az eltérés egyrészt a betöltött funkciók gyakorisági megoszlását, másrészt az egyes elemek frekvenciális sajátosságait is érinteni fogja. Másképpen, a férfiak és a nők nem azonos gyakorisággal élnek a vizsgált kifejezésekkel fokozó elem, valamint szentimentkifejezés funkciójában, továbbá nem azonos gyakorisággal használják az egyes elemeket sem.

A kutatás eredményei

A vizsgálat igazolta a hipotéziseinket, és további figyelemre érdemes sajátosságokat is feltárt. Példaként említsük meg a kutatás következő tanulságait: 1) A nők jóval több elemet használnak a vizsgált elemek közül fokozó funkcióban a férfiaknál, és ezek gyakorisági megoszlása is kiegyenlítettebb. Velük szemben a férfiak néhány elemet preferálnak, a többi a vizsgált elemek közül vagy csak ritkán, vagy egyszer sem fordult elő a korpuszban. 2) A nők gyakran használják a vizsgált elemeket szentimentkifejezés funkciójában valódi értékelési cél nélkül, pusztán a meglepetés érzelm kifejezésére (pl. *dedurvaaa micsoda egybeesések*). Velük szemben a férfiak nyelvhasználatára ez egyáltalán nem jellemző. Amennyiben ők valamely vizsgált elemet ebben a funkcióban használják, azzal vagy negatívan, vagy pozitívan, de mindenképpen értékelnek.

IRODALOM

- Andor J. 2011. De durva ez a téma! – Megfigyelések a melléknévi polaritásváltásról. In *Hungarológiai Évkönyv* **12**. 33–42.
- Drávucz, F., Szabó, M. K., Vincze V. Szentiment- és emóciósztárak eredményességének mérése emóció- és szentimentkorpuszokon. Benyújtva: *XIII. Magyar Számítógépes Nyelvészeti Konferencia*. Szeged, 2017. január 26-27.
- Kugler N. 2014. A nyelvi polaritás kifejezésének egy mintázata, avagy milyen a félelmetesen jó? *Magyar Nyelvőr* **138**(2). 129–139.
- Szabó M. K. 2015. A polaritásváltás problémája a szentimentelemzés szempontjából. In: Váradi T., ed. *IX. Alkalmazott Nyelvészeti Doktoranduszkonferencia konferenciakötete*. Budapest, MTA Nyelvtudományi Intézet. 51–61.
- Szabó M. K., Nyíri Zs., Morvay G., Lázár B. A polaritásmódosulás vizsgálata szemantikai–pragmatikai szempontból annotált korpuszon. Benyújtva: *XIII. Magyar Számítógépes Nyelvészeti Konferencia*. Szeged, 2017. január 26-27.