

LINGDOK 15.

NYELVÉSZDOKTORANDUSZOK DOLGOZATAI

Szerkesztette: Gécseg Zsuzsanna

Szegedi Tudományegyetem
Nyelvtudományi Doktori Iskola
Szeged
2016

Sorozatszerkesztő
Kenesei István

Technikai szerkesztő
Nagy Ágoston

Letölthető az SZTE Nyelvtudományi Doktori Iskolájának honlapjáról:
http://nydi.bibl.u-szeged.hu/SZTE_NYDI/

ISSN 2416-0601

TARTALOM

A szerzők adatai.....	5
Előszó.....	7
Andrási Krisztina: Magyar infintívuszok: logoforikus és predikatív kontroll.....	9
Georgieva, Ekaterina: Kített alanyok az udmurt nem véges mellékmondatokban.....	31
Georgieva, Ekaterina & Ótott-Kovács Eszter: Nem véges alárendelői mellékmondatok a Volga-vidéki nyelvi areában.....	51
Horváth Márton Gergely: A szabad topik a beszélt franciában.....	69
Huszthy Bálint: Újraértelmezhető-e az olasz nyelv fonológiája az olaszok idegen akcentusán keresztül?.....	95
Nagy Anna: A diskurzusjelölők szerepe humoros szövegekben.....	117
Pintér Lilla: A <i>csak</i> partikula értelmezése óvodáskorban.....	133
Szabó Martina Katalin: A nyelvi értékelés mibenléte a számítógépes értékeléselemzés (szentimentelemzés) szempontjából.....	153
Szalay Tünde: Az /l/ hang és az /l/-törlés a magyar nyelvben.....	173
Szőke Bernadett: Az azonosító értelmező vizsgálata a magyarban.....	191
Turi Gergő: Kétféle unakkuzatív predikátum: részkimozgatás magyar alanyokból.....	213

A SZERZŐK ADATAI

ANDRÁSI Krisztina: *ankrisztina@gmail.com*
SZTE Nyelvtudományi Doktori Iskola
Elméleti nyelvészet doktori program

GEORGIEVA Ekaterina: *katina.geo@gmail.com*
SZTE BTK Nyelvtudományi Doktori Iskola
Finnugor nyelvészet doktori program

HORVÁTH Márton Gergely: *horvath.marton.gergely@btk.ppke.hu*
ELTE Nyelvtudományi Doktori Iskola
Romanisztika program

HUSZTHY Bálint: *huszthy.balint@gmail.com*
PPKE Nyelvtudományi Doktori Iskola
Romanisztika műhely

NAGY Anna: *nagyanna.de@gmail.com*
Debreceni Egyetem Nyelvtudományok Doktori Iskola
Angol nyelvészet alprogram

ÓTOTT-KOVÁCS Eszter: *okeszter@gmail.com*
SZTE BTK Nyelvtudományi Doktori Iskola
Altajisztika doktori program

PINTÉR Lilla: *pinterlilla87@gmail.com*
PPKE Nyelvtudományi Doktori Iskola
Elméleti nyelvészeti műhely

SZABÓ Martina Katalin: *szabomartinakatalin@gmail.com*
SZTE Nyelvtudományi Doktori Iskola
Magyar nyelvészet doktori program

SZALAY Tünde: *tnd.szalay@gmail.com*
ELTE anglisztika mesterképzés / elméleti nyelvészet szakirány

A szerzők adatai

Szóke Bernadett: *bettyszoke@gmail.com*
SZTE Nyelvtudományi Doktori Iskola
Elméleti nyelvészet doktori program

Turi Gergő: *gergo.turi@gmail.com*
PPKE BTK Nyelvtudományi Doktori Iskola
Elméleti nyelvészet műhely

A NYELVI ÉRTÉKELÉS MIBENLÉTE A SZÁMÍTÓGÉPES ÉRTÉKELÉSELEMZÉS (SZENTIMENTELEMZÉS) SZEMPONTJÁBÓL¹

Szabó Martina Katalin

Bevezetés

Dolgozatomban a számítógépes értékeléselemzés (*sentiment analysis*) leg-alapvetőbbnek tekinthető dilemmáját, a nyelvi értékelő jelentés mibenlétét tárgyalom részletesen.

A *szentimentelemzés* a számítógépes nyelvészet egy részfeladata, amely arra irányul, hogy az értékelő kifejezéseket és az értékelő tartalmakat megtalálja a szövegekben, meghatározza ezeknek az értékeknek a típusát, vagyis pozitív vagy negatív voltát (számítógépes nyelvészeti terminussal: *polaritását*)², esetlegesen e polaritás intenzitásának a mértékét, valamint hogy megállapítsa az értékelés tárgyát (számítógépes nyelvészeti terminussal: *targetét*), tehát azt, hogy az értékelés mire irányul. Mindezt a természetesnyelv-feldolgozás (NLP) kínálta automatikus eszközökkel igyekszik megvalósítani. A feladat az alábbi egyszerű példával szemléltethető:

(1) a. A szentimentelemzés nagyon bonyolult.

Az (1) alatti példa a szentimentelemzésről fogalmaz meg értékelést, hiszen nagyon bonyolultnak minősíti azt. A *target* ebben az esetben tehát a *szentimentelemzés*, amely a kifejezésben negatív értékelő minősítést kap, s ezt a minősítést a *nagyon* elem fokozza. A kifejezés szentimentszerkezete a következőképpen sematizálható:

(1) b. [_{target} A szentimentelemzés] [_{szent} [_{int} nagyon] bonyolult].

A szentimentelemzés a nyelvtechnológiai kutatások palettáján még viszonylag új keletűnek tekinthető, ugyanakkor egyre nagyobb figyelmet kap annak nyelvészeti kihívásai, valamint gazdasági hasznosítási lehetőségei miatt (vö.

¹ Köszönettel tartozom témavezetőimnek, Bibok Károlynak és Vincze Veronikának a dolgozatírás során nyújtott hathatós segítségükért, valamint Várad Tamásnak és névtelen bírálómnak, hogy értékes megjegyzéseikkel hozzájárultak a munkámhoz.

² A polaritás fogalma a nyelvtechnológiában nem azonos azzal a polaritás-fogalommal, amelyet az elméleti nyelvészet alkalmaz (vö. Neubarth 2006; Moghaddam–Popowich 2010). A szentimentelemzésben az adott kifejezés szentimentértékére utal.

Pang–Lee 2008; Liu 2012). Ami a magyar nyelvű szövegek szentimentelemzését illeti, a feladattal behatóbban, nyelvészeti szempontokat is figyelembe vevő vizsgálat keretében csekély számú dolgozat foglalkozik (vö. Berend–Farkas 2008; Miháltz 2010, 2013; Hangya et al. 2015). Problémát jelent továbbá, hogy ezek a dolgozatok sem tárgyalják a feladat elméleti vonatkozásait, illetve nem nyújtanak kellő információt az elméleti kérdések vonatkozásában alkalmazott megoldásokról. Ugyanakkor mindezek tárgyalása nélkülözhetetlen egy olyan elemző rendszer létrehozásához, amely képes a szövegek hatékony feldolgozására. A jelen dolgozat e hiányzó elméleti megfontoláshoz kíván adalékul szolgálni.

A szentimentelemzés számos elméleti nyelvészeti és nyelvtechnológiai feladata közül a legalapvetőbb annak a kérdésnek az alapos megfontolása, hogy hogyan határozható meg magának a nyelvi értékelésnek a mibenléte. Jelenleg nincs tudomásom olyan dolgozatról, amely a nyelvi értékelés mibenlétének a kérdését a szentimentelemzés feladata szempontjából vizsgálat tárgyává tenné.

A szentimentelemzést célzó kutatásokra általánosságban jellemző, hogy a feladatot az úgynevezett *szentimentkifejezések*³ (*sentiment words*) szótári formába rendezett listáival igyekeznek – legalább részben – megoldani (vö. Vázquez–Bel 2012; Liu 2012: 12–13). A szentimentelemzés terén már-már vezérelvnek tekinthető az a megállapítás, hogy az értékelő szemantikai tartalmak legfontosabb indikátorai e pusztán szentimentkifejezések, vagyis azok a legkisebb nyelvi egységek, amelyek értékelő tartalmat hordoznak. E megfontolás alapján a szentimentszótárak a legtöbb elemzésben kulcsszerepet töltenek be.

Belátható, hogy egy szentimentlexikon létrehozása csupán abban az esetben lehetséges, ha feltételezzük, hogy bizonyos nyelvi elemeknek pozitív vagy negatív értékelő jelentéskomponense van a kontextuális sajátságoktól függetlenül; másképpen, egy meghatározott polaritási sajátsággal rendelkező szemantikai szerkezetük részeként (vö. Moghaddam–Popowich 2010: 3; Baroni–Vegnaduzzo 2004; Kamps et al. 2004; Vegnaduzzo 2004). Így például Moghaddam és Popowich (2010: 3) a (2a) alatti példákat pozitív, a (2b) alattiakat negatív polaritásúaknak ítéli, s velük szemben a (3) alattiakat a

³ Abból a megfontolásból használok a *szentimentkifejezés* terminust (eltérően az angol *sentiment word* terminus szó szerinti fordításától), hogy szentimentkifejezésként nem csupán egy-, de többszavas frazeológiai egységek is szolgálhatnak; például *értéktelen ~ egy fillért nem ér*.

szentimentérték szempontjából semleges mellékneveknek tartja (a szerzők dolgozatában bemutatott listából most csupán az első 5-5 elemet közlöm)⁴:

- (2) a. *good* 'jó', *nice* 'jó, kedves', *awesome* 'tök jó', *excellent* 'kiváló',
great 'nagyszerű'
b. *bad* 'rossz', *awful* 'borzasztó', *defective* 'hibás, tökéletlen', *faulty*
'hibás', *poor* 'szegény'
- (3) *mediocre* 'közepes, közepszerű', *average* 'átlagos', *enough* 'elég,
elégséges', *fair* 'korrekt, méltányos', *okay* 'rendben'

A megoldás problematikussága véleményem szerint a semlegesnek ítélt kifejezéseknél a legszembeötlőbb. Mindazok a nyelvi elemek ugyanis, amelyek valamilyen minőséget fejeznek ki, e sajátágukból fakadóan jelenthetnek pozitív vagy negatív értékelést az aktuális kontextus függvényében. Ennek megfelelően a (3) alatti elemek pozitív és negatív szentimentértéket egyaránt hordozhatnak. Az elmondottakkal összefüggésben egyre elterjedtebb az úgynevezett doménfüggő szentimentszótárak alkalmazása, amelyek tartalmukban és a bennfoglalt kifejezések polaritását tekintve igyekeznek az aktuális szövegdomén sajátágaihoz igazodni (lásd például Hangya et al. 2015).

Számos újabb kutatás (például Feldman et al. 2010; Ruppenhofer–Rehbein 2012) a szentimentszótárak mellett különböző szemantikai kompozíciós szabályok megalkotását és alkalmazását is célozza. A törekvés azon a megfigyelésen alapul, hogy a szentimentkifejezés szövegkörnyezetének elemei hatást gyakorolhatnak az adott kifejezés értékére. Például a szentimentérték törölhető tagadó operátorral (4), illetve fokozható (5a) vagy csökkenthető úgynevezett intenzifikáló kifejezések segítségével (5b) (vö. Bolinger 1972; Quirk et al. 1985; Pete 2009; Ruppenhofer–Rehbein 2012).

- (4) Nem hazudott.
(5) a. Rendkívül okos.
b. Valamennyire tehetséges.

Olyan szentimentelemző rendszer megalkotását szorgalmazzák tehát, amely képes e módosításokat hatékonyan kezelni (vö. Feldman et al. 2010: 3).

A fentebb elmondottak alapján a következőket állapíthatjuk meg: A legtöbb szentimentelemzést célzó vállalkozás a szentimentlexikonok létrehozására, s így a lexikai szintű nyelvi értékelésre, valamint a szemantikai kompozíciós szabályszerűségekre fókuszál. Ezzel összefüggésben a szentimentelemzést a nyelvtechnológiai kutatások alapvetően szemantikai vonatkozású

⁴ Az angol nyelvű kifejezések jelentéseit a *SZTAKI Szótár* alapján adtam meg.

problémaköréként tartják számon.⁵ Csupán elvétve találni olyan, a szentimentelemzés tárgykerében született dolgozatot, amely egyáltalán említést tenne a feladat pragmatikai vonatkozásairól (ilyen például Honkela et al. 2012).

A jelen dolgozatban ezért arra vállalkozom, hogy részletesen megvizsgáljam a nyelvi értékelés mibenlétének a kérdését. Dolgozatom célja kettős: egyrészt a nyelvi értékelés néhány sajátosságának körüljárása alapján amellet fogok érvelni, hogy a nyelvi értékelés mibenléte túlmutat a lexikai és a kompozíciós szemantika szféráján, ezáltal a pragmatikai sajátságok jelentősebb figyelembe vétele – a lexikai pragmatikai elméleti keret legalább részleges nyelvtechnológiai implementációja – javíthat a számítógépes értékelemzés jelenlegi eredményein. Másrészt azt is meg kívánom mutatni, hogy a hazai elméleti nyelvészeti szakirodalom nyelvi értékelésről tett megállapításai sem tekinthetők eléggé kidolgozottaknak ahhoz, hogy az alkalmazott szentimentelemzési kutatásokat hatékonyan támogathassák.⁶ Végül javaslatot fogok tenni a nyelvi értékelés mibenlétének egy újabb, pragmatikai megközelítésére, amely reményeim szerint egy nyelvtechnológiai implementációhoz is elméleti alapot adhat majd a jövőben.

1. A nyelvi értékelés megközelítési módjai a hazai elméleti nyelvészeti irodalomban

Laczkó (2008: 326) minősítésnek nevezi azt az eljárásmodot, amelyben egy élő személyt, egy élettelen tárgyat, egy eseményt vagy egy folyamatot minősége, jellege alapján valamilyennek nyilvánítunk. A szerző szerint a magyar nyelvben a minősítés kifejezésére a leggyakrabban a minősítő és fajtajelölő mellékneveket használjuk, ugyanakkor az ezekben kifejezett szemantikai tartalom „szerkezetekkel vagy az irodalmi nyelvben gyakori formával, körülírással is” átadható. A „szerkezeteket” illetően Laczkó (2008: 326) megemlíti, hogy azok többféle formában létezhetnek, tekintve, hogy bizonyos esetekben a mondat szavai grammatikai szabályok segítségével kapcsolódnak össze, más esetekben azonban a tagokat „logikai, pragmatikai jelentés vagy „csupán” a nyelvszokás tartja össze”. Kiemeli továbbá, hogy a minősítés nyelvi kifejezése alá- és mellérendelő szintagmákkal egyaránt lehetséges, ily módon minősíthetünk alany–állítmány-visztonnyal (*Péter tanuló*), tárgyas szintagmá-

⁵ Ezt mutatja az a jelenség is, hogy a szentimentértéket az angol nyelvű terminológiában több szerző is (például Moghaddam–Popowich 2010: 1) szemantikai orientációként (*semantic orientation*) emlegeti.

⁶ A jelen dolgozat keretei nem tették volna lehetővé, hogy a nemzetközi elméleti nyelvészeti irodalom vonatkozó tételeit is kimerítően áttekinthessem.

val (*Péter szeret olvasni*), határozós szintagmával (*uszodába jár*), minősítő jelzős szerkezettel (*magas diák*), azonosító értelmező szerkezettel (*Péter, a barátom*), valamint a mellérendelő szintagmák különböző altípusaival (*érdekes, sőt izgalmas; lassan, megfontoltan dolgozott*). Mindemellett említést tesz a minősítés kifejezésére alkalmas szerkezetek sajátos típusáról, az állandósult szókapcsolatokról is (*letargiába esik; szegény, mint a templom egere*).

Szilágyi (1996: 11) értékjelentésnek nevezi a szójelentésnek azt az összevetőjét, amelyben „kifejeződik a beszélőközösségnek a jel tárgyához való viszonya”. Megállapítja, hogy az értékjelentés pozitív és negatív attitűdöt tükrözhet, illetve léteznek az értékjelentés szempontjából semleges kifejezések is (például *nátrium-klorid, fonéma* stb.). A szerző (1996: 11–12) szerint értékjelentésről csupán akkor beszélhetünk, ha az a szótári jelentés részét képezi; tehát „az értékjelentés vizsgálata [...] a szójelentésre irányul”.

Berényi (1988: 75) Szilágyi (1996)-hoz hasonlóan ugyancsak szemantikai komponensként tételezi a nyelvi értékelést, amikor a főnevek és a mellénevek sajátosságaként említi az esetlegesen megjelenő, úgynevezett „értékelő jegy”-et. Véleménye szerint a főnév „önmagában lebecsülésre, vagy éppen ellenkezőleg, felértékelésre utaló szemantikai jeggyel rendelkezik”. Mindennek fényében ellentmondásosnak tűnhet több megállapítás is, amelyeket Berényi (1988) dolgozatában más helyütt tesz. Egyrészt arról beszél, hogy a denotatív lexikai elemeket a hozzájuk társuló melléknév affektív jegye is felrúházhhatja „a beszélői becslésre utaló szubjektív értékkel” (Berényi 1988: 77). Másrészt felhívja a figyelmet arra a jelenségre is, amikor a jelzős összetétel elemei közül önmagában egyik sem telített a szubjektív megítélés jegyével, de az összetétel már rendelkezik ezzel a sajátossággal (vö. Berényi 1988: 77). Ez utóbbi szemantikai változás véleménye szerint mind a kapcsolódás pusztán tényéből, mind pedig a tágabb szövegösszefüggés értelmezésre gyakorolt hatásából fakadhat. Berényi (1988: 77) a jelenségre *a külgazdasági politika egyéni arculata* és *a gazdasági tömörülés vezető ereje* kifejezéseket hozza példaként.

Fentebbi észrevételein túl Berényi (1988) a nyelvi modalitás jelenségét is tárgyalja az értékelés kapcsán. Az ígéket illetően megállapítja, hogy a propozicionális tartalomra vonatkozó beszélői becslés „a bizonyosságtól a bizonytalanságig, a fenntartásig vagy az elutasításig terjed”, és mindezek kifejezése történhet igemódokkal és módbeli segédigékkel egyaránt (Berényi 1988: 78). Emellett olyan úgynevezett „modalizáló elemeket” is feltételez, amelyek mondathatározóként explicit vagy implicit módon utalnak a beszélői állásfog-

lásra az állított tartalommal kapcsolatban; például *természetesen, kétségtelen, köztudott* stb. (vö. Berényi 1988: 76, 77–78).

Berényihez hasonlóan Péter Mihály (2008: 2) is a modalitás fogalmával hozza összefüggésbe a nyelvi értékelést, amikor azt a modalitás egyik fajtájaként határozza meg, s szemantikai kategóriának tekinti. Ugyanakkor megjegyzi, hogy „a nyilatkozat modális kereteként a pragmatika szférájába tartozik” (Péter 2008: 2). Berényi (1988)-cal és Szilágyi (1996)-tal megegyezően ő is amellett érvel, hogy az értékelés szerves részét képezi a nyelvi elemek jelentésszerkezetének (vö. Péter 1991: 47; 2008: 2). Ami a nyelvi értékelés típusait illeti, Péter (1991, 2008) két különböző szempont alapján a következő kategorizálási megoldásokat tartja lehetségesnek: egyrészt megkülönböztethetjük az értékelés „általános” és „speciális” típusát (vö. Péter 2008: 2), s az elsöre a *jó* és a *rossz* mellékneveket, a másodikra a *szép*, a *hasznos*, a *rút* és a *haszontalan* fogalmakat hozza példaként. Másrészt megkülönböztethetjük az „emocionális” és a „racionális” értékelés típusát (vö. Péter 1991: 46). A szerző az előbbire a *főnököm remek ember*, az utóbbira a *habbeton rossz hővezető* mondatokat hozza példaként. Az emocionális típushoz kapcsolódik a szerző 2008-as dolgozatában tett megállapítása, miszerint az értékelés rendszerint társadalmilag és történelmileg meghatározott (például egy jó autó ismérvei napjainkban nem azonosak a múlt század harmincas éveinek vonatkozó ideáival), s „a társadalmi érvényű értékelések mindenkor bizonyos normákra, a valóság idealizált modelljének megfelelő elképzelésekre, követelményekre vonatkoznak” (Péter 2008: 2). Ugyanakkor megjegyzi, hogy az értékelés lehet nem e társadalmi norma által meghatározott, hanem egyéni is.

A fentebb tárgyalt szerzők tehát Laczkó (2008) kivételével rendre a modalitás fogalmához kapcsolják a nyelvi értékelést. Vegyük hát górcső alá a nyelvi modalitás szakirodalmának vonatkozó tételeit!

Kiefer (1986: 24–27) a nyelvi modalitást illetően a *faktív* és a *nem faktív attitűd* fogalmaival operál. Közülük a faktív attitűdöt kifejező szerkezetek a szerző rendszerében olyan faktív attitudinális operátorokat tartalmaznak, amelyek „a beszélőnek az adott tényálláshoz való érzelmi-értelmi viszonyulását” fejezik ki (Kiefer 1986: 25):

- (6) a. Pisti *sajnos* beteg.
- b. Pisti *szerencsére* már egészséges.

Murvai (2001: 59) és Pete (2002: 189–190) megkülönböztet egy úgynevezett *értékelő modalitás*-típust, azonban e fogalmat a szerzők teljesen eltérően értelmezik. Murvai (2001: 59) az értékelő modalitást – a dolgozatában közölt

példák alapján (*jó–rossz; kívánatos–sajnálatos*) – a kieferi (1986: 24–27) faktív attitűdöt kifejező attitudinális operátorokhoz hasonlóan képzelem el, lévén, hogy ezek az elemek is a beszélőnek a propozíciós tartalomhoz való érzelmi–értelmi viszonyulását fejezik ki. Pete (2002: 189–190) véleménye szerint ugyanakkor az értékelő modalitás „a beszélő ismeretei bizonyosságának a fokát [...] jelöli” (Pete 2002: 189). Pete (2002: 189) rendszerében tehát az értékelő modalitás a kieferi nem faktív attitűd fogalmával rokonítható, ahol a szerkezet olyan attitudinális operátort tartalmaz, amely a beszélő bizonyossági fokára utal:

(7) Pisti *valószínűleg* beteg.

Pete (2002: 189–190) alapján tehát például a (8) alatti megnyilatkozások „nem tartoznak az értékelő modalitás körébe” (Pete 2002: 189).

(8) a. Éva rendkívül okos lány.
b. Éva nem buta lány.

Összegzőképpen a következőket mondhatjuk el: az áttekintett, nyelvi értékeléssel kapcsolatos irodalmi tételek megállapításai alapvetően három sajátság köré csoportosulnak. Egyrészt az értékelést a legtöbb szerző a nyelvi modalitáshoz tartozó fogalomként képzelem el. Másrészt azt alapvetően szemantikai komponensnek, a szótári jelentés részének tekintik. Harmadrészt az értékelés kapcsán a szerzők rendre a beszélői szubjektivitás fogalmával operálnak, s egyetlen kivételtől (Péter 1991) eltekintve nem szólnak az objektív nyelvi értékelés lehetőségéről.

2. A nyelvi értékelés mibenlétének vizsgálata

A dolgozat további részében részletes vizsgálat tárgyává teszem a nyelvi értékelés mibenlétét, valamint górcső alá veszem annak néhány sajátságát, amelyeket a számítógépes értékeléselemzésnek a bevezetőben összegzett megközelítési módjaival, valamint az elméleti nyelvészeti szakirodalom fentebb (az 1. részben) tárgyalt megállapításaival egyaránt ütköztetni fogok. A fejezetben tett megállapítások alapján végül javaslatot teszek egy a nyelvi értékelés mibenlétének egy újabb, pragmatikai megközelítésére.

2.1. A nyelvi értékelés mint modális kategória

Amint azt az 1. részben megállapítottuk, a legtöbb szerző a nyelvi értékelést szemantikai kategóriának tekinti, és a modalitás fogalmával hozza kapcsolat-

ba (vö. Kiefer 1986; Berényi 1988; Murvai 2001; Pete 2002; Péter 2008). Ugyanakkor amellet érvelek, hogy a modalitás elméleti keretében az értékelés fogalma a faktivitás fogalmával fonódik össze, ami nem szerencsés a számítógépes értékeléselemzés szempontjából; a szentimentelemzés során ugyanis nem a beszélői bizonyosság fokát, hanem a kifejeződő értékítéletet keressük a szövegben.

Láttuk, hogy Kiefer (1986: 24–27) *faktív attitűd* fogalma olyan attitudinális operátorok meglétét tételezi a megnyilatkozásban, amelyek „a beszélőnek az adott tényálláshoz való érzelmi-értelmi viszonyulását” fejezik ki (Kiefer 1986: 25). Ugyanakkor a szentimentelemzés nem korlátozódhat azokra a megnyilatkozásokra, amelyek faktív attitudinális operátort tartalmaznak, hiszen a feladat szempontjából érdemi információt hordoznak az ilyen operátort nem szerepeltető, értékelést megfogalmazó adatok is:

- (9) a. Pisti beteg.
- b. Pisti már egészséges.

Jegyezzük meg, hogy a Kiefertől idézett meghatározás ettől a kitéeltől eltekintve is szűkös a számítógépes értékeléselemzés számára: a szentimentelemzés ugyanis nem korlátozódhat kizárólagosan a Péter (1991: 46) által „emocionális”-nak nevezett, azaz a beszélő szubjektív ítéletét megfogalmazó értékeléstípus vizsgálatára; a „racionális” értékelés éppúgy releváns információ (vö. Péter 1991: 46).

Pete (2002) és Berényi (1988) az értékelő modalitást a propozíciós tartalommal szembeni bizonyosság ítéleteként képzelik el. Pete (2002: 189) véleménye szerint a beszélő ismereteinek bizonyossági foka különböző operátorok, Berényi (1988: 78) szerint igemódok és módbeli segédigék segítségével juthat kifejeződésre. A fentebb, (8) és (9) alatt közölt példák tehát e szerzők rendszere alapján sem tartoznak az értékelő modalitás kategóriájába.

A modalitás elméleti kerete ugyanakkor más szempontból is problematikus a szentimentelemzés számára. Az indirekt beszédaktusokban megfogalmazott értékelést például ez a rendszer nem teszi kezelhetővé (vö. Wilson–Sperber 2004). Kiefer (1986: 32) szerint a kérdés egy olyan beszédaktus, amely „nem fejez ki kognitív vagy emotív viszonyulást” a benne megfogalmazott tényállással szemben, ez alapján pedig a kérdő formájú adatokat ki kellene zárni a vizsgálat köréből. Ugyanakkor tudvalevő, hogy a kérdő mondatfajta számos funkciója között informáló szerepet is betölthet (vö. Perrault–Allen 1980: 170). Nyelvi produktumaink számtalan olyan kérdő

A nyelvi értékelés mibenléte a számítógépes értékeléselemzés szempontjából

formájú megnyilatkozást tartalmaznak, amelyeknek nem információkérési, hanem információközlő szándékuk van:

- (10) a. Hallottad, milyen rossz minőségű fék van rajta?
- b. El tudsz képzelni ennél borzalmasabb telefont?

A fentebb tárgyaltak alapján tehát a nyelvi értékelés, valamint a beszélői bizonyosság két önálló dimenzióként rajzolódik ki, és nem célszerű megpróbálni azokat egy problémaként kezelni a modalitás elméleti keretében.

2.2. A nyelvi értékelés mint szemantikai komponens

Amint azt a nyelvi értékelés megközelítési módjait áttekintő részben láttuk, Berényi (1988), Péter (1991) és Szilágyi (1996) egyaránt a szó szemantikai komponenseként tételezi a nyelvi értékelést. Szilágyi (1996: 11–12) az értékjelentést a szótári jelentés részeként kezeli, ennek értelmében az értékjelentés vizsgálatát a szó szemantikai sajátosságának vizsgálataként képzei el. Hasonlóan, Péter (1991: 47; 2008: 2) is a nyelvi elemek jelentésszerkezetének szerves részeként tekint az értékelésre, Berényi (1988: 75) pedig egy úgynevezett „értékelő jegy”-ről beszél, amely a főnevek és a melléknevek sajátóságaként hordozza a lebecsülés vagy felértékelés szemantikai tartalmát. A pragmatikai jelentés értékelésbeli szerepéről csupán Laczkó (2008: 216) tesz említést.⁷

Annak céljából, hogy megállapítsam, vajon tekinthető-e az értékelés a szó jelentéskomponensének, a következő sajátóságokat, illetve kérdéseket vizsgálom meg részletesen: a domén-, illetve kontextusfüggés problémáját, a szemantikai kompozíciós szabályok alkalmazhatóságának a kérdését, valamint az úgynevezett polaritásváltási jelenséget.⁸

2.2.1. A kontextusfüggés jelensége a szentimentelemzés szempontjából

Számos kifejezés egy adott szövegdoménben, illetve egy adott target értékelőjeként eltérő polaritással rendelkezik valamely más doménbeli, illetve targetfüggő sajátosságához képest (vö. Ahn et al. 2012; Vázquez–Bel 2012; Vechtomova et al. 2014). A doménfüggés esetére említsük példaként a *beolt* igét, amely sem az orvoslás, sem a biológia doménjében alapvetően nem fejez ki

⁷ Tekintettel arra, hogy a különböző modalitáselméletek értékelés-fogalmait több alapvető okból sem tartottam a szentimentelemzés számára megfelelő elméleti keretnek, e probléma szempontjából ezeket nem tárgyalom.

⁸ Az ironia jelenségével a dolgozat terjedelmi korlátai miatt nem foglalkozom.

értékelést (11a), azonban politikai tárgyú diskurzusban rendre negatív értékelő tartalmat hordoz (11b).⁹ Tekintsük az alábbi, valós nyelvi példákat!

- (11) a. A Tremex fuscicornis nevű fadarázs nőstény gombaspórákat hordoz, és azzal oltja be a fát, mikor belepetézik.
b. A saját sikertelenségét akarja a fidesz az MSZP kormányra tolni! (...) Mindegyik esztét vesztí már az ellenségképtől? Rendesen beoltotta őket Orbán!

A targetfüggés esetét példázzuk a *nagy* melléknévvel, amely egy telefontöltő vonatkozásában általában negatív, egy számítógép-memória tulajdonságaként azonban pozitív sajátságának tekinthető.

A domén- és targetfüggés jelenségével kapcsolatosan, Ahn és munkatársai (2012: 5–6) két polaritástípust tételeznek fel, s rendszerüket a koreai mellékneveken keresztül demonstrálják: az úgynevezett *abszolút polaritást* (*Absolute Polarity*) és a *relatív polaritást* (*Relative Polarity*). A szerzők (Ahn et al. 2012: 6) szerint, amennyiben egy melléknév abszolút polaritással rendelkezik, úgy polaritása nem függ az adott doméntól, amelyben szerepel, illetve targettől, amit minősít. Amennyiben azonban a melléknév relatív polaritású, úgy polaritása az aktuális domén, illetve target függvényében áll.¹⁰ (12) alatt közlünk néhány példát a szerzők rendszeréből.

- (12) Abszolút pozitív polaritás: *cohta* 'jó'; *chinhata* 'barátságos'
Abszolút negatív polaritás: *telepta* 'piszkos'; *sikkulepta* 'zajos'
Relatív polaritás: *mukepta* 'nehéz'; *kanghata* 'erős'

Ahn és munkatársai (2012: 6) összesen 2135 melléknévi predikátumot megvizsgálva arra a megállapításra jutnak, hogy azok 58,6%-a rendelkezik relatív, s csupán 41,4%-a abszolút polaritással.

A szentimentkifejezések polaritásának e nagymértékű domén- és targetfüggése problematikussá teszi a megközelítést, amely a szó szemantikai komponenseként, a lexika szintjén rögzített sajátságként tételezi a pozitív vagy

⁹ A szemantikai tartalom egyéb eltéréseivel e helyütt nem foglalkozom, tekintettel arra, hogy a szentimentelemzés szempontjából a polaritásbeli eltérések a relevánsak.

¹⁰ A bemutatott kategóriák nem azonosíthatóak Péter (1991: 46) korábban tárgyalt emocionális és racionális értékelés fogalmaival (l. fentebb, 1.). Amíg ugyanis Péter (1991: 46) rendszere az érvényességi kör alapján osztályozza a nyelvi értékelést, vagyis aszerint, hogy azok egyéni megállapításon vagy általános igazságon alapulnak-e, addig az utóbbi felosztás a „puszta” target–szentiment-vizsgálattal operál. Ezzel összefüggésben tehát Péter (1991: 46) mindkét korábban idézett példája (l. fentebb, 1.) abszolút polaritású, tehát targettől függetlenül rendre ugyanazzal a polaritással rendelkező értékelő elemet (*remek* és *rossz*) hordoz.

negatív polaritást. Ez az elmélet legfeljebb az abszolút polaritású elemek esetében állja meg a helyét, amennyiben elfogadjuk, hogy bizonyos lexémák (l. *undorító, selejtes, alantas* stb.) inherens polaritással rendelkeznek. A relatív polaritás minősége azonban – annak relatív természetéből fakadóan, a fentebb tárgyaltak értelmében – semmiképpen sem lehet a lexika szintjén eleve kódolt információ. Ezzel összefüggésben, az értékjelentés vizsgálata nem irányulhat pusztán a szó jelentésére (l. fentebb, 1.).

A dolgozat bevezetőjében már szóltunk Moghaddam és Popowich (2010: 3) megoldásáról, miszerint a szerzők a mellékneveket kontextustól függetlenül pozitív, negatív, illetve semleges kategóriákba sorolják. Amellett érveltem, hogy mindazok a nyelvi elemek, amelyek valamilyen minőséget fejeznek ki, az aktuális kontextusban pozitív és negatív értékítéletet egyaránt tarthatnak. Ily módon kontextusfüggetlen semlegességük véleményem szerint nem tételezhető.

A kontextus szerepe kapcsán szeretnék említést tenni az úgynevezett polaritási skála alkalmazhatóságáról. A szentimentelemzés nemzetközi gyakorlatában (vö. Moghaddam–Popowich 2010; Honkela et al. 2012; Ruppenhofer et al. 2014) bevettnek számít, hogy a különböző elemek szentimentértékbeli eltéréseiket úgy igyekeznek kezelni, hogy azokat elhelyezik egy matematikai értelemben vett diszkrét skálán (vö. Moghaddam–Popowich 2010). A polaritási skála lehetőséget ad arra, hogy a pozitív és a negatív értéktartományon belül további konszekutív értékeket tegyenek fel az elemzők (vö. Honkela et al. 2012; Ruppenhofer et al. 2014).¹¹

Ahhoz, hogy rámutassunk a megoldás kérdéses voltára, tekintsünk két elemet Ruppenhofer és munkatársai (2014) kérdőíven alapuló skaláris rendszeréből. A szerzők polaritási skáláján a *second-rate* ('másodosztályú') kifejezés negatívabb minősítést kap a *so-so* ('tűrhetően') kifejezésnél. Kétségtelen azonban, hogy a fentebbi elemek a velük kifejezésre juttatott ítéletet tekintve igen változatosan lehetnek a különböző kontextusokban, ezért nem volna helytálló megoldás, ha minden előfordulásukhoz ugyanazokat a szentimentértékeket rendelnénk az itt bemutatott értékhierarchiára támaszkodva.

2.2.2. A kompozíciós szemantika lehetőségei a szentimentelemzésben

Amint azt a dolgozat bevezetőjében megmutattam, a legtöbb szentimentelemzést célzó munka a szentimentlexikonok mellett a szemantikai kompozíciós szabályszerűségekre fókuszál (például Feldman et al. 2010; Ruppenhofer–

¹¹ A megoldás alapjának tekinthető többek között Osgood et al. (1957).

Rehbein 2012). Olyan szabályalapú szentimentelemző rendszer megalkotását célozzák, amely képes a szentimentkifejezés környezetében levő nyelvi elemeknek a szentiment-kifejezésre gyakorolt hatását kezelni (vö. Feldman et al. 2010: 3; Liu 2012: 60).

Az angol nyelvű terminológia szentimentshiftereknek (*sentiment shifters*) nevezi azokat a nyelvi elemeket, amelyek befolyásolják a lexémák elsődleges, azaz szótári szentimentértékét (vö. Liu 2012: 60). Ahogyan azt a bevezetőben megmutattuk, a szentimentérték így például törölhető tagadó operátorral, illetve fokozható vagy csökkenthető úgynevezett intenzifikáló kifejezések segítségével (vö. Székely 2007; Pete 2009; Ruppenhofer–Rehbein 2012). A korábban (4) és (5) alatt közölt példákat itt (13) és (14) alatt megismételjük:

- (13) Nem hazudott.
- (14) a. Rendkívül okos.
b. Valamennyire tehetséges.

Számos újabb kutatás (például Feldman et al. 2010; Ruppenhofer–Rehbein 2012) arra törekszik tehát, hogy olyan szemantikai kompozíciós szabályokat állapítson meg, amelyek segítségével a szentimentszótárok elemeinek lexikai polaritásából kiindulva kiszámítható a kompozíciós szentimentérték.

Ugyanakkor a vázolt eljárás mód igencsak problematikus, hiszen az értékelést kifejező nyelvi egységek gyakorta nem teljesen kompozicionálisak. Ahogyan arra több szerző (Israel 2004: 711; Feldman et al. 2010: 4) is rámutat, a szentimentérték közel sem kalkulálható ki minden esetben maradéktalanul szemantikai kompozíciós szabályok segítségével. Tekintsük az alábbi példákat!

- (15) a. Mari *jó* tanuló.
b. Mari *nem jó* tanuló. → Mari rossz tanuló
- (16) a. Mari *buta*.
b. Mari *nem buta*. ↯ Mari okos

A példák azt mutatják, hogy amíg a pozitív polaritású elem tagadása esetén negatív értékeléssel kell számolnunk, addig ugyanez a művelet a negatív érték esetében nem feltétlenül vezet hasonló eredményre. Véleményem szerint ugyanis a (16b) alatti példa tartalma az aktuális kontextustól függően semleges és pozitív polaritású egyaránt lehet.

A kompozíciós szemantikai megközelítési mód azért problematikus, mert – akárcsak a szótáralapú elemzési megoldás – a nyelvi jel egyfajta kódmodellszerű felfogását tükrözi. A humán kommunikáció e pragmatikai modellje

a kommunikációt úgy tekinti, mint „explicit módon közölt üzenetek kódolása és dekódolása” (Németh T. 2011: 45). Eszerint konvencionális kapcsolat áll fenn a jelzéseként felfogott szavak és az üzenetként értékelt szójelentések között, azaz a jelentések bizonyos értelemben „áttetszenek a mondatokon” (Reboul–Moeschler 2000: 23). Belátható azonban, hogy a kompozíciós szemantikai megközelítés inherensen korlátozott, ezáltal a módszer önmagában nem hozhat a nyelvtechnológia számára kimagasló eredményt.

2.2.3. A polaritásváltás jelensége a szentimentelemzés szempontjából

Több szerző (Tolcsvai Nagy 1988; Laczkó 2007; Andor 2011; Kugler 2014) is felhívja a figyelmet egy az értékelést kifejező elemek, valamint azok intenzifikálói kapcsán megfigyelhető terjedőben levő jelenségre, miszerint „a negatív jelentéstartalmú és használatú lexikai egységek pozitív irányú jelentésváltozását vagy jelentésbővülését, jelentésük kiterjesztését figyelhetjük meg” (Andor 2011: 33). Andor (2011: 33) megemlíti, hogy a változás az ellentétes polaritás irányába is lehetséges, azaz pozitívból negatívba, ugyanakkor ez a változat jóval ritkább előfordulású.¹²

Vizsgálati eredményeim alapján a polaritásváltás az alábbi szintaktikai szerkezetekben következhet be:

A vizsgált kifejezés állítmányként önmagában pozitív értékelést fejez ki (vö. Andor 2011: 37):

(17) [...] tők *durva* volt, amit a hegedűs művelt a hangversenyen.¹³

A kifejezés egy főnév jelzőjeként annak pozitív voltát fejezi ki (vö. Andor 2011: 37):

(18) Nagyon *durva* sportkupét mutatott be a Volvo.

A kifejezés határozói funkcióban egy szentimentkifejezés értékét intenzifikálja:

(19) A *mocskosul* jól kinéző CGI-vérnek például búcsút inthetünk.

A kifejezés jelzőként egy szentimentkifejezés értékét intenzifikálhatja:

¹² A polaritásváltás jelensége feltételezhetően nyelvfüggetlen sajátosság (vö. Constant et al. 2009. 10–13; Andor 2011: 39–41; De Smedt–Daelemans 2012).

¹³ A megnyilatkozás Andor (2011: 37) valós nyelvi példája. Bár a kifejezésnek lehet negatív értékű olvasata is, számunkra ez a jelen probléma szempontjából nem releváns, ezért e helyütt ezzel nem foglalkozunk.

(20) a. (...) a miskolci Jameson Cinefesten *baromi* jó filmeket lehet nézni tökingyen.

b. Beyonce rajongó vagy? Biztos *jó* hülye lehetsz!

Végezetül említsük meg, hogy az utóbbi típusnak nem ritka az egybe írt változata sem:

(21) Harcsa Veronika jazz-énekesnő mesélt mindenféle izgi dolgot, és egy *marhajó* playlistet hozott nekünk!

Tolcsvai Nagy (1988: 402) amellett érvel, hogy amíg az általa „hagyományosnak” nevezett jelzői csoport tagjai, így például a *kiváló*, a *nagyszerű* vagy a *csodálatos* elemek magukban hordozzák „a szóval jelölt érték valódiságát”, addig az *állati*, *baromi* jelzők „a jelzett értékek relativitására utalnak”. Ennek következtében azok jelentésében „ironizáló magatartás” mutatkozik meg. Tolcsvai Nagy (1988) tehát nem tesz fel olyan lehetőséget, amelyben a vizsgált elemek teljes mértékben polaritást válthatnának, vagy elsődleges polaritásukat veszítve fokozó értelmű intenzifikálóként funkcionálhatnak. Székely (2007: 30, 45) ugyanakkor úgy véli, hogy ezek az elemek szemantikailag gyakorta motiválatlanok, s ezt a szemantikai motiválatlanságot a szerző (Székely 2007: 43–45) szerint két sajátság is alátámasztani látszik: egyrészt a vizsgáltakhoz hasonló szerkezetekben olyan elemek is kollokálódhatnak, amelyekben az intenzifikáló elem önmagában a módosított elem szemantikai tartalmával ellentétes tartalmat hordoz (például *pokoli hideg*). Másrészt bizonyos szókapcsolatok esetében a kifejezések között semmiféle szemantikai–logikai kapcsolat nem tételezhető (például *borzalmasan szeret*).

Véleményem szerint a polaritásváltáshoz kapcsolódó szemantikai motíváltság és motiválatlanság között az intenzifikáló elemek esetében egy skála tételezhető fel, amelyen a kollokációk a szemantikai motíváltság felől a szemantikai motiválatlanság felé „mozognak”, s ezzel együtt pragmatikai motíváltságuk megnövekszik. A jelenséget a (22) alatti kifejezésekkel példázzuk.¹⁴

- (22) a. dög bűdös
b. dög rossz
c. dög meleg¹⁵
d. dög jó

¹⁴ A példák mindegyike valós nyelvi.

¹⁵ A példa egy fokozó értelmű jelző és egy főnév kapcsolatából áll, az írásmódját is ez a szintaktikai sajátság indokolja. Ugyanakkor a szókapcsolatnak létezik egy lexikalizálódott változata is, amely tipikusan főnévként használatos, pl. *Ebben a dögmelegben nem lehet aludni*.

A *dög* szó – részben – a kellemetlen szag képzetével társul, ily módon a (22a) alatti példa szemantikailag motivált: a kifejezéssel jellemzett dolognak olyannyira kellemetlen szaga van, mint egy dögnek. A (22b) alatti példában ugyanakkor már nem emelődik ki az előbbihez hasonlóan egyetlen szemantika komponens; a *dög* teljes negatív jelentéstartamával fokozza a modifikált elemet. A (22c) alatti kifejezést illetően, Székely (2007: 43–45) alapján amellett érvelek, hogy a módosító és a módosított elem között nincs szemantikai–logikai kapcsolat, tekintve, hogy a *dög* lexéma lexikai–szemantikai reprezentációja nem tartalmazza a MELEG szemantikai komponenset. Ugyanakkor, a *dög* nem teljes mértékben deszemantizálódott fokozó elem a kifejezésben, hiszen a szókapcsolat a negatív értékelő tartalmát csupán annak negatív lexikai szintű polaritásából nyerheti, lévén, hogy a *meleg* relatív polaritású elem (erről részletesebben a 2.2.1. részben esett szó). A módosító elem elsődleges polaritásának tehát mindenképpen szerepet kell játszania a kapcsolat interpretációjában. A (22d) alatti példában a vizsgált elem már olyan kifejezéssel kapcsolódik, amely ellentétes annak lexikai szintű polarításával, emiatt szükségszerűen negatív elsődleges polaritás nélküli intenzifikátorként kell funkcionálnia. Ezzel összefüggésben úgy vélem, hogy az ellentétes polarítású elemek kapcsolatát pragmatikai ok teszi lehetővé, hiszen a szentimentszerkezettel a beszélő – a módosító elem affektív vagy intenzitási jegyét kiaknázva – hatás-keltési céllal él.

2.3. A nyelvi értékelés egy lehetséges pragmatikai megközelítése

A dolgozatban áttekintett nyelvi jelenségek alapján amellett érvelek, hogy a számítógépes értékeléselemzés lexikai és kompozíciós szemantikai megközelítése abban az esetben hozhat kimagasló eredményt, ha – legalább bizonyos nyelvi jelenségek esetében – sikerül lexikai pragmatikai szempontokat is érvényesíteni az elemzésben. Bár a megközelítés részletes bemutatására ebben a dolgozatban nem vállalkozhatom, az alábbiakban azt igyekszem megmutatni, hogyan számolhatunk el a lexikai szinten polaritással rendelkező intenzifikáló elemekkel a lexikai pragmatika elméleti keretében, valamint röviden utalok arra is, hogyan képelem megvalósíthatónak a tárgyaltak nyelvtechnológiai implementációját.

Wilson és Sperber (2004) alapján a megnyilatkozás kommunikált tartalma messze túlmutat azon a tartalomra, amelyet a pusztán nyelvi elemek és szerkezetek kódolnak, ezért a jelentést a hallgatónak kell kidolgoznia az aluldeterminált jelentésből a kontextus alapján. Ezzel összhangban a lexikai pragmatika elméleti keretében a lexéma alulspecifikált szemantikai reprezentációval

rendelkezik, és teljes jelentését a kontextusban, az aktuális pragmatikai következtetések útján nyeri el (vö. Bibok 2014: 229). Bibok (2014: 227–228) amellett érvel, hogy az alulspecifikált szemantikai reprezentációban foglalt enciklopédikus információk a pragmatikai tudás prototipikus vagy sztereotipikus struktúráiban gyökereznek. Amennyiben az aktuális kontextus összhangban áll a lexikai-szemantikai reprezentáció valamely prototipikus vagy sztereotipikus struktúrájával, úgy a lexéma az aktuális kontextusban problémamentesen interpretálható. Amennyiben azonban az aktuális kontextus nem áll összhangban a lexéma lexikai-szemantikai reprezentációjának egyik prototipikus vagy sztereotipikus struktúrájával sem, „távolabbi” enciklopédikus információk elérésére van szükség.

A fentebbi elméleti keretre támaszkodva a lexikai szinten polaritással rendelkező intenzifikálók problémája a következőképpen volna kezelhető: az elmélet alapján a lexikai szinten polaritással rendelkező intenzifikáló elemek alulspecifikált szemantikai reprezentációval rendelkeznek. Amennyiben az aktuális kontextus – jelen esetben maga az intenzifikált elem – összhangban áll az intenzifikáló elem lexikai szemantikai reprezentációjának valamely prototipikus vagy sztereotipikus értékelő struktúrájával, úgy a kifejezés problémamentesen interpretálható. Ezt láttuk például a (22a) kapcsolat esetében, ahol az intenzifikált elem összhangban áll a *dög* lexikai szemantikai reprezentációjában meglévő **KELLEMETLEN SZAG** szemantikai komponenssel. Amennyiben azonban az aktuális kontextus – tehát az intenzifikált elem – nem áll összhangban az intenzifikáló elem lexikai szemantikai reprezentációjának egyik prototipikus vagy sztereotipikus struktúrájával sem, úgy egyéb információkra van szükség a kapcsolat szemantikai tartalmának kidolgozásához. Amint láttuk, a (22c) példa esetében a módosított elem nem áll összhangban a módosító elem lexikai szemantikai tartalmával. Másrészt, mint arra rámutattunk, a módosított elem relatív, azaz kontextusfüggő polaritású. A kapcsolat értékelő jelentésének kidolgozásához ezért a hallgatónak a módosító elem negatív minőségi jegyére kell támaszkodnia. Ugyanakkor a (22d) példa megfelelő interpretációjához a korábbi esetekben segítségül hívott enciklopédikus információk egyike sem elegendő. Amellett érvelek tehát, hogy a vizsgált intenzifikáló elem szemantikai reprezentációja az argumentumstruktúra mellett magában foglal egy affektív vagy intenzitási minőségi jegyet (vö. Bibok 2000: 59), s a jelentést a hallgató a (22d)-hez hasonló szerkezetek esetében az intenzifikáló elem affektív vagy intenzitási minőségi jegye alapján dolgozza ki.

Úgy vélem, hogy a fentebb bemutatott lexikai pragmatikai megközelítési mód alkalmazott nyelvészeti implementációja hatékonyan támogathatná a lexikai szinten polaritással rendelkező fokozó értelmű kifejezések automatikus kezelését. Meglátásom szerint ugyanis a (22) alatti példákon bemutatott következtetési folyamatok – egy bizonyos szinten legalábbis – algoritmizálhatóak volnának. Ehhez egyrészt arra volna szükség, hogy a szentimentelemző rendszer ne csupán az abszolút, de a relatív polaritású elemeket is kezelni tudja, tehát mindkét polaritástípus elemeinek szótárakba rendezésére szükség volna. Másrészt a következtetési folyamatokat egyszerű kompozíciós szabályok formájában kódolni kellene a mesterséges intelligencia számára. Így például algoritmizálni kellene azt a feldolgozási mechanizmust, hogy amennyiben egy a lexikai szintjén egy bizonyos polaritással rendelkező intenzifikáló elem egy relatív polaritású elemet módosít, úgy a teljes kapcsolat polaritását az intenzifikáló elem szótári polaritása határozza meg. Ugyanakkor ha egy a lexikai szintjén egy adott polaritással rendelkező intenzifikáló elem egy vele ellentétes polaritású elemet módosít, úgy az egész konstrukció polaritása az alaptag polaritása szerint alakul.

Összegzés

A dolgozatban a számítógépes értékeléselemzés legalapvetőbbnek tekinthető dilemmáját, a nyelvi értékelő jelentés mibenlétét tárgyaltam részletesen.

Megállapítottam, hogy az automatikus szentimentelemzést célzó munkák a lexikai szintű nyelvi értékelésre, valamint a szemantikai kompozíciós szabályszerűségekre helyezik a hangsúlyt. Ezzel összefüggésben a szentimentelemzést alapvetően a nyelvtechnológiai kutatások szemantikai vonatkozású problémaköréért tartják számon. A nyelvi értékelés néhány sajátosságának körüljárása alapján azonban amellet érveltem, hogy a lexikai és a kompozíciós szemantikai megközelítés eredményein javíthatnánk, ha lexikai pragmatikai szempontokat is érvényesítenénk az elemzésben.

A nyelvi értékelést tárgyaló hazai irodalmi tételeket áttekintve megmutattam, hogy az értékelést a legtöbb szerző a nyelvi modalitáshoz tartozó fogalomként képzeli el, alapvetően szemantikai komponensnek tekinti, továbbá nem szól az objektív nyelvi értékelés lehetőségéről. Ugyanakkor rámutattam, hogy a hazai elméleti irodalom nyelvi értékelésről tett megállapításai nem nyújthatnak megfelelő alapot egy hatékony szentimentelemző rendszer létrehozásához.

Dolgozatom végén a lexikai szinten polaritással rendelkező intenzifikáló elemek példáján keresztül bemutattam, hogyan számolhatunk el ezzel a –

számítógépes értékelésemzés számára jelenleg nehezen kezelhető – jelenléggel a lexikai pragmatika elméleti keretében. Végül javaslatot tettem a bemutatott megközelítési mód nyelvtechnológiai implementációjára.

Hivatkozások

- Ahn, A – Laporte, É. – Nam, J. 2012. Semantic Polarity of Adjectival Predicates in Online Reviews. *CoRR abs/1211.4161*
- Andor József 2011. De durva ez a téma! – Megfigyelések a melléknévi polaritásváltásról. In *Hungarológiai Évkönyv* 12. 33–42.
- Berend, G. – Farkas, R. 2008. Opinion Mining in Hungarian based on textual and graphical clues. In *Proceedings of the 4th Intern. Symposium on Data Mining and Intelligent Information Processing*. Santander. 408–412.
- Bibok Károly 2000. A lexikai és az aktuális jelentés kapcsolata a kétszintű konceptuális szemantikában és a generatív lexikon elméletében. In Gecső T. (szerk.) *Lexikális jelentés, aktuális jelentés*. Budapest, Tinta Könyvkiadó, 56–63.
- Bibok, Károly 2014. Lexical semantics meets pragmatics. *Argumentum* 10. Debreceni Egyetemi Kiadó, 221–231.
- Baroni, M. – Vegnaduzzo, S. 2004. Identifying Subjective Adjectives through Web-based Mutual Information. In Buchberger, E. (ed.) *Proceedings of KONVENS*. Vienna, GAL, 17–24.
- Berényi Pálné 1988. Szaknyelv és szubjektivitás. *Magyar Nyelvőr* 112(1), 71–80.
- Bolinger, D. 1972. *Degree Words*. The Hague, Mouton.
- Constant, N. – Davis, C. – Potts, C. – Schwarz, F. 2009. The pragmatics of expressive content: Evidence from large corpora. *Sprache und Datenverarbeitung* 33(1–2), 5–21.
- De Smedt, T. – Daelemans, W. 2012. "Vreselijik mooii!" (terribly beautiful): A Subjectivity Lexicon for Dutch Adjectives. In *Proceedings of the 8th Language Resources and Evaluation Conference (LREC'12)*, 3568–3572.
- Feldman, R. – Rozenfeld, B. – Breakstone, M.Y. 2010. SSA – A Hybrid Approach to Sentiment Analysis of Stocks. *The Israeli Seminar for Computational Linguistics (ISCOL)*, Tel Aviv University. [http://web.mit.edu/michab/www/ISCOL_Paper.pdf]
- Hangya V. – Farkas R. – Berend G. 2015. Entitáorientált véleménydetekció webes híryanagokból. In Tanács A. – Varga V. – Vincze V. (szerk.) *XI. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2015)*. Szeged, Szegedi Tudományegyetem, 227–234.
- Honkela, T. – Izzatdust, Z. – Lagus, K. 2012. Text Mining for Wellbeing: Selecting Stories Using Semantic and Pragmatic Features. In *Lecture Notes in Computer Science* 7553, 467–474. [http://link.springer.com/chapter/10.1007/978-3-642-33266-1_58]

- Kamps, J. – Marx M. – Mokken, R. – Rijke m. 2004. Using WordNet to Measure Semantic Orientations of Adjectives. In *Proceedings of the fourth international conference on Language Resources and Evaluation*. Lisbon, Portugal, 1115–1118.
- Kiefer Ferenc 1986. A modalitás fogalmáról. *Nyelvtudományi közlemények* **88**(1–2), 1–38.
- Kugler Nóra 2014. A nyelvi polaritás kifejezésének egy mintázata, avagy milyen a félelmetesen jó? *Magyar Nyelvőr* **138**(2), 129–139.
- Laczkó Mária 2007. Napjaink tizenéveseinek beszéde szóhasználati jellemzők alapján. *Magyar Nyelvőr* **131**(2), 173–184.
- Laczkó Mária 2008. A minősítés kifejezésére használt szerkezetek a tanulók spontán beszédében. *Magyar Nyelvőr* **132**(3), 326–340.
- Liu, B. 2012. *Sentiment Analysis and Opinion Mining*. Draft.
[<http://www.cs.uic.edu/~liub/FBS/SentimentAnalysis-and-OpinionMining.pdf>]
- Miháltz M. 2010. OpinHu: online szövegek többnyelvű véleményelemzése. In Tanács A. – Vincze V. (szerk.) *VII. Magyar Számítógépes Nyelvészeti Konferencia* (MSZNY 2010). Szeged: Szegedi Tudományegyetem. 14–23.
- Miháltz M. 2013. OpinHuBank: szabadon hozzáférhető annotált korpusz magyar nyelvű véleményelemzéshez. In Tanács A. – Vincze V. (szerk.) *IX. Magyar Számítógépes Nyelvészeti Konferencia* (MSZNY 2013). Szeged: Szegedi Tudományegyetem. 343–345.
- Moghaddam, S – Popowich, F. 2010. Opinion polarity identification through adjectives. *Computing Research Repository* (CoRR) *abs/1011.4623*.
[<http://arxiv.org/ftp/arxiv/papers/1011/1011.4623.pdf>]
- Murvai Olga 2001. Szemantikai és pragmatikai kiegészítések a modalitás vizsgálatához. In *Officina Textologica* **5**, Debrecen, 59–69.
- Németh T. Enikő 2011. A humán kommunikáció modelljei és az ember-gép kommunikáció. In Németh T. Enikő (szerk.) *Ember-gép kapcsolat. A multimodális ember-gép kommunikáció modellezésének alapjai. Segédkönyvek a nyelvészet tanulmányozásához* **133**. Budapest, Tinta. 43–61.
- Neubarth, F. 2006. *Polarity, Alternatives, and Scales*. Doctoral dissertation. University of Vienna.
[http://www.ofai.at/~friedrich.neubarth/linguistics/FN_Diss_2006.pdf]
- Osgood, C.E. – Suci, G.J. – Tannenbaum, P.H. 1957. *The Measurement of Meaning*. Urbana, University of Illinois Press.
- Pang, B.–Lee, L. 2008. Opinion mining and sentiment analysis. In *Foundations and Trends in Information Retrieval* **2**(1–2), 1–135.
- Perrault, C.R.–Allen, J.F. 1980. A Plan-Based Analysis of Indirect Speech Acts. *American Journal of Computational Linguistics* **6**(3–4), 167–182.
- Pete István 2002. A nyelvi modalitás három fő típusának hierarchiája a magyarban. *Magyar nyelv* **98**(2). 173–191.
- Pete István 2009. A fokozás mibenlétéről. *Magyar Nyelvőr* **133**(3), 295–310.
- Péter Mihály 1991. *A nyelvi érzelmek kifejezés eszközei és módjai*. Budapest, Tankönyv.

- Péter Mihály 2008. A magyar aspektusról – más aspektusból. *Magyar Nyelv* **CIV**(1), 1–11.
- Quirk, R.–Greenbaum, S.–Leech, G.–Svartvik, J. 1985. *A Comprehensive Grammar of the English Language*. Harlow, Longman.
- Reboul, A. – Moeschler, J. 2000. *A társalgás cselei*. Budapest, Osiris.
- Ruppenhofer J. – Rehbein, I. 2012. *Anchoring sentiment analysis in semantic frames*. [<http://www.uni-hildesheim.de/ruppenhofer/pubs/longversion.pdf>]
- Ruppenhofer, J. – Wiegand, M. – Brandes, J. 2014. Comparing methods for deriving intensity scores for adjectives. In *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics 2*, Short Papers. Gothenburg, Sweden, Association for Computational Linguistics, 117–122.
- Székely Gábor 2007. Egy sajátos nyelvi jelenség, a fokozás. *Segédkönyvek a nyelvészet tanulmányozásához* **66**. Budapest, Tinta.
- Szilágyi N. Sándor 1996. *Hogyan teremtsünk világot?* Kolozsvár, Erdélyi Tankönyvtanács.
- Tolcsvai Nagy Gábor 1988. A mai magyar nyelv normarendszerének egy jelentős változásáról az „ifjúsági nyelv” kapcsán. *Magyar Nyelvőr* **112**(4), 398–406.
- Vázquez, S. – Bel, N. 2012. A Classification of Adjectives for Polarity Lexicons Enhancement. In Calzolari, N. – Choukri, K. – Declerck, T. et al. (eds.) *Proceedings of the Eight International Conference on Language Resources and Evaluation* (LREC 12); 2012. Paris, ELRA, 3557–3561.
- Vechtomova, O. – Suleman, K. – Thomas J. 2014. An Information Retrieval-Based Approach to Determining Contextual Opinion Polarity of Words. In *Proceedings of the 36th European Conference on Information Retrieval* (ECIR), Amsterdam.
- Vegnaduzzo S. 2004. Acquisition of subjective adjectives with limited resources. In *AAAI Spring Symposium Technical Report: Exploring Affect and Attitude in Text*. [<http://www.aaai.org/Papers/Symposia/Spring/2004/SS-04-07/SS04-07-033.pdf>]
- Wilson, D. – Sperber, D. 2004. Relevance Theory. In Ward, G.– Horn, L. (eds.) *Handbook of Pragmatics*. Oxford, Blackwell, 607–632.

Forrás

SZTAKI Szótár [<http://szotar.sztaki.hu/>]