

Szótárfejlesztés a MetaMorpho fordítóprogram keretében: elméleti és gyakorlati problémák

A számítógépes fordítóprogramok egyik legfontosabb alkotóeleme a szótár maga. Ennek minősége és mérete nagyban befolyásolja a fordítás eredményességét, illetve elfogadhatóságát. Előadásunkban a MetaMorpho magyar-angol fordítóprogram kétnyelvű szótárának előállításánál során tapasztalt elméleti és gyakorlati problémákról számolunk be.

A Szegedi Tudományegyetem Informatikai Tanszékcsoportjának Nyelvtechnológiai Csoportjában végzett munkánk során a szótár bővítésének keretein belül a főnévi csoportok (NP), melléknévi csoportok (ADJP) és kollokációk angolra fordításával foglalkozunk. Előadásunkban elsőként példákon keresztül ismertetjük a fordítószabályok formátumát és a benne foglalt információkat. Mivel az automatikus fordítás legtöbbször nem kínált fel elfogadható angol megfelelőt, magunknak kellett megkeresni a helytálló angol kifejezést. Ennek során számos elméleti és gyakorlati problémával és nehézséggel találtuk magunkat szemben, melyeket példákkal illusztrálunk.

A szótárakkal kapcsolatban a következő problémakörök merültek fel. Az egynyelvű szótárakban található körkörös definíciók nem segítik elő az adott szó értelmezését, így az ekvivalens keresését sem könnyítik meg (*pedagógus-továbbképzés*). Gondot okoz az is, ha a keresett szó nem szerepel egyik kétnyelvű szótárban sem, ilyenkor különösen nehéz kiindulási pontot találni a fordításhoz (*rottyogás*). Találkoztunk olyan esettel is, amikor a kétnyelvű szótár teljesen félrevezető ekvivalenst adott meg, mely tényre csak az egynyelvű szótárban történt ellenőrzés derített fényt (*gazdasági épület – outhouse*). Külön problémát jelentenek azok az esetek is, amikor a magyar-angol szótárban található angol szó fordítása nincs átfedésben az eredeti magyar szóval (*venyigés – sarmentose – indás*).

A fordítás során más jellegű problémákkal is találkoztunk. A többjelentésű szavak sajátossága, hogy bizonyos esetekben nem tudtuk egyértelműen eldönteni, a szónak melyik jelentése a (leg)gyakoribb (*ráhajtás, partnercsere, fazon*). A kultúraspecifikus szavak fordítása is igen nehézkes, hiszen sokszor a másik nyelvterület nem is ismeri az adott dolgot a maga valójában vagy pedig teljesen más képzeteket társít hozzá (*máglyarakás, tanyarendszer*).

A melléknévi kifejezések fordítása során felmerült problémákat hat csoportba osztottuk: (1) önmagukban sohasem fordulnak elő (*nevű, színű*); (2) egyik jelentésükben csak más szó kíséretében fordulnak elő, de másik jelentésük lefordítható (*éves*); (3) a melléknév fordítása attributív vagy predikatív pozícióban eltérő (*kérdéses*); (4) kettős szófajú szavak (*angol*); (5) gyakran olyan szókapcsolatban fordul elő, amelynek angol megfelelője egy szó (*nyolcvanas (évek) vs. eighties*) és (6) a melléknév pontos fordítása az utána következő főnévtől függ (*nagy, szép*).

A szótárban kollokációk is szerepelnek. Ezek fordítása a legnehezebb, hiszen a kollokációk nem teljes mértékben kompozicionálisak, ebből következően a kifejezés részeinek lefordításából előállt szókapcsolat a legtöbb esetben nem tekinthető a kifejezés angol megfelelőjének (*tanítóképző főiskola – derivational suffix college – teacher training college*). A kollokációk fordítása automatizálható a legkevésbé, ezért itt a legnagyobb a jelentősége a helytálló szótári információknak.

Előadásunk végén néhány, internetes fordítóprogrammal előállított példamondaton keresztül szemléltetjük, hogy miért is van jelentősége a megfelelő szótáraknak a gépi fordításban.