

Igei igenevek problémája számítógépes nyelvészeti szempontból

Nádasdi Péter

Szegedi Tudományegyetem, Általános Nyelvészeti Tanszék
H-6722 Szeged, Egyetem u. 2.
nadasdi_peter@freemail.hu

Kivonat: Az igei igenevek nemcsak a leíró és elméleti nyelvészeti kutatások számára jelentenek problémát, hanem a gépi fordítás számára is, ezért fontos követelmény, hogy nem szavanként, hanem egységes szerkezetként kezeljük a vizsgált konstrukciót, így a gépi program csak azokat a szerkezeteket jelenítené meg, amelyek beletartoznak ebbe a típusba, de azok közül mindet visszaadná.

1 Bevezetés

A mai magyar nyelvben létezik egy olyan szerkezet (például *a Vágó István vezette vetélkedő*), amely egy főnév (*Vágó István*) és egy *-t/-tt* toldalékos verbális elem (*vezette*) szoros kapcsolatából áll, és ez a két elem együtt egy másik főnevet (*vetélkedő*) módosít. Feltűnő tulajdonsága, amely elkülöníti más jelzős és/vagy igeves szerkeztől, hogy személyjelölő található a szerkezet igei tagján: *vezett-e*. A szerkezet a gépi fordítás számára kihívást jelent, hiszen nincs egységes elképzelés e szerkezetről számítógépes nyelvészeti szempontból, ezért a szintaxis alkalmazása fontos tényező a vizsgált szerkezet számára.

2 A szerkezet jellemzői

Két fő elképzelés létezik erről a szerkezetről a generatív nyelvészeti szakirodalomban: ige (mai egyedüli képviselője [3]) és igeves [1, 2], attól függően, hogy a szerkezet verbális elemét (*vezette*) igeinek vagy befejezett melléknévi igenévnek tartják. A szerkezet általános jellemzőit a következőkben foglalhatjuk össze:

1. A szerkezet egy nominális elem és egy verbális elem kapcsolata, amelyek mindig ebben a sorrendben követik egymást: *Vágó István* és *vezette*.
2. Egy *-t/-tt* toldalékos verbális elem található a szerkezetben: *vezet-t-e*.
3. Egy esetjelölés nélküli nominális elem van a szerkezetben: *Vágó István*.

4. A szerkezet több szóból is állhat, de tartalmaznia kell legalább egy nominális és egy verbális elemet, egyik vagy másik nélkül agrammatikus a szerkezet: *Vágó István vezette*¹.
5. Személyjelölő található a szerkezetben a verbális elemen: *vezett-e*.
6. Jelzői szerepben áll egy módosított főnév (antecedens) előtt: (*Vágó István vezette*) *vetélkedő*.
7. A szerkezet két tagja között predikatív viszony van, azaz az esetjelölés nélküli nominális elem az alanyi argumentuma a verbális elemnek: a *Vágó István* a *vezette* elemnek.
8. A vizsgált szerkezetben tárgyias vagy alkalmilag tárgyias (egynél több argumentumú) igék fordulnak elő, például a nem tárgyias *lakik* is.
9. A szerkezetet általában névelő vezeti be, amely lehet határozott vagy határozatlan is, de alkalmilag el is maradhat: (*a*) *Vágó István vezette vetélkedő*.

A vizsgált konstrukció szintaktikai szerkezete tehát az alábbiak szerint ábrázolható:

Névelő	alárendelt (jelzői) tagmondat	módosított főnév (antecedens)
<i>a</i>	[<i>Vágó István</i> (tárgy \emptyset_1) <i>vezet-t-e_i</i>]	<i>vetélkedő</i>

A személyjelölő tárgyias személyragként értelmezhető, amely egy rejtett névmásra referál (\emptyset), és ez a rejtett névmás a módosított főnevet (*vetélkedő*) képviseli tárgyként a jelzői (alárendelt) tagmondatban.

3 Gépi fordítás

A szerkezet gépi fordítására rányomja a bélyegét, hogy nincs egységes elképzelés e szerkezetről, és nem kompakt egészként tekintenek erre a konstrukcióra. Az annotálása sem egyöntetű a szerkezet verbális tagjának, hiszen az MNSZ-ben ige-névnek is jelölik, vagy éppen ismeretlen kategóriának, ezért fontos lenne egységes jelölést alkalmazni. A félig kompozicionális főnév + ige szerkezetekkel mutat nagyfokú rokonságot ez a szerkezet [4], hiszen a kérdéses konstrukció is egy főnév (*Vágó István*) és egy ige(név) (*vezette*) kapcsolata: egy tagmondatot alkot (*Vágó István vezette*), és a szerkezet jelentését csak e tagmondat és a módosított főnév (*vetélkedő*) viszonyaként adhatjuk meg. Produktivitása és szintaktikai önállósága ellenére léteznek olyan szókapcsolatok, amelynek szavai gyakran fordulnak elő együtt, ezért is tekinthetők kollokációknak, mint például az *X.Y. vezette kormány, valamilyen természeti erő sújtotta terület*. A számítógépes nyelvészeti alkalmazásoknál is a legnagyobb problémát a kollokációk, illetve a kollokációszerű szerkezetek jelentik [4]. Nagyban megkönnyítené a gépi fordítását, ha nem szavanként, hanem egységes szerkezetként

¹ Legfeljebb igekötő (igemódosító) és tagadószó kerülhet a nominális és a verbális elem közé, mint például *a Lajos felásta föld* vagy *az ember nem lakta föld*, de adjunktum nem. A személyes névmásos alakoknál a személyes névmás ugyan elmaradhat, de ezek már nem produktívak, mint például *a szerette város*.

kezelnénk, így csak azokat a szerkezeteket jelenítené meg a gépi program, amelyek beletartoznak ebbe a konstrukcióba, de azok közül mindet visszaadná. A szerkezet ugyanis két részből áll: 1. egy beágyazott tagmondat (*Vágó István vezette*), 2. egy módosított főnév (*vetélkedő*). Természetesen fontos kiemelni, hogy ez a szerkezet nagyon közel áll a nem személyjelölt befejezett melléknévi igenes szerkezetekhez, és ez a konstrukció könnyen átalakítható nem személyjelöltté: *a Vágó István által vezetett vetélkedő*. Éppen ezért idegen nyelvre, például angolra való fordításánál figyelembe kell vennünk, hogy kétféleképpen tudjuk elérni, hogy jó fordítás jöjjön létre: 1. posztnominális vonatkozó szerkezettel (angolban *wh*-konstrukció vagy *that*), 2. nem személyjelölt melléknévi igenes szerkezettel (angolban *past participle+by*-os szerkezettel). A (*Vágó István vezette*) *vetélkedő* esetében a két változat az alábbiak szerint fordítható: 1. 'Az a vetélkedő, amelyet Vágó István vezet' = 'The quiz show (**that/which**) István Vágó presents.' 2. 'A Vágó István által vezetett vetélkedő' = 'The quiz show presented **by** István Vágó.'

A fenti konstrukció szintaktikai elemzése két fő részre osztja a szerkezetet: 1. *a vetélkedő*, 2. *Vágó István vezette*, ahol 1. egy DP, amelybe be van ágyazva egy többnyire két elemből álló tagmondat, amelynek az első eleme a beágyazott tagmondat alanya, a második eleme pedig az állítmánya. Fontos kiemelni, hogy vonatkozó szerkezetként való fordításánál az igei elemet a múlt idejű (befejezett melléknévi igenévi) forma ellenére is jelen idővel kell fordítani: '*amelyet rendszerint szokott vezetni*' (Simple Present). A gépi fordításnál elengedhetetlen, hogy a program felismerje, hogy az adott főnév és igei elem összetartozik [4]. Ennél a szerkezetnél pedig ezután azt is kódolnia kell, hogy a módosított főnév is a szerkezet része, de nem annak argumentuma, hiszen a módosított főnév bármilyen esetet felvehet, aszerint, hogy milyen mondatba foglaljuk bele: *Láttam a Vágó István vezette vetélkedőt*. Ugyanis ez az ige+főnév szerkezet egy beágyazott tagmondat, amely a főnevet (antecedens) módosítja. A Google fordító és a *webforditas.hu* a példamondatunk fordításakor rosszul formált szerkezetet hoz létre, viszont például *az Angela Merkel vezette kormány* esetében a két program közül a Google fordító jó fordítást hoz létre: *the government led by Angela Merkel*, de csak az igenevest alkalmazza, viszont ha kicseréljük a módosított főnevet az alábbi módon: *az Angela Merkel vezette vetélkedő*, akkor már rossz eredményre jut. Feltételezhető, hogy nem szerkezeti elemzéssel, hanem heurisztikákkal dolgozik, és statisztikai alapon jut jó megoldásra.

4 Összegzés

A cikkben azt a problémát vizsgáltam, hogy a kérdéses szerkezet kihívást jelent a számítógépes nyelvészeti alkalmazás szempontjából a gépi fordítás, illetőleg az annotálás számára, és a szintaxis alkalmazását tartom fontos lépésnek a konstrukció szempontjából a nehézséget jelentő megoldandó feladatok számára.

Hivatkozások

1. Kenesei, I.: On the Role of the Agreement Morpheme in Hungarian. *Acta Linguistica Hungarica* Vol. 36 (1986) 109–120
2. Laczkó T.: Néhány lexikai-funkcionális gondolat a személyragos *-(t)T* igenévről. In: Andor J. et al. (szerk.): Színes eszmék nem alszanak... Szépe György 70. születésnapjára. *Lingua Franca* Csoport, Pécs. (2001) 741–759
3. Nádasdi P.: Mozgatás nélkül? Egy prenominális vonatkozó szerkezet minimalista elemzése a magyarban. In: Gárgyán G., Sinkovics B. (szerk.): *LingDok5. Nyelvész-doktoranduszok dolgozatai*. JATEPress, Szeged (2006) 61–86
4. Vincze V.: Félig kompozicionális főnév + ige szerkezetek a Szeged Korpuszban. In: Tanács A. et al. (szerk.): VI. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2009). JATEPress, Szeged (2009) 390–393