

12. fejezet

1. a) *Mi az üzenet?*

b) -- --. .. - --. --- --
-- --- .-. .. - .- .-.- .- .- .. - .- .. ---

2. Ha az elírt szó is helyes alak, csak mást jelent. Továbbá ismeretlen összetett szavaknál, ad hoc formáknál.

3. Az ideális keresőprogram biztonságosan megkülönbözteti a magyar és nem magyar tartalmú oldalakat (ha magyar a keresőszó), a lehető legtöbb oldalt megtalálja, és a legfontosabbakat listázza elsőként, nem csak szóalakokat, hanem a toldalékolt alakokat is megkeresi. Idézet esetén olyan releváns oldalakat keres, amelyekben a beírt kifejezés minden szava az adott sorrendben szerepel...stb.

4. A keresési beállítások

szóalak hoz szófaj: ige...

ami mellett van közvetlenül

szóalak szófaj: névszói jellegű...

eset: tárgy

Részkorpuszra korlátozás:

Régió: Magyarország

Szlovákia

Kárpátalja

Erdély

Vajdaság

Stílusréteg: sajtó

A lekérdezett korpusz mérete: 71 012 205 szó
Találatok száma: 2362 (33,26 / millió szó)
Ez alapján a leggyakoribb a véletlen mintában a *döntést hoz* kifejezés.
intézkedéseket **hoz**. Lakatos Zoltán, a
döntést **hoz** az ember? Spórolni kezd
hoz változást a szeptember: az intézményeknek
összeget **hoz** idén az államháztartás konyhájára -
ítéletet **hoz** a Magyar Televízió Közalapítvány
kuratóriuma
változást **hoz**. A vezeték nélküli internetelés
hoz profitot? - tehetnénk fel a
hoz döntést, hogy a versenyben maradt
forintot **hoz** Paisék és Kaposi Erzsébet konyhájára
céget **hoz** létre azzal a céllal,
részpiacokat **hoz** létre, amelyeken ráadásul eltérő
alapot **hoz** létre Az Intel 250 millió
ügynökséget **hoz** létre pristinai vagy skopjei

székhellyel

bizottságot **hoz** létre a kelet-timori túlkapások kivizsgálására

Testőrákadémiát **hoz** a cseheknek Ceausescu volt biztonsági

bevételt **hoz** a konyhára az istállóknak.

programokat **hoz** létre, amelyekbe az alacsonyabb

felvirágzást **hoz** majd Ausztriában, hogy vagy

hoz eredményt, akkor a harmadik fordulóban

hoz enyhülést. Riporter: - A

hozamokat **hoz** biológiai körülmények között. Ezért

technológiát **hoz** be, munkahelyet teremt,

fűszereket **hoz**, Omara Portuondo pedig arra

intézményt **hoz** létre. Ám nemhogy új

üzemet **hoz** létre Csehországban. A tajpeji

hoz döntést az ügyben. (é.

nyereséget **hoz** számára. Tőlünk is elvárták

szenet **hoz** be, ráadásul az első

fejlődést **hoz**, mint ami az átalakulás

Szerencsét **hoz** a tizenharmadik Magyar Nagydíj?

konzorciumot **hoz** létre, s átadja a

példát **hoz** rá Fontaine személyében. Ebben

fordulatot **hoz** az eseményekbe Milly feltűnése.

pénzalapot **hoz** létre a bűncselekmények áldozatai anyagi

bevételt **hoz** az országnak. A Magyar

döntést **hoz** a lakás áráról, várni

kenyeret **hoz**, italt, vidámakat mond

változást **hoz** magával, a globalizációtól a

hoz pezsgést az a nemrégiben elindult honlap

központot **hoz** létre. Az 1994-ben megkezdett

hasznát **hoz**: a jelenlegi átvételi árak

hoz előrelépést a privatizációban, az állam

adóbevételt **hoz**. A fő gond szerintük

társaságot **hoz** létre a Magyar Televízió Bojtár

növekedést **hoz** a keleti kibővítés Szelestey Lajos

állapotot **hoz** létre. Még akkor is

döntést **hoz**. Lapunk érdeklődésére a

Környezetvédelmi

határozatot **hoz**, amely tartalmazza a jogorvoslati

fejlődést **hoz**. Eddig inkább az élmények

döntést **hoz** a beruházás elindításáról, s

5. A Nyelvi Jogok Egyetemes Nyilatkozatáról van szó. Hivatalos magyar fordítása:

„Ez a nyilatkozat nyelvi közösségnek tekint minden olyan - egy bizonyos térségben történelmileg kialakult - emberi társadalmat, akár ezt a térséget elismerik, akár nem, mely népként határozza meg magát, és tagjai saját közös nyelvet fejlesztettek ki a kommunikáció és a kulturális kohézió természetes eszközeként. Az a kifejezés, hogy egy területre jellemző nyelv, egy történelmileg létrejött közösség nyelvére vonatkozik egy ilyen térségben.” (<http://www.mnyknt.hu/>)

Hibákat okozhat: egyes kifejezéseknek nincsen meg a megfelelőjük a fordító szótárában (Declaration – Nyilatkozat), az ekvivalencia hiánya, pl. a forrásnyelvben homonima, a célnyelvben nem az (history – történet, történelem, space – tér, űr).

6. A szabály alapú módszer nyelvi kompetenciaként, míg a statisztikai performanciaként működik. Mivel a szabály alapú a számítógépes nyelvész nyelvérzékén alapul, ennek egyedüli alkalmazásával

nem lehet kizárni bizonyos fokú önkényességet a szabálymegadásban, ezt viszont az adatbázis konkrét nyelvi tényeinek figyelembe vételével ki lehet küszöbölni.

7. Leggyakoribbnak gondolt kifejezések: névelők (a, az, egy) kötőszavak (és, meg, hogy, de...stb), a létige. A szövegben — nyelvészeti szakszövegről lévén szó — nem azok a kifejezések a leggyakoribbak, amelyek a hétköznapi szövegekben, valószínűsíthető a *nyelvészet*, *nyelv*, *szöveg*, *számítógépes* szavak nagy gyakorisága. Az sem mindegy, hogy a toldalékolt változatokat is a kifejezés részének fogjuk-e fel, illetve a kifejezéshez tartozó összetevőket bármilyen szintaktikai sorrendben egy kifejezésnek értelmezzük-e (pl. igekötő+ige kapcsolata).

A szövegben vizsgált kifejezések:

van/vannak: 50 előfordulás, igekötő nélküli, jelen idő, kijelentő mód 3. személyű alakban

kapcsolat: 18 előfordulás, ha minden olyan kifejezést beleszámítunk, ahol ez a szótó (azaz *kapcsolatos* melléknév is!)

„*kapcsolatban van*”: 2 előfordulás

szól: 9 előfordulás, ha minden olyan kifejezést beleszámítunk, ahol ez a szótó (azaz *szóló* igenév is!)

Érdekesség: a szöveg első bekezdésének elemzése a Szószablya számítógépes gyakorisági szótár webkorpuszának segítségével a szó gyakorisági eloszlása szerint rendezve: (<http://szotar.mokk.bme.hu/szoszablya/searchw.php>)

word (szó)	lemma (szótó)	freq (a szó gyakorisága a webkorpuszban)	lemmafreq (a szótó gyakorisága a webkorpuszban)	analysis (elemzés)
a	a	52706200	71143086	ART
és	és	12674699	12674699	CONJ
vagy	vagy	2282845	2282845	CONJ
ez	ez	2104567	5696473	NOUN
van	van	1788400	7009268	VERB
minden	minden	1191571	1400393	NOUN
mert	mert	1033220	1033220	CONJ
olyan	olyan	890239	1060194	ADJ
pedig	pedig	782771	782771	CONJ
amely	amely	617921	1617001	NOUN
kapcsolatban	kapcsolat	160520	473128	NOUN<CAS<INE>>
újabb	új	146437	934334	ADJ[COMPAR]/ADJ
valóban	valóban	132641	132641	ADV
természetes	természetes	131167	340661	ADJ
olyan	olyan	118661	1060194	ADV
nehéz	nehéz	116525	218255	ADJ
tartozik	tartozik	107946	276378	VERB
vagy	van	102171	7009268	VERB<PERS<2>>
műszaki	műszaki	96242	97426	ADJ
ide	ide	87043	116134	ADV
gyakorlati	gyakorlati	81687	81964	ADJ
rendkívül	rendkívül	79488	79521	ADV
adni	ad	79231	742526	VERB<INF>
tevékenység	tevékenység	74618	345697	NOUN
leginkább	leginkább	70685	70685	ADV
ennél	ez	58130	5696473	NOUN<CAS<ADE>>
foglalkozik	foglalkozik	40951	178098	VERB
nyelvű	nyelvű	40036	40422	ADJ
elméleti	elméleti	38556	38813	ADJ

számítógépes	számítógép	37284	121553	NOUN[ATTRIB]/ADJ
tudomány	tudomány	34211	64093	NOUN
ide	ide	29091	116134	PREV
szolgáltatásokat	szolgáltatás	17406	279378	NOUN<PLUR><CAS<ACC>>
szövegek	szöveg	11157	140722	NOUN<PLUR>
nevén	név	9103	425941	NOUN<POSS><CAS<SUE>>
takar	takar	4718	20545	VERB
szerteágazó	szerteágazó	3700	4369	ADJ
nyelvészet	nyelvészet	2245	3077	NOUN
feldolgozásával	feldolgozás	1842	50020	NOUN<POSS><CAS<INS>>
meghatározást	meghatározás	1322	66782	NOUN<CAS<ACC>>
nyelvekkel	nyelv	442	223640	NOUN<PLUR><CAS<INS>>
mert	mer	133	40652	VERB<PAST>
szabatosabb	szabatos	72	2046	ADJ[COMPAR]/ADJ

8. A megfejtés: *állófogadás*. Mivel a morfológiai elemző kontextustól függetlenül vizsgál, és nincs megfelelő világismerete, meg fogja találni a szóalakon belül az értelmes morfokat, annak ellenére, hogy ez az összetett kifejezés a feladatban leírt jelentésében nem használatos a magyarban (bár nem elképzelhetetlen, de bizonyára nem ebben a jelentésben fordul elő a leggyakrabban).

9. A beszédzintézis és a gépi beszédfelismerés problematikája tartozik ide.