

B"H

## **Egy kvantitatív szemiotika felé?**

Takács Boglárka

Bécsi Orvostudományi Egyetem, Ausztria / Mikata Kft.

Hogyan lehet megállapítani egy jelsorozatról, hogy természetes nyelvi jellegű-e? Egyáltalán lehetséges ez, el tudjuk-e valamilyen kvantitatív eljárással különíteni a természetes nyelveket nem-nyelvi szimbólumsorozatokról, vagy olyan több tekintetben hasonló mesterséges struktúráktól, mint például a programozási nyelvek?

Az elmúlt két évben több számítógépes nyelvészeti módszer is született a probléma megoldására, de ezek mind hagynak kívánnivalókat maguk után. A kritikák eddig főleg arra fókuszáltak, hogy olyan sorozatok is előállíthatóak, amelyek bizonyíthatóan nem természetes nyelvek, de mégis megfelelnek az adott módszerek kritériumainak. Viszont meglehetősen keveset tudunk arról, milyenek a való életben fellelhető, de nem természetes nyelvi szimbólumsorozatok *nyelvészeti* szempontból, annak ellenére, hogy a modern számítógépes nyelvészet, korpusz nyelvészet stb. eszközeivel számszerűsíthető következtetéseket is tehetünk. A jelekkel általánosságban foglalkozó tudományág, a szemiotika jelenleg leginkább - habár nem kizárólagosan - kvalitatív módszerekkel dolgozik, ami szintén nagyon fontos, de ebben a konkrét kérdésben kevésbé visz minket előre.

Tud-e újat mondani a jelekről a nyelvészet? Esetleg csak annyiban, hogy felhívja a figyelmet arra, hogy különböző alkalmazott matematikai eljárások, amelyek a természetes nyelvek esetében beváltak, más szimbólumrendszerek esetében is hasznosak lehetnek? Vagy akár konkrét nyelvészeti elméletek, nyelvmodellek is bevethetőek e célra? Ha igen, hogyan?

Az előadáson szó lesz az aktuális helyzet összefoglalása és a definíciós kérdések mellett néhány konkrét érdekes jelenségről és gyakorlati alkalmazási lehetőségről is.