

Filmek a magyarnyelv-oktatás szolgálatában

Vendégoktatói Konferencia
2024. augusztus 22.

Király Dávid

lektor, Osi Állami Egyetem

Sass Bálint

HUN-REN Nyelvtudományi

Kutatóközpont

Lexikológiai Intézet

kiralydavidmid@gmail.com

sass.balint@nytud.hu



vázlat

1. út a korpuszig
2. demó
3. a korpusz belülről
4. példák
5. hozzáférés
6. a korpusz és a MID
7. tanácsok
8. tervek

ötlet

kutatási terület alapja

célnyelvi környezet hiánya

didaktizált MID videók korlátozott száma

memóriánk kapacitása

technikai fejlettség

ötlet

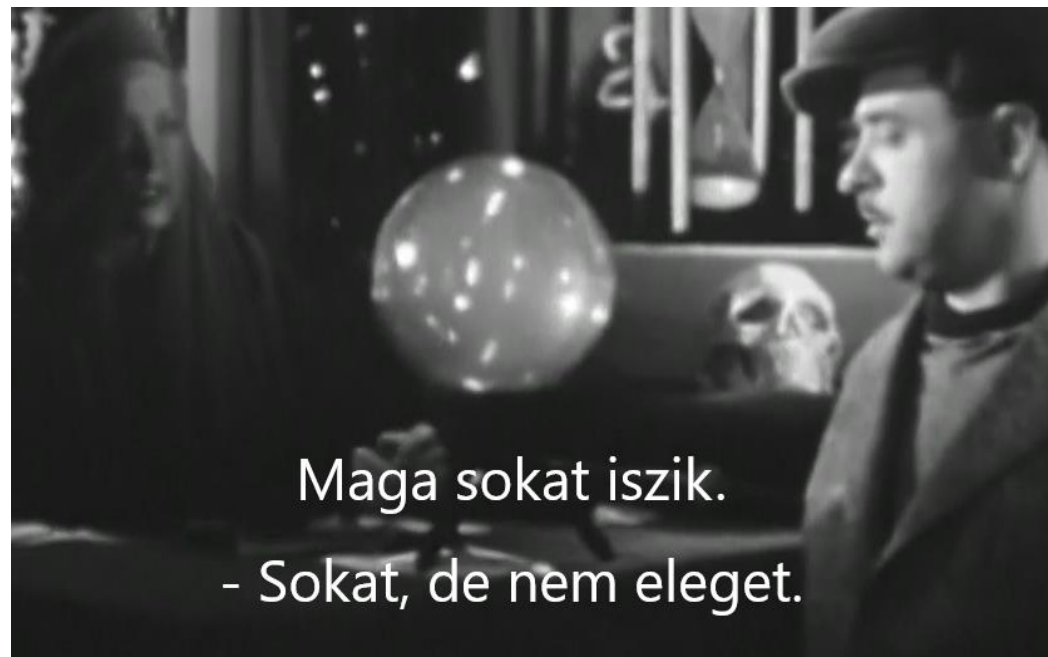
kutatási terület alapja

célnyelvi környezet hiánya

didaktizált MID videók korlátozott száma

memóriánk kapacitása

technikai fejlettség



ötlet

kutatási terület alapja

célnyelvi környezet hiánya

didaktizált MID videók korlátozott száma

memóriánk kapacitása

technikai fejlettség

1

00:00:43,680 --> 00:00:50,028

RÁDIÓ: A jelentős UV-sugárzás

miatt vörösriasztás van érvényben.

2

00:00:50,240 --> 00:00:55,553

Rekordmelegre és lehűlésre,

csapadékra egyaránt számíthatunk.

ötlet

kutatási terület alapja

célnyelvi környezet hiánya

didaktizált MID videók korlátozott száma

memóriánk kapacitása

technikai fejlettség

Num	Pause	Show	Hide	Dura...	Text
1	-	00:01:16,040	00:01:18,110	2,070	- Sakk! - Lehet, hogy matt?
2	0,040	00:01:18,150	00:01:21,310	3,160	- Lehet! No pasarÃ~n! - No pasarÃ~n!
3	0,030	00:01:21,340	00:01:24,000	2,660	- ElvtÃ~rsak, Ã~n nagyon tisztelem!Ã~nÃ~ket.
4	0,040	00:01:24,040	00:01:28,030	3,990	KÃ~rem, olvassÃ~k el ezt. SzeretnÃ~nk szerzÃ~dÃ~st kÃ~t.
5	0,040	00:01:28,070	00:01:30,780	2,710	Az iskolÃ~knak ugyanis szÃ~ksÃ~ge lenne egy veterÃ~nra.
6	0,040	00:01:30,820	00:01:33,510	2,690	- Sajnos tÃ~rgytalan.
7	1,940	00:01:35,450	00:01:37,790	2,340	(Ã~vÃ~s hangja)
8	0,030	00:01:37,820	00:01:40,190	2,370	- Ã~gy van ez.
9	0,040	00:01:40,230	00:01:42,940	2,710	EgyÃ~bkÃ~nt Ã~n mÃ~r kaptam egy jobb ajÃ~nlatot.
10	0,030	00:01:42,970	00:01:45,030	2,060	- De kÃ~rem szÃ~pen, mi takarÃ~tanÃ~nk is.
11	0,040	00:01:45,070	00:01:47,060	1,990	Ablakot is pucolnÃ~nk, mert fÃ~dszinti a lakÃ~s.
12	0,030	00:01:47,090	00:01:49,920	2,830	Tejet is hoznÃ~nk. - SzÃ~t sem lehet rÃ~la!
13	0,030	00:01:49,950	00:01:52,610	2,660	Ã~s egyÃ~bkÃ~nt sem szabÃ~lyos.
14	0,030	00:01:52,640	00:01:56,960	4,320	- Ne itt keresse az emberÃ~t, pajtÃ~s, hanem lenn a pincÃ~ben.
15	1,990	00:01:58,950	00:02:00,880	1,930	- PincÃ~ben? - Ott hÃ~t.
16	0,030	00:02:00,910	00:02:02,940	2,030	Ott vannak a tÃ~slÃ~k.
17	3,630	00:02:06,570	00:02:09,250	2,680	(robbanÃ~s hangja)
18	1,750	00:02:11,000	00:02:14,840	3,840	- ElkÃ~stem. (zene)
19	0,030	00:02:14,870	00:02:16,820	1,950	ElkÃ~stem.
20	0,030	00:02:16,850	00:02:18,760	1,910	(zene)
21	12,090	00:02:30,850	00:02:34,150	3,300	- Jelentsd. - ElkÃ~stem.
22	12,740	00:02:46,890	00:02:48,870	1,980	- RadnÃ~tival prÃ~bÃ~kozott mÃ~r?

demó

1. filmek előkészítése
2. felület: <https://mmkk.nytud.hu>
3. amire kíváncsiak vagyunk: *Mit tudom én!*
4. → időadatok
5. ... voilà!

korpusz



számítógépen tárolt, nagy mennyiségű, *nyelvileg elemzett* szöveg
keresünk benne – nyelvészetileg releváns kérdéseket tehetünk fel

Magyar Mozgóképek Korpusz: 11 millió szó ~ 25000 oldalnyi szöveg

korpusz – belülről

Mit tudom én!

szóalak szótő elemzés

Mit	mi	[/Pro] [Acc]
tudom	tud	[/V] [Prs.Def.1Sg]
én	én	[/Pro] [1Sg] [Nom]
!	!	[Punct]

...

korpusz – időadattal

Mit tudom én!

szóalak szótő elemzés

<time v="01:26:13,927">

← struktúra, időadat

Mit	mi	[/Pro] [Acc]
tudom	tud	[/V] [Prs.Def.1Sg]
én	én	[/Pro] [1Sg] [Nom]
!	!	[Punct]

...

</time>

korpusz – időadattal

Mit tudom én!

szóalak szótő elemzés

<time v="01:26:13,927">

← struktúra, időadat

Mit mi [/Pro] [Acc]

tudom tud [/V] [Prs.Def.1Sg]

én én [/Pro] [1Sg] [Nom]

! ! [Punct]

...

</time>

→ a keresőfelületen a korpuszban kódolt infókra tudunk rákeresni

példák

- *jaj* → sok találat → *jaj* , *de*
- *mit tudom én!* ← többszavas
- *asztal* ← alapból szótóként keresi
- feltételes mód: CQL = " . *Cond. * " + gyakorisági lista
- *fontos* + kollokációk
- *nehogy már* – beszéltnyelvi fordulatok – *bakker*, *basszus* ...

példák

- *jaj* → sok találat → *jaj* , *de*
- *mit tudom én!* ← többszavas
- *asztal* ← alapból szótóként keresi
- feltételes mód: CQL = " . *Cond. * " + gyakorisági lista
- *fontos* + kollokációk
- *nehogy már* – beszéltnyelvi fordulatok – *bakker*, *basszus* ...

további funkciókról és részletekről tájékozódhatunk:

- a felületen az ABOUT részekben,
- a kattintható kérdőjeleknél és
- a SketchEngine youtube csatornáján

hozzáférés

<https://mmk.nytud.hu>

felhasználói név: demo

jelszó: mmkk

Szabadon kipróbálható.

Visszajelzésnek örülünk.

Király Dávid

`kiralydavidmid@gmail.com`

Sass Bálint

`sass.balint@nytud.hu`

kódok felfedezése

1. konkrét szóra keresés
2. kód megjelenítése
3. kód feljegyzése
4. kód használata CQL lekérdezésben: ". *KÓD . *"
default attribute: ana

korpusz és MID

eszköz a nyelvtanításhoz

~1600 film és ~120 sorozat kereshető

számtalan tanórai felhasználási lehetőség

korpusz és MID

eszköz a nyelvtanításhoz

~1600 film és ~120 sorozat kereshető

számtalan tanórai felhasználási lehetőség



Berlinben él a kisunokámmal, Hansszal.

Van egy _____, Ana.

tanácsok

Filmio ~ 700 film

Nemzeti Archívum ~ 500 film és 50 sorozat

más streamingszolgáltatások

VPN vagy dedikált IP-cím

állomáshelyi tájékozódás

tervek

folyamatos bővítés

mideo.hu

zeneszöveg korpusz

Filmek a magyarnyelv-oktatás szolgálatában



Vendégoktatói Konferencia
2024. augusztus 22.

Király Dávid

lektor, Osi Állami Egyetem

Sass Bálint

HUN-REN Nyelvtudományi

Kutatóközpont

Lexikológiai Intézet

kiralydavidmid@gmail.com

sass.balint@nytud.hu