

1956 és 1989 között keletkezett propagandaszövegek nyelvi sajátosságai

Szabó Martina Katalin^{1,2}, Ring Orsolya¹, Vincze Veronika³

¹ Társadalomtudományi Kutatóközpont, CSS-RECENS Kutatócsoport
1097 Budapest, Tóth Kálmán u. 4.

²Szegedi Tudományegyetem, Informatikai Intézet
6720 Szeged, Árpád tér 2.

³MTA-SZTE Mesterséges Intelligencia Kutatócsoport
6720 Szeged, Tisza Lajos körút 103.

ring.orsolya@tk.mta.hu, {martina,vinczev}@inf.u-szeged.hu

Kivonat Elemzésünkben a Magyar Szocialista Munkáspárt Központi Bizottságának (MSZMP KB) hivatalos havi lapja, a *Pártélet* első és utolsó két évének lapszámait vizsgáljuk, és összehasonlítjuk a két időszak szövegeinek statisztikai, morfológiai, szintaktikai és szemantikai jellemzőit. A vizsgált szövegek bizonyosan a totalitárius nyelvhasználatot reprezentálják, ezért kutatási eredményeinket felhasználhatónak tekintjük a politikai propaganda azonosítására és elemzésére irányuló kutatásokban és fejlesztésekben.

Kulcsszavak: információkinyerés, propaganda, megtévesztés, diskurzus-elemzés, szocializmus, politika

1. Bevezetés

A magyar történelem 1956 és 1989 közötti időszaka a történettudományban és a társadalomtudományban a gyakran vizsgált korszakok közé tartozik. Politikai diskurzusának nyelvi jellemzői azonban eddig nem képezték elemzés tárgyát. Ezt a problémát felismerve korpuszt építettünk a *Pártélet* nevű folyóiratból, az MSZMP KB hivatalos ideológiai lapjából, majd azt számítógépes nyelvészeti eszközökkel és módszerekkel elemeztük.

A *Pártélet* című lap vizsgálata lehetővé teszi az országot irányító állampárt hivatalos diskurzusának tüzetes elemzését. A lap célja a politikai ideológia terjesztése, tehát a közvetlen agitáció és propaganda volt. A *Pártélet*, amely 54 150 példányban jelent meg, elsősorban nem az átlagemberekhez, hanem az állampárt különböző tisztségviselőihez szólt. Bár nincsenek pontos adataink arra vonatkozóan, hogy pontosan kik olvasták, a nagy példányszámból arra következtethetünk, hogy a kiadvány valószínűsíthetően a párhierarchia egészét célozhatta. Dolgozatunkban két, megközelítőleg egyforma hosszúságú időszakot vizsgálunk: a 1957 januárja és 1958 decembere közötti (17 szám), valamint a 1988 januárja és 1989 áprilisa közötti időszakot (16 szám). A két időszakot a következő szempontok alapján jelöltük ki: Az első időszak a Kádár-korszak kezdete, közvetlenül az 1956-os forradalom után következő, hipotézisünk szerint az agitáció

és a propaganda szempontjából igen aktív periódus. A második időszakot, amely közvetlenül megelőzi a rendszerváltást és a Kádár-korszak végét jelenti, az első ellenpontjának tekinthetjük. Az első a rendszer megszilárdulásának, míg a második annak fokozatos felbomlásának időszaka, így azt feltételezzük, hogy a kettő diskurzusa eltérést mutat.

Célunk a propaganda nyelvi jellemzőinek feltárása volt, különös tekintettel a politikai nyelvhasználatot gyakorta jellemző előítéletekre és a megtévesztésre (Propaganda Analysis, 1938; Jalilifar és Alavi, 2011; Barrón-Cedeño és mtsai, 2019). Azt vizsgáljuk, hogy a politikai célok miként és milyen mértékben befolyásolják a nyelvhasználatot, a diskurzus szervezését érintő döntéseket (Jalilifar és Alavi, 2011). Elemzésünk egyben egy jövőbeli, szélesebb körű – a propagandát az online térben vizsgáló – kutatás megalapozásának tekinthető.

Annak ellenére, hogy a megtévesztés és elfogultság jelenségei különböző típusú diskurzusokban azonosíthatóak, a legtöbb szerző ezeket vagy hangzó szövegekben vizsgálja (Fetzer, 2008; Fraser, 2010; Scheithauer, 2007; Simon-Vandenberg és mtsai, 2007) vagy automatikus ténykivonatolási és bináris osztályozási feladatként kezeli (Greene és Resnik, 2009; Rubin és mtsai, 2015; Wang, 2017; Thorne és Vlachos, 2018; Graves, 2018). Bár a propaganda mint diskurzus kvantitatív és kvalitatív elemzése mind a számítógépes nyelvészet, mind a történettudomány és a politológia szempontjából fontos terület, eddig nagyon kevés kísérlet történt nagy mennyiségű propagandaszöveg szisztematikus elemzésére (Propaganda Analysis, 1938; Rashkin és mtsai, 2017; Barrón-Cedeño és mtsai, 2019). Különösen jelentős a hiány a magyar nyelvű szövegek vonatkozásában. Olyan dolgozatról pedig egyáltalán nincs tudomásunk, amely a magyar totalitárius diskurzus jellemzőit NLP-módszerekkel vizsgálná.

2. A vizsgálati korpusz

A vizsgálatokhoz használt korpuszt, a Pártélet című lap számait az Arcanum Digitheca¹ oldalról töltöttük le. A lap szkennelt, PDF-formátumban közzétett dokumentumait a letöltés után további komplex feldolgozási folyamatoknak vetettük alá, amelyek eredményeképpen megkaptuk a szövegek elemezhető és megfelelő minőségű, szöveges (txt) formátumú változatát (Szabó és mtsai, 2019).

A kész korpuszból a jelen dolgozatban bemutatott elemzésekhez kiválasztottuk a korpusz első két, és utolsó két évnyi lapszámát. E két alkorpusz alapvető statisztikai adatait az 1. táblázat közli.

korpusz	lapszám	mondatszám	tokenszám
1956-57	17	23573	503047
1988-89	16	28229	531962
Összesen	33	51802	1035009

1. táblázat. Az alkorpuszok alapvető adatai.

¹ <https://adtplus.arcanum.hu/en/collection/Partelet/>

Az alkorpuszok kiegyenlített mennyiségi adatai elősegítik az alkorpuszokban mért adatok összevetését egymással.

3. A vizsgálat módszere

Kutatásunk kiindulópontja az a hipotézis, hogy a *Pártélet* szerzői a megtévesztés és manipuláció különböző nyelvi eszközeit alkalmazták arra, hogy elrejtsek az akkori politikai vezetés számára nem kívánatos tényeket, és egyben befolyásolják, meggyőzzék a szövegek olvasóit (Girlea és mtsai, 2016; Rashkin és mtsai, 2017). Mivel elemzésünk alapja egy ideológiai folyóirat, amely a párt egyik fő szócsöve volt, azt feltételeztük, hogy valószínűleg a fenti jellemzők mindkét időszakban meghatározóak voltak, és az első időszakban elsősorban a forradalomhoz, a másodikban pedig a rendszerváltáshoz kapcsolódtak. Ugyanakkor azt is feltételeztük, hogy a két alkorpusz nyelvi jellemzői több sajátság tekintetében is különböznek egymástól, lévén, hogy az első alkorpusz szövegei a rendszer megerősödésekor, míg a második alkorpusz szövegei a rendszer gyengülésekor keletkeztek.

Első lépésben a szövegekre lefutattuk a magyarlanc nyelvi elemzőt, így megkaptuk a szövegek mondat- és tokenszámát, lemmatizált változatát, majd morfológiai és szintaktikai elemzését (Zsibrita és mtsai, 2013). Ezt követően a szövegek szemantikai és pragmatikai sajátságait különböző szótárak segítségével elemeztük, például a nyelvi bizonytalanságot jelölő elemek (Vincze, 2014), a szentimentek és az emóció szótáraival (Szabó és mtsai, 2016; Szabó, 2015), és kiszámoltuk az egyes lexikális elemek számát és gyakoriságát. Az automatikus elemzés alapján kiszámított nyelvi jellemzőket az alábbiakban mutatjuk be.

– **Statisztikai jellemzők:**

- mondatok száma,
- szavak száma,
- lemmák száma,
- mondatok átlagos hossza.

– **Morfológiai jellemzők:**

- főnevek, igék, melléknevek, határozószavak, tulajdonnevek, névmások, számnevek, kötőszavak, írásjelek és ismeretlen szavak száma és aránya a szószámhoz képest,
- múlt és jelen idejű igék száma és aránya az igék számához képest,
- feltételes és felszólító módú igék száma és aránya az igék számához képest,
- műveltető, ható és gyakorító igék száma és aránya az igék számához képest,
- E/1. és T/1. igék száma és aránya az igék számához képest,
- mutató névmások száma és aránya a szavak számához képest,
- felsőfokú és középfokú melléknevek száma és aránya a melléknevek számához képest.

– **Szintaktikai jellemzők:**

- alanyok, tárgyak, jelzők, határozók, alárendelő és mellérendelő mondatok száma és aránya.

– **Szemantikai jellemzők:**

- tagadószavak száma és aránya a szavak számához képest,
- tartalmas szavak és funkciószavak aránya,
- pozitív és negatív jelentésű szavak száma és aránya a szószámhoz képest (a listákat Szabó (2015) alapján készítettük el),
- bizonytalanságra utaló szavak száma és aránya a szószámhoz képest (Vincze (2014) alapján felállított osztályokba sorolva),
- érzelmekre utaló szavak száma és aránya a szószámhoz képest (Szabó és mtsai (2016) alapján felállított osztályokba sorolva),
- befelé forduló cselekvésre utaló igék („private verbs”) száma és aránya Quirk és mtsai (1985) alapján magyarítva,
- kifelé forduló cselekvésre utaló igék („public verbs”) száma és aránya Quirk és mtsai (1985) alapján magyarítva,
- érvelésre utaló igék („suasive verbs”) száma és aránya Quirk és mtsai (1985) alapján magyarítva.

A szemantikai jellemzők vizsgálatakor egyszerű listaillesztéses módszert használtunk: amennyiben az egyes szavak lemmája megegyezett bármelyik lista-elemmel, akkor találatként számoltuk.

– **Pragmatikai jellemzők:**

- beszédaktust jelentő igék száma és aránya,
- diskurzusjelölők száma és aránya Dér és Markó (2007) alapján,
- szó szerinti idézetek és nyilatkozatok száma és aránya.

Utóbbiakat a szövegben előforduló idézőjelek és mondatkezdő gondolatjelek számával mértük.

A két alkorpuszon kapott elemzési eredményeket végül összevetettük egymással statisztikai szignifikanciavizsgálatok (t-próba) segítségével. Eredményeinket összehasonlítottuk a Propagandaelemző Intézet (Institute for Propaganda Analysis) kutatási eredményeivel is (Propaganda Analysis, 1938).

Annak céljából, hogy a diskurzusban megjelenő időbeli változást még jobban vizsgálhassuk, a gyakorisági jellemzőket lapszámonként számítottuk ki, és eszerint ábrázoljuk majd azokat az eredményeket tárgyaló fejezetben is.

4. Eredmények

Vizsgálataink több sajtóság tekintetében szignifikáns eltérést mutattak a két alkorpusz között, l. a 2. táblázat. Az alábbiakban bemutatunk néhányat a leginkább figyelemre méltóak közül.

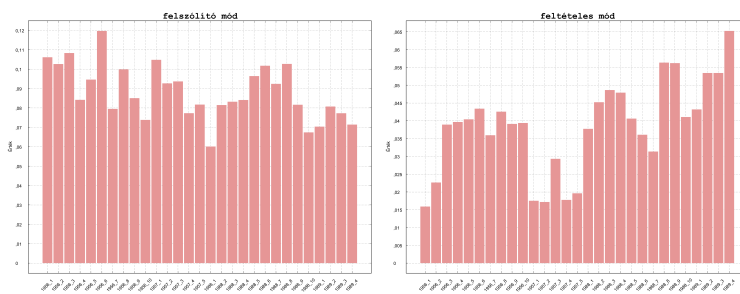
4.1. Morfológiai eltérések

A különböző igemódok használatát illetően azt tapasztaltuk, hogy az első periódusban jóval gyakrabban szerepelnek a felszólító módú igék. Ebből arra következtethetünk, hogy a diskurzus egyes cselekvések elvégzésének szükségességét hangsúlyozhatja az állampárt akaratával összhangban. A fentebbivel szoros

Jellemző	p-érték	Jellemző	p-érték
mondatszám	0,0047	alanyok aránya	<0,0001
tokenszám	<0,0001	tárgyak aránya	<0,0001
mondat hossz	<0,0001	jelzők aránya	0,0002
ismeretlen szavak száma	0,0431	határozók aránya	0,0062
ismeretlen szavak aránya	0,0326	mellérendelések aránya	0,0002
mellénevek száma	0,0171	episztemikus bizonytalanság száma	0,0005
tulajdonnevek száma	0,0465	feltételes bizonytalanság száma	0,0454
mellénevek aránya	0,0017	doxasztikus bizonytalanság száma	0,0003
számnevek aránya	0,0236	episztemikus bizonytalanság aránya	0,0001
felsőfok száma	0,0106	weasel aránya	0,0466
középfok száma	0,0003	peacock aránya	0,0109
felsőfok aránya	<0,0001	hedge aránya	0,0130
középfok aránya	0,0003	doxasztikus bizonytalanság aránya	<0,0001
E/1. igék száma	0,0016	félelem szavainak száma	0,0263
T/1. igék száma	0,0055	düh szavainak száma	0,0151
múlt idő aránya	<0,0001	szorongás szavainak száma	0,0001
jelen idő aránya	0,0029	öröm szavainak aránya	0,0167
feltételes mód száma	0,0007	szorongás szavainak aránya	<0,0001
felszólító mód aránya	0,0258	emotív negatív szavak száma	0,0073
feltételes mód aránya	0,0002	negatív szavak aránya	<0,0001
E/1. igék aránya	0,0005	emotív negatív szavak aránya	0,0010
T/1. igék aránya	0,0013	befelé forduló igék száma	0,0105
ható igék száma	0,0003	meggyőzés igéinek száma	0,0017
műveltető igék száma	0,0471	befelé forduló igék aránya	<0,0001
ható igék aránya	<0,0001	meggyőzés igéinek aránya	0,0001
műveltető igék aránya	0,047	idézetek aránya	0,0037

2. táblázat. Statisztikailag szignifikáns nyelvi jellemzők.

összefüggésben, ugyanebben az alkorpuszban a feltételes módú igék előfordulása jelentősen ritkább, míg a második periódust kifejezetten jellemzi a feltételes módú igék használata, amelyet a bizonytalanság jeleként értelmezhetünk (vö. az 1. ábra).



1. ábra: A felszólító és feltételes módú igék időbeli gyakorisági megoszlása.

A felsőfokú melléknevek és határozószók gyakori előfordulása az első periódusban egyfajta propagandatechnikának is tekinthető: használatuk önbizalmat fejez ki, segítségükkel kiemelhető az adott politikai ideológia jelentősége, valamint a rendszer megkérdőjelezhetetlensége.

Eredményeink korrelálnak egy korábbi dolgozattal (Propaganda Analysis, 1938), amely a propagandaszövegekben fellelhető nyelvi eszközöket vizsgálta. Ilyen például a „glittering generality” jelensége, melynek értelmében a kiemelkedően nagyra becsült fogalmak és hiedelmek a hozzájuk kapcsolódó jelenségek általános és indoklás nélküli elfogadását váltják ki, említésükkel tehát mintegy manipulálható az olvasói elkötelezettség a szöveg tartalma iránt. Az úgynevezett „cherry picking”-elv alapján pedig a bizonyítékok elfedése, vagy a hiányos bizonyítékokon alapuló téves következtetések bizonyos adatokat és eseteket igazolnak, míg az ezeknek ellentmondó adatokat ignorálják.

Az igeidők vizsgálata a két alkorpuszban ugyancsak szignifikáns különbséget mutat. Így például az 1956–57-es periódusban jelentősen gyakoribbak a múlt idejű igék. Az adatok kézi elemzése megmutatta, hogy a múlt idejű igékkal a szövegek az 1956-os forradalom előzményeire, és az ahhoz vezető politikai döntésekre utalnak, az azokhoz vezető okokat, problémákat taglalják.

Megvizsgáltuk a két alkorpuszban előforduló igealakok személyét és számát is. Itt azt tapasztaltuk, hogy a második alkorpuszban lényegesen magasabb számban fordulnak elő a többes szám első személyű alakok, mint az elsőben. A két időszak adatainak összevető, kézi elemzése megmutatta, hogy amíg az 1956–57-es szövegek nagyrészt más, a pártvezetésen kívüli politikai és társadalmi szereplőkkel foglalkoznak (például azokkal a csoportokkal, amelyek cselekedetei és döntései a forradalomhoz vezettek), az 1988–89 a szövegek szerzői (akik maguk is a párt vezető tisztségviselői) javarészt saját magukról írnak.

4.2. Szemantikai eltérések

A szentiment- és emócióelemzés eredményeit illetően, a korszak elejének alkorpuszát kifejezetten jellemzik a pozitív érzelmet kifejező elemek, mint például az öröm és a szeretet. Mindez összhangban van Propaganda Analysis (1938)-nek a már korábban említett eredményeivel: a propagandaszövegekben gyakoriak a pozitív érzelmet kiváltó fogalmak és hiedelmek (mint például a *szerelem*, az *ország* és az *otthon*, valamint a *béke*, a *szabadság*, a *dicsőség* és a *becsület*) (vö. „glittering generality”). Ezzel szemben a negatív érzelmei, mint a félelem és a szorongás, gyakrabban fordulnak elő a 1988–89-es időszakban keletkezett szövegekben, amelyek jelzik a kommunista rezsim helyzetével, a közelgő változásokkal kapcsolatos félelmet.



2. ábra: Különböző érzelmek időbeli gyakorisági megoszlása.

A fentiek mellett azonosítottuk a bizonytalanság nyelvi elemeinek különböző típusait is. Vizsgálataink azt mutatják, hogy háromféle bizonytalanság, a weasel (bizonytalan forrás), a peacock (túlzó kifejezések) és a hedge (mennyiségre vonatkozó bizonytalanság) sokkal gyakrabban jelenik meg az első alkorpuszban, mint a második időszakban (Vincze, 2013). A szemantikai bizonytalansággal szemben (Szarvas és mtsai, 2012) a diskurzusszintű bizonytalanság esetén „a hiányzó vagy szándékosan kihagyott információ nem a mondat propozíciós tartalmához kapcsolódik, hanem más tényezőkhöz”, pl. a *néhány*, *gyakran*, *sok* stb. kifejezések nem adják meg a szóban forgó dolog vagy esemény pontos számát, illetve gyakoriságát. Eredményeink alapján a fentebbiekhez hasonló kifejezések okozta torzí-

tás a propagandadiskurzus jellegzetes vonásának tekinthető. A diskurzusszintű bizonytalansággal szemben az úgynevezett szemantikai bizonytalanság elemei (pl. *lehet, lehetséges, hisz* stb.) csupán ritkán fordulnak elő az első időszakban. Szemantikai bizonytalanság esetében a bizonytalanságjelölő szemantikai tartalma felelős a bizonytalanságért (Vincze, 2014). Például, az első alkorpuszban a szemantikai bizonytalanság episztemikus és doxasztikus (azaz hiedelmekre vonatkozó) típusai elég ritkának mondhatóak. Eredményeink ismét korrelálnak a propagandaszövegek nyelvi eszközeivel kapcsolatos korábbi kutatási eredményekkel (Propaganda Analysis, 1938): a „glittering generality” elvével összefüggésben, a propagandaszövegek egyik fontos jellemzője a pozitív érzelmet kiváltó kifejezések gyakori használata, amelyek inherensen hordozzák a meggyőződést és nem igényelnek indoklást. Ezekből az eredményekből arra következtethetünk, hogy az első alkorpusz szövegei implicit módon megtévesztőbbek a második alkorpusz szövegeinél.

Az igéket további, szemantikai vizsgálatoknak is alávetettük a következőképpen: Megvizsgáltuk a „kifelé fordulás” igéit (public verbs, pl. *bejelent, fenntart, igazol* stb.), a „befelé fordulás” igéit (private verbs, pl. *reménykedik, ítéel, feltételez* stb.), valamint az érvelés igéit (suasive verbs, pl. *egyetért, javasol, elismer* stb.). Az eredmények közül a legszembeötlőbb az volt, hogy a második vizsgált korszakban jelentősen gyakoribbak a befelé fordulás igéi a kifelé fordulás igéivel szemben, ami utalhat a rendszer elbizonytalanodására, a hatalom erejének a csökkenésére. Az érvelés igéi ugyancsak a második korszakban gyakoribbak, amire egy lehetséges magyarázat, hogy a gyengülő rendszerben a *Pártélet* szerzői fokozott figyelmet fordítottak a párt elveinek propagálására.

Az alkorpuszok szemantikai tartalmára vonatkozó vizsgálati eredményeink alapján összességében elmondható, hogy amíg az első időszak szövegei az erős és magabiztos, ugyanakkor a megtévesztés nyelvi jegyeit erősen magán hordozó kommunikáció jegyeit viselik magukon, a politikai korszak végét egy jelentősen kevesebb erőt sugárzó, félelemmel és idegességgel teli, önreflexív diskurzus jellemzi.

4.3. Lexikális eltérések

Végezetül összehasonlítottuk a két időszakban a különböző főnevek, igék és melléknevek gyakorisági megoszlását. Általánosságban elmondható, hogy alapvetően ugyanazok az elemek fordulnak elő mindkét alkorpuszban az egyes szófajok esetében, ugyanakkor az elemek gyakorisága jelentősen eltér. Az előbbi sajátosság egyrészt a szövegek tematikájával, másrészt a propaganda jelleggel is magyarázható. Az utóbbi tulajdonság ugyanakkor arra mutat, hogy a hangsúlyok jelentősen mások a politikai korszak végén annak elejéhez képest.

Az első alkorpuszban a legtöbbször előforduló főnevek között sok olyat találunk, amelyek szemantikailag a vezetőség hatalmával, rendelkezéseivel kapcsolatosak (pl. *terv, teljesítés* stb.). Ugyanakkor a második alkorpuszban a leggyakrabban előforduló főnevek között sok, az eltérő vélemények és a választás lehetőségére utaló elem szerepel (pl. *lehetőség, döntés, vélemény* stb.).

ta korszak propagandanyelvének elemzésére. Annak céljából, hogy megállapításainkra alapozva a vizsgált diskurzusról általános érvényű jellemzést adhassunk, a kutatás fontos, további lépése volna, hogy az itt alkalmazott elemzési módszerek és eszközök segítségével a vizsgált korszakban keletkezett, nem propagandisztikus szövegek elemzését is elvégezzük, eredményeinket azok nyelvi sajátágaival összehasonlítsuk. Itt azonban fontos megjegyeznünk, hogy a kommunista korszakból – a párt direkt politikai agitációjával összefüggésben – nem áll rendelkezésünkre olyan kiadvány, amelynek tartalmát bizonyosan semleges, propagandától „mentes” szövegnek tekinthetnénk. A fentebbi összevető vizsgálatot tehát nem tartjuk lehetségesnek. E problémával összefüggésben azt tervezzük, hogy a kutatás következő lépéseként rendszerváltást követően publikált sajtóanyagokat vizsgálunk, és azokat vetjük össze az itt bemutatott eredményeinkkel.

Munkánk hosszú távú célja egy olyan elemző algoritmus létrehozása, amely képes hatékonyan elkülöníteni a politikai agitációt más, propagandától mentes szövegektől.

Köszönetnyilvánítás

A kutatást részben az Emberi Erőforrások Minisztériuma támogatta (TUDFO/47138-1/2019-ITM).

Hivatkozások

- Barrón-Cedeño, A., Jaradat, I., Martino, G.D.S., Nakov, P.: Propy: Organizing the news based on their propagandistic content. *Information Processing Management* 56(5), 1849–1864 (2019), <https://doi.org/10.1016/j.ipm.2019.03.005>
- Dér, Cs.I., Markó, A.: A magyar diskurzusjelölők szupraszegmentális jelöltsége. In: *Nyelvelmélet–nyelvhasználat*, pp. 61–67. Tinta, Székesfehérvár–Budapest (2007)
- Fetzer, A.: „And I Think That Is a Very Straightforward Way of Dealing With It”– The Communicative Function of Cognitive Verbs in Political Discourse. *Journal of Language and Social Psychology* 27, 384–396 (12 2008)
- Fraser, B.: Chapter 11. hedging in political discourse. In: *Perspectives in Politics and Discourse*, pp. 201–214 (01 2010)
- Girlea, C., Girju, R., Amir, E.: Psycholinguistic Features for Deceptive Role Detection in Werewolf. In: *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. pp. 417–422. Association for Computational Linguistics, San Diego, California (Jun 2016)
- Graves, L.: Understanding the promise and limits of automated fact-checking. Tech. rep., Reuters Institute – University of Oxford (2018)
- Greene, S., Resnik, P.: More than words: Syntactic packaging and implicit sentiment. In: *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*. pp. 503–511. Association for Computational Linguistics, Boulder, Colorado (Jun 2009), <https://www.aclweb.org/anthology/N09-1057>

- Jalilifar, A.R., Alavi, M.: Power and Politics of Language Use: A Survey of Hedging Devices in Political Interviews. *The Journal of Teaching Language Skills (JTLS)* 3, 43–66 (2011)
- Propaganda Analysis, I.f.: How to Detect Propaganda. In: *Propaganda Analysis. Publications of the Institute for Propaganda Analysis. vol. I*, pp. 210–218 (1938)
- Quirk, R., Greenbaum, S., Leech, G., Svartvik, J.: *A Comprehensive Grammar of the English Language*. Longman, London (1985)
- Rashkin, H., Choi, E., Jang, J.Y., Volkova, S., Choi, Y.: Truth of varying shades: Analyzing language in fake news and political fact-checking. In: *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. pp. 2931–2937. Association for Computational Linguistics, Copenhagen, Denmark (Sep 2017)
- Rubin, V.L., Chen, Y., Conroy, N.J.: Deception detection for news: Three types of fakes. In: *Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community*. pp. 83:1–83:4. ASIST '15, American Society for Information Science, Silver Springs, MD, USA (2015), <http://dl.acm.org/citation.cfm?id=2857070.2857153>
- Scheithauer, R.: Metaphors in election night television coverage in Britain, the United States and Germany. In: *Political Discourse in the Media: Cross-cultural perspectives*, pp. 75–106 (01 2007)
- Simon-Vandenberg, A.M., White, P., Aijmer, K.: Presupposition and 'taking-for-granted' in mass communicated political argument An illustration from British, Flemish and Swedish political colloquy. In: *Political Discourse in the Media*, pp. 31–74 (01 2007)
- Szabó, M.K.: Egy magyar nyelvű szentimentlexikon létrehozásának tapasztalatai és dilemmái. In: *Segédkönyvek a nyelvészet tanulmányozásához 177*, pp. 278–285. Tinta, Budapest (2015)
- Szabó, M.K., Ring, O., Nagy, B., Kiss, L., Koltai, J., Berend, G., Vidács, L., Kmetty, Z.: Mapping the dynamic change of the concept „industry” and „agriculture” in the Hungarian Socialist era using a word embedding model. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*. Publikálásra benyújtva (2019)
- Szabó, M.K., Vincze, V., Morvay, G.: Magyar nyelvű szövegek emócióelemzésének elméleti nyelvészeti és nyelvtechnológiai problémái. In: *Távlatok a mai magyar alkalmazott nyelvészetben*, pp. 282–292. Tinta, Budapest (2016)
- Szarvas, Gy., Vincze, V., Farkas, R., Móra, Gy., Gurevych, I.: Cross-Genre and Cross-Domain Detection of Semantic Uncertainty. *Computational Linguistics* 38, 335–367 (June 2012)
- Thorne, J., Vlachos, A.: Automated fact checking: Task formulations, methods and future directions. In: *Proceedings of the 27th International Conference on Computational Linguistics*. pp. 3346–3359. Association for Computational Linguistics, Santa Fe, New Mexico, USA (Aug 2018), <https://www.aclweb.org/anthology/C18-1283>
- Vincze, V.: Weasels, Hedges and Peacocks: Discourse-level Uncertainty in Wikipedia Articles. In: *Proceedings of the Sixth International Joint Conference*

- on Natural Language Processing, Nagoya, Japan, October 2013. pp. 383–391 (2013)
- Vincze, V.: Uncertainty Detection in Natural Language Texts. Ph.D.-értekezés, University of Szeged, Szeged, Hungary (2014)
- Wang, W.Y.: "Liar, Liar Pants on Fire": A New Benchmark Dataset for Fake News Detection. In: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers). pp. 422–426. Association for Computational Linguistics, Vancouver, Canada (July 2017), <http://aclweb.org/anthology/P17-2067>
- Zsibrita, J., Vincze, V., Farkas, R.: magyarlanc: A toolkit for morphological and dependency parsing of Hungarian. In: Proceedings of RANLP. pp. 763–771 (2013)