

Mennyire elérhetőek az antecedensek?

Névmási visszautalások vizsgálata az elérhetőségi elmélet alapján.

Újdonságok a szemantikai és pragmatikai kutatásokban 2018. 04. 27.

Kovács Viktória





I Elméleti háttér

(Ariel 1990, 2001, 2014)

Az elérhetőségi elmélet alapelvei

- A beszélő egy bizonyos információ átadására, úgy választja meg a kifejezést, hogy a hallgató a diskurzus adott pillanatnyi állapotából értelmezni tudja azt.
- A kifejezés formája jelzi azt, hogy az információ mennyire elérhető a címzett számára.
- Különböző típusú kifejezések különböző fokú mentális elérhetőséget feltételeznek.


Az elérhetőséget mutató tényezők:

- Informativitás
- Rigidség
- Fonológiai méret

További tényezők

- Topik pozíció
- Alany eset
- Szövegen belüli pozíció
- Élő/ Élettelen

[1] Accessibility Marking Scale

A vertical double-headed arrow is positioned to the left of the scale items, pointing upwards towards 'Low Accessibility' and downwards towards 'High Accessibility'.

	Low Accessibility
a	Full name + modifier
b	Full ('name') name
c	Long definite description
d	Short definite description
e	Last name
f	First name
g	Distal demonstrative + modifier
h	Proximal demonstrative + modifier
i	Distal demonstrative (+NP)
j	Proximal demonstrative (+NP)
k	Stressed pronoun + gesture
l	Stressed pronoun
m	Unstressed pronoun
n	Cliticized pronoun
o	Extremely High Accessibility Markers (gaps, including pro, PRO and <i>wh</i> traces, reflexives, and Agreement – some of which will be discussed in Part II).
	High Accessibility

Hogyan lesz valami elérhető?

I. Fizikai kontextus

- észleljük/érezkeljük

II. Diskurzusuniverzum

- Szaliencia
- Említések száma
- Új diskurzusraferens

Anaforikus viszonyok az elérhetőségi elmélet nézőpontjából

- Az anafora és az antecedens között is érvényesül az elérhetőség
- Ha prominens az objektum, akkor valószínű, hogy rendszeresen említjük, tehát visszautalunk rá
- Minden egyes említéssel egyre elérhetőbbé válik az objektum
- Minél elérhetőbb valami, annál messzebből is vissza tudunk rá utalni (globális topik)



II Kutatási kérdések

A fizikai kontextus és a diskurzusuniverzum befolyásolja, hogy mely objektumok elérhetőek




I Eltérő képet mutathatnak-e az elérhetőség szempontjából az írott szövegek a spontán beszédétől?

(Konrod, Engel 2001; Toole 1996; Fox 1987;)

II Milyen tulajdonságai vannak az írott beszélt nyelvnek az elérhetőségi elmélet szempontjából?

III A névmási visszautalás stratégiájában vannak-e különbségek a három szövegtípus között

- A visszautaló szó típusában
- Az antecedens típusában
- A visszautalás távolságában



III A vizsgálat

Adatok

Írott szöveg

- Magyar Tudomány bevezetői
- Kortárs novellarészletek
- Élet és Tudomány rövid cikkei
- Index cikkei

Írott beszélt nyelv

- Blogbejegyzések (Prohardver - Fujitsu blog, Játéknapló blog, Könyvkritikák blog, Otthonédes blog, Webisztán blog, Filmvilág blog)

Beszédátirat

- Vincze, Hoffmann, Szatlóczki, Bíró, Gosztolya, Tóth, Pákáski, Kálmán a Szegedi Tudományegyetemen zajló az enyhe kognitív zavar automatikus azonosítására törekvő projektjéhez készült korpusz kontroll csoportra vonatkozó része.

(Vincze et al. 2015)

A kategóriák általános adatai

	Szövegek száma	Szavak száma	szöveg/ szó	szöveg/ visszaütalás	NP-k száma	CP-k száma
BESZÉLT	30	2996	99,87	3,1	1159	760
ÍROTT-BESZÉLT	30	12449	414,97	10,2	4897	2138
ÍROTT	30	10308	343,6	7,67	3921	1757

Módszertan

- Magyarlanc előelemzése (parse modul (szegmentálás, POS-tag, dependencia és konstituens elemzés))
- Az előelemzett szövegek átalakítása az MMAX2-höz
- Annotálás az MMAX2-vel
- Az annotált szövegekből kinyert adatok értékelése

Annotálási elvek

Névmási visszautalószó lehetséges címkéi

- Személyes névmás
- Mutató névmás
- Vonatkozó névmás
- Visszaható névmás
- Ismétlés

Antecedens lehetséges címkéi

- Módosítóval ellátott név
- Név
- Hosszú határozott névelős kifejezés (+2 tartalmas szó)
- Rövid határozott névelős kifejezés (2 tartalmas szó)
- Módosítóval ellátott távolra mutató névmás (+ NP)
- Módosítóval ellátott közelre mutató névmás (+ NP)
- Távolra mutató névmás (+ NP)
- Közelre mutató névmás (+ NP)
- Névmás
- Határozatlan névelős kifejezés

(teljes propozíció, kvantifikált kifejezés)



IV Eredmények

NP-k megoszlása az egyes szövegekben

%	Name+Mod.	Name	Longdef.Descr.	Shortdef. Descr.	Distdem.+Mod.	Proxdem.+Mod.	Distdem.	Proxdem.	Pronoun
BESZÉLT	0	0	1,898	34,254	0,259	0,863	0,259	1,381	19,586
ÍROTT-BESZÉLT	0,179	6,555	10,557	29,630	0,694	0,184	0,572	1,062	12,640
ÍROTT	0,179	3,647	12,089	27,085	0,306	0,153	0,765	1,377	14,614

Az antecedensek megoszlása

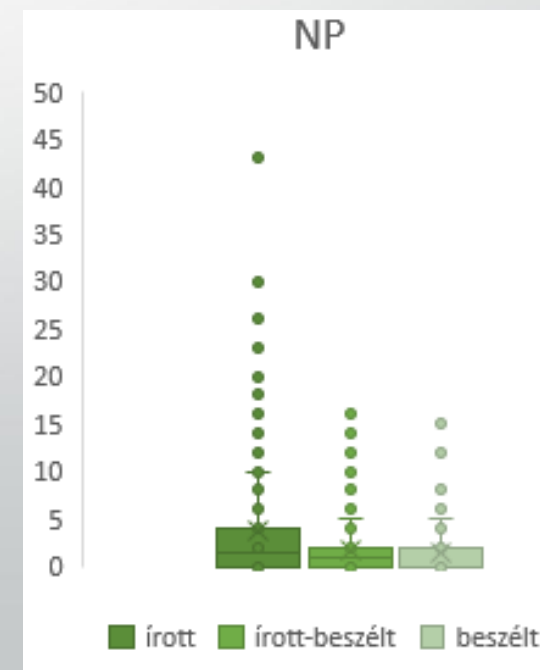
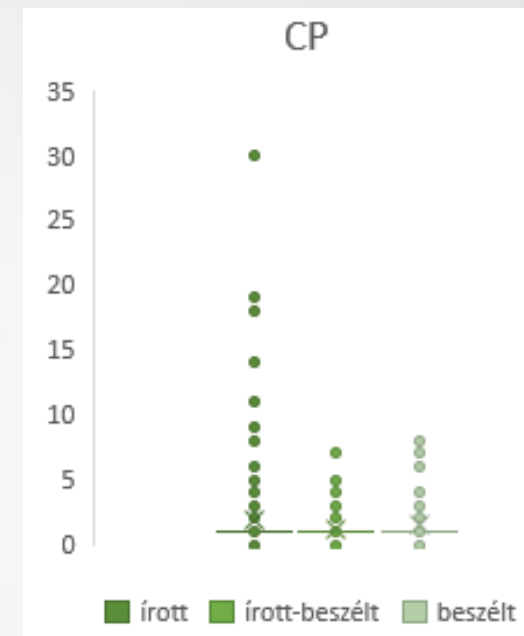
%	Name+Mod.	Name	Longdef.Descr.	Shortdef. Descr.	Distdem.+Mod.	Proxdem.+Mod.	Distdem.	Proxdem.	Pronoun	Indef.Descr.
BESZÉLT	0	0	0	37,634	0	2,151	2,151	3,226	9,677	35,484
ÍROTT-BESZÉLT	1,645	10,855	8,224	22,368	2,303	0,329	4,276	1,645	5,592	8,224
ÍROTT	0,877	16,230	13,597	12,281	1,316	0	6,579	0,877	9,211	7,895

A visszautaló névmások megoszlása

%	Személyes névmás	Mutató névmás	Vonatkozó névmás	Visszaható névmás	Ismétlés
BESZÉLT	22,581	30,108	45,161	1,075	1,075
ÍROTT-BESZÉLT	13,726	32,353	52,288	1,307	0,327
ÍROTT	22,609	30,435	41,304	3,913	1,739

A visszautalás távolsága

	Átlagosan a versengő antecedensek száma	A visszautaló szó és az antecedens közötti távolság	A visszautaló szó és az antecedens azonos tagmondatban (%)	A visszautaló szó és az antecedens szomszédos tagmondatban (%)
BESZÉLT	1,355	1,446	5,376	76,344
ÍROTT-BESZÉLT	1,729	1,1405	2,941	87,582
ÍROTT	3,783	1,791	2,609	74,348



A visszautalás távolsága 2

NP	Name+Mod.	Name	Longdef.Descr.	Shortdef. Descr.	Distdem.+Mod.	Proxdem.+Mod.	Distdem.	Proxdem.	Pronoun	Indef.Descr.
BESZÉLT	0,000	0,000	0,000	0,657	0,000	0,500	0,000	0,667	5,222	1,250
ÍROTT- BESZÉLT	1,286	1,795	1,706	1,507	1,000	0,000	0,400	2,000	2,900	0,656
ÍROTT	2,000	8,368	1,679	3,387	0,667	0,000	0,200	0,500	7,238	1,000
CP										
BESZÉLT	0,000	0,000	0,000	1,143	0,000	1,000	1,000	1,333	3,111	1,406
ÍROTT- BESZÉLT	1,000	1,205	0,971	1,293	1,143	1,000	1,000	1,000	1,550	0,969
ÍROTT	1,000	1,658	1,036	2,419	1,000	0,000	1,000	1,000	3,810	1,056

T-próba

	p
Name + mod.	0,1177023
Name	0,420200679
LongDef.Descr.	0,32917663
ShortDef.Descr.	0,003085503
Dist.Dem.Mod.	0,086014614
Prox.Dem.Mod.	0,162790994
Dist.Dem.	0,3768
Prox.Dem.	0,217994
Pronoun	0,343284
Indef.Descr.	0,1836566

	p
PersonalPron.	0,318456
DeterminativePron.	0,0724
RelativePron.	0,00146
ReflexivePron.	0,0841344
Repetition	0,124

	p
NP	0,03267037
CP	0,091481942
azonos CP	0,299538
szomszédos CP	0,004002392

Konklúzió

- A szövegben található NP-k megoszlásában az írott beszélt nyelv inkább az írott szöveghez áll közelebb az elérhetőségi elmélet alapján.
- Az antecedensek megoszlásában az írott beszélt nyelv az írott szöveg és a spontán beszéd között helyezkedik el.
- A visszautalások megoszlásában nincs különbség.
- A visszautalás hatóköre szempontjából az írott beszélt nyelv a spontán beszédhez áll közelebb.

A kutatás hasznosíthatósága

- További tulajdonságok, amelyek figyelembe vehetőek az automatikus anaforafeloldás során
- Árnyaltabb kép az írott beszélt nyelvről
- Határozatlan névelős kifejezések elhelyezése az elérhetőségi skálán

További kutatási lehetőségek

- Az összehasonlításra alkalmasabb spontán beszédanyag gyűjtése
- EKZ-s minta összehasonlítása a kontroll mintával az elérhetőségi elmélet szempontjából valamint a visszautalási távolság szempontjából
- Az írott beszélt nyelv további szövegtípusainak vizsgálata (chat, e-mail, twitter/ facebook bejegyzés)



Köszönöm a figyelmet!

Felhasznált irodalom

Ariel Mira. 1990, 2014. *Accessing noun-phrase antecedents*. Croom Helm linguistics series. Routledge, London ; New York.

Ariel Mira 2001 Accessibility theory: An overview. In Ted Sanders, Joost Schliperoord and Wilbert Spooren eds. Text representation. John Benjamins (Human cognitive processing series)

Bódi Zoltán. 2004. A szóbeliség kifejeződése az internetes kommunikációban. *Információs társadalom : társadalomtudományi folyóirat* (4. évf. 1. sz.):26–38.

Crystal, David 2001. *Language and The Internet*. Cambridge University Press.

Érsok Nikoletta. 2003. Írva csevegés – virtuális írásbeliség. *Magyar Nyelvőr* (127.évf. 1.sz.), January.

Fox Barbara (1987). *Discourse structure and anaphora*. Cambridge: Cambridge University Press.

Kronrod Ann, Engel Orit (2001) Accessibility theory and referring expressions in newspaper headlines, *Journal of Pragmatics* 33, 683-699,

Müller Christoph, Strube Michael (2006): Multi-Level Annotation of Linguistic Data with MMAX2. In: Sabine Braun, Kurt Kohn, Joybrato Mukherjee (Eds.): *Corpus Technology and Language Pedagogy. New Resources, New Tools, New Methods*. Frankfurt: Peter Lang, pp. 197-214. (English Corpus Linguistics, Vol.3).

Nielsen, Jakob 1997. How Users Read on the Web. <http://www.useit.com/alertbox/9710a.html>.

Toole Janine 1996. The effect of genre on referential choice. In: Fretherim Thorstein and Jeanette K. Gundel, eds., *Reference and referent accessibility*, 263-287. Amsterdam: Benjamin's.

Veszelszki Ágnes. 2010. Digilektus a lektusok rendszerében. In Márta Illés-Molnár, Zsuzsa Kaló, Laura Klein, and Andrea Parapatics, editors, *Félúton 5.: az ELTE BTK Nyelvtudományi Doktori Iskolájának konferenciája : 2009. április 22-23.*, pages 199–215. L'Harmattan, Budapest.

Példák

	ÍROTT	ÍROTT-BESZÉLT	BESZÉLT
Name+mod.	a jól ismert Stroop-teszt	a tizenkét éves Rutát	-
Name	Péteri Gáborét	Clicker Heroes	-
Long.Def.Descr.	az elsőként hazánkban elvégzett funkcionális mágneses képalkotási vizsgálatot	a világon mindenhol nagy sikerrel vetített filmről	-
Short.Def.Descr.	a felújításnál	A kvantumszámítógépek	az idétlen inas
Dist.Dem. + mod.	olyan erőszakos konfliktus	azok a csodálatos illusztrációk	-
Prox.Dem. + mod.	-	Ez a látványos , kisebbfajta csarnok	ez a pincérszerű alak
Dist.Dem.	az az elképzelés	az a játék	az a szereplő
Prox.Dem.	Ebben az utcában	ez a kötet	ennek a hölgynek
Pronoun	ennek	ezt	ezeket
Indef.Descr.	egy régi vastelep	egy emberszerű lényel	Egy kicsi dobozból