

A MorphoLogic és a magyar számítógépes nyelvészet

A MorphoLogic és a magyar számítógépes nyelvészet című rendezvényen a Neumann János Számítógép-tudományi Társaság Informatikatörténeti Fórum (NJSZT iTF) a hazai nyelvészeti kutatások és fejlesztések történetét mutatta be az Óbudai Egyetemen.

A rendezvény időpontja: 2017. november 24. (péntek) 14:00–18:00
Helyszín: Óbudai Egyetem (Budapest III. Bécsi út 96/B) F01 terem

A program:

Tick József: *Megnyitó* ▶

Prószéky Gábor: *A számítógépes nyelvészet rövid történeti áttekintése* ▶

Kiefer Ferenc: *A gépi nyelvészet hazai megjelenése (Emlékek a valaha volt gépi fordító csoportról)* ▶

Prószéky Gábor, Kornai András: *Papp Ferenc és az „újrafelhasznált” Szóvégmутató szótár* ▶

Prószéky Gábor, Votisky Zsuzsa, Kis Ádám: *A „Számítógépes nyelvészet” c. könyv létrejötte* ▶

Prószéky Gábor: *A MorphoLogic indulása* ▶

Tihanyi László: *Videoüzenet Luxemburgból* ▶

Pál Miklós: *Húsz év 10 percben* ▶

Naszódi Mátyás: *Csatlakozik a SZTAKI...* ▶

Prószéky Gábor: *A MorphoLogic további sikeres termékei és a MorphoLogic Lokalizáció létrejötte* ▶

Merényi Csaba: *MetaMorpho, az első működő magyar gépi fordító rendszer* ▶

Prószéky Gábor: *A MorphoLogic pályázati tevékenysége és nemzetközi jelenléte* ▶

Várad Tamás: *Hazai kutatási együttműködések a MorphoLogic és az akadémiai szféra között* ▶

Olaszy Gábor: *A beszédtechnológia és a MorphoLogic kapcsolatai* ▶

Kis Balázs (Kis Ádám előadásában): *Kilgray – egy MorphoLogic spin-off* ▶

Prószéky Gábor: *Zárszó a nyelvtechnológia jelenéről és jövőjéről* ▶

Prószéky Gábor

**A számítógépes nyelvészet
rövid történeti áttekintése**

1950-60: ötletek

(vannak már gépek)

1960-70: kísérletek

(kialakulnak az igények)

1970-80: programok

(megjelennek a „használható” gépek)

1980-90: termékek

(a gépek kapacitása megnő)

1990-től: technológia

(a kommunikációs helyzet megváltozik)

2000-től: ipar

(egyre több a feldolgozandó szöveg)

2010-től: internet

(mindenhol)

Matematikai nyelvészet és gépi
fordítás az MTA Számítástechnikai
Központjában 1962-1972

- 1960 körül Európa számos országában folytak már számítógépes nyelvészeti kutatások. A számítógépes alkalmazáshoz a létezőknél sokkal pontosabb nyelvtanokra, felhasználásukhoz pedig a számítógép számára érthető elemző algoritmusokra volt szükség. 1960-ban a *Voproszy jazykoznanija* c. szovjet nyelvészeti folyóiratban jelent meg Grigorjev akadémikus rövid összefoglalása a szovjet Akadémia állásfoglalásáról, amely szorgalmazza a gépi fordítási és strukturális nyelvészeti kutatások előmozdítását.

- Az MTA Nyelv- és Irodalomtudományi Osztálya 1961-ben szervezett egy ún. nyelvészaktívát, amely a szovjet akadémia szellemében foglalt állást a számítógépes és strukturális nyelvészeti kutatások előmozdításáról. Kalmár László matematikus akadémikus azt javasolta, hogy az MTA Nyelvtudományi Intézetében jöjjön létre egy strukturális nyelvészeti, az MTA Számítástechnikai Központjában (=SZK) pedig egy számítógépes nyelvészeti kutatócsoportot.

- 1962 márciusban az SZK, a Nyelvtudományi Intézet (NYTUD) és az MTA Elnökségének Kibernetikai Bizottsága konferenciát szervezett a matematikai nyelvészet kérdéseiről. A résztvevők között találjuk a Műszaki Egyetem Nyelvi Intézetének néhány munkatársát (Hell György, Sipőczy Győző), Papp Ferencet (Debreceni Egyetem) és Dömölki Bálintot, Kiefer Ferencet (SZK) és nem utolsósorban Kalmár László és Rényi Alfréd akadémikusokat.

- 1962 őszén az SZK-ban megalakul a számítógépes nyelvészeti, az NYTUD-ban pedig a strukturális nyelvészeti kutatócsoport. Az előbbi három teljes (Kiefer Ferenc, Varga Dénes, Petőfi S.János) és három félállású kutatóval indult és 1973-ban vált önálló kutatócsoporttá, amely 1972-ig működött.
- Az SZK-ban 1960 óta működött már az M-3 számítógép (amely 1965-ben Szegedre került). A számítógép operatív memóriája egy 1024 szó tárolására képes mágnesdob volt, amelyet fokozatosan bővítettek.

- Az M-3 770 elektroncsőből és 2000 diódából állt, nagy meleget termelt, és sok probléma volt a hűtésével. Ennél nagyobb probléma volt, hogy a gép operációs rendszerrel, fordítóprogramokkal nem rendelkezett, és kizárólag gépi kódban lehetett programozni.
- A nyelvi elemzés algoritmusainak gépi kódban történő programozása rendkívül körülményes volt, ezért a kutatók egyre inkább elméleti kérdések megoldására összpontosították figyelmüket.

- 1965-ben az SZK új korszerűbb számítógépet kapott: az URAL 2-t. A számítógép központi egysége elektroncsövekre épült. Elsősorban tudományos számítások végzésére használták. Perifériái közé elektronikus írógép, lyukkártya- és lyukszalag-olvasó, nyomtató, mágnesdob, mágnesszalag tartoztak. Az URAL 2 már alkalmas volt bizonyos nyelvi elemzések elvégzésére is.

- A SZK matematikai nyelvészeti kutatócsoportja 1963-ban indította el a *Computational Linguistics* c. kiadványt. A publikált cikkek a magyarországi számítógépes nyelvészeti munkákról tájékoztattak. A szerzők között nemcsak a SZK munkatársai, hanem egyetemi (Szeged, Debrecen, BME) kutatók is szerepelnek.

- A cikkek túlnyomó többsége az orosz-magyar gépi fordítás kérdéseivel foglalkozott. Az ALPAC (1965) jelentés után azonban a gépi fordítási kísérletek háttérbe szorultak, és a kutatók érdeklődése a számítógépes elemzés egyéb területeire is kiterjedt: folklór, irodalmi művek, népdalok. Ugyanakkor nagyobb hangsúlyt kaptak az elméleti munkák is (a formális grammatikai modellek vizsgálata).

- Az M-3 jelentősége: „idézet”

- A Számítógépes Nyelvészeti Csoport 1972-ig működött. Akkor szűnt meg, amikor az SZK és az AKI (Automatizálási Kutatóintézet) összeolvadásából 1973-ban létrejött a SZTAKI (=Számítástechnikai és Automatizálási Kutatóközpont). A *Computational Linguistics* utolsó száma 1969-ben jelent meg. 1973-ban *Computational Linguistics and Computer Languages* címen jelent meg újra a SZTAKI kiadványaként.

Prószéky Gábor, Kornai András

**Papp Ferenc és az
„újrafelhasznált”
Szóvégmutató szótár**

Éke(zete)s anyanyelvünk

MIKROVILÁG: Beszélgetésünk résztvevői három különböző korosztályt képviselnek. A közös vonás: a számítógépes nyelvészet szeretete. Honnan ered ez az érdeklődés?

PAPP F.: Közel harminc éve, első generációs — tehát elektroncsöves — gépek kezdtem el a munkát. Nagyon lelkesek voltunk. Emlékszem, hogy Kalmár László is előadást tartott számítógépes nyelvészetről, sőt még szociológusok is érdeklődtek a téma iránt.

PRÓSZÉKY G.: 1979-ben, egyetemistaként csöppentem bele a számítógépes nyelvészet izgalmas világába. Akkoriban alkalmazták először PROLOG-ot, s az MTA Nyelvtudományi Intézet és a SZTA KI közös vállalkozásaként magyar morfológiai elemző rendszert dolgoztunk ki.

BESSENYEI G.: Szerencsés csillagzat alatt születtem, hiszen gimnazistaként én már személyi számítógéppel dolgozhatok. Olyan programokat írok, amelyek a nyelvtan tanítását segíthetik.

Szavak, szavak,
szavak

MIKROVILÁG: Meglepő, de tény, hogy a nyelvészek elsők között kezdtek el foglalkozni a



DR. PAPP FERENC
akadémikus
**ELTE Általános
és Alkalmazott
Nyelvészeti Tanszék**
„Örömmel javítottam a hibákat”



PAJZS JÚLIA
MTA Nyelvtudományi
Intézet
*„A számítógépes
mondatgenerálás
rendkívül izgalmas”*



PRÓSZÉKY GÁBOR
Országos Pedagógiai
Könyvtár és Múzeum
*„Az UNESCO is támogatja
a nemzeti törekvéseket”*



BESSENYEI GÁBOR
Berzsenyi Dániel
Gimnázium
*„Szerencsés csillagzat
alatt születtem”*

val szemléltetném. Előfordulhat, hogy azt a szót: *csónak*, a számítógép a „csón” szó többes számaként értelmezi. Az sem lehetetlen, hogy egy ige többes szám harmadik személyű alakjának tekinti, mint például a „futnak” esetében. De az is lehet, hogy egy főnév birtokos személynévragjának értelmezi. Ami egy iskolás gyerek számára is természetes, az komoly nehézséget jelenthet a számítógépnek.

PAJZS J.: Egy másik izgalmas terület a számítógépes mondatgenerálás. A gép képes arra, hogy nyelvtanilag tökéletes mondatokat hozzon létre véletlenszerűen a szótárából. Ilyeneket például: a kutyája jól ugat a szobában este. Ez azonban csak formális szerkezet, a mondatnak nincs „lelke”, gondolati tartalma. Meg tudjuk határozni, hogyan épüljön fel (szintaxis), de azt nem, hogy mit mondjon (szemantika).

PRÓSZÉKY G.: Az ember — beleérző képessége révén — képes arra, hogy nyelvtanilag nem tökéletes mondat alapján is megértse a beszélő szándékát. A gép ilyen szempontból tökéletlenebb. A számítógépes nyelvészek nagy álma, hogy olyan rendszert fejlesszenek ki, amellyel társaloghatunk. Ez nagyon hasznos

MIKROVILÁG

A VSZM Közösségi Házban
a számítógépes nyelvészettel foglalkozó
szakembereket kérdeztük kutatási
eredményeikről.

1987. június

**Mikolás Zoltán, Votisky Zsuzsa,
Kis Ádám**

**Néhány szó a
„Számítógépes nyelvészet”
c. könyv létrejöttéről**

PRÓSZÉKY GÁBOR

SZÁMÍTÓGÉPES NYELVÉSZET

TERMÉSZETES NYELVEK
HASZNÁLATA
SZÁMÍTÓGÉPES RENDSZEREK BEN



Prószéky Gábor

A nyelvtechnológia
(és)
alkalmazásai

Antikvrium.hu

PRÓSZÉKY GÁBOR - NYELVÉSZET

Számítógéppel emberi nyelven

ellenőrzés
helyesírás

elválasztás
el-va lasz-tás

keresés
????????

fordítás
translation



Prószéky Gábor

A MorphoLogic indulása



1991. június

Komjádi Béla utca, 1991



cseréjét, törlését vagy beszúrását engedi meg.

A NyelvÉSZ jelenlétét felismerő Ékszer alapváltozatának ára 19 900 forint + áfa, a hozzá tartozó NyelvÉSZ verzió — amely a Softinvest által forgalmazott változathoz hasonlóan kulcselemmel védett — 10 000 + áfa.

Egy rivális program: Helyes-e

Az Ifabón mutatták be a NyelvÉSZ vetélytársát, a Helyes-e helyesírás-ellenőrző programot. Prószéky Gábor, a fejlesztőcsoport vezetőjét kérdeztük meg a programról, és megkérdeztük, hogy milyen véleménye az elkészült termék jellemzőiről.

Prószéky úr elmondta, hogy tavaly decemberben Tihanyi Lászlóval együtt kezdték a fejlesztést. A rendszerterv,

amely Papp Ferenc és Elekfi László nyelvészeti kutatásainak továbbfejlesztésén alapul, Prószéky Gábor munkája. Alapkonceptióként a mai magyar köznyelv helyesírás-ellenőrzésére kívántak koncentrálni, ezért a program szótárában az *Értelmező kézisztár* szóanyagát, valamint az összes ma használt ragot, jelet, igekötőt és a ma is produktív képzőket találjuk meg. A program a szótárában nem szereplő jelöletlen főnévi összetételeket is felismeri, a különböző szófajok közötti összetételeket pedig letiltja. A program nem ad „javítási tanácsot”, mivel szerzőik véleménye szerint az ilyen jellegű segítség többet árt, mint használ. Későbbi fejlesztésnél is csak a korábbi nemzedékek írógéphasználatából eredő ékezethibák „javítási tanácsadását” tartják megvalósíthatónak. A Helyes-e maga nem tartalmaz elválasztáskezelő modult sem, de összetett szavak esetén visszaadja az összetételhatárt, mivel Prószéky úr szerint az alkalmazói programokba épített elválasztáskezelő modulok — ha megfelelő módon támogatják őket — hatékonyan ellátják a feladatukat. A Helyes-e program önálló változatában 130 kilobájt, szövegszerkesztőbe beépített változatában pedig 90

kilobájt tárat foglal el, lemezigénye pedig 380 kilobájt. Prószéky úr elmondta, hogy a tárigény a jelenleg 60 000 szót tartalmazó adatbázis további töltése esetén sem változik.

Jelenlegi tesztállapotában a program a Computer Media által forgalmazni kívánt Írnokba integrálva működik, és várható, hogy néhány héten belül a szövegszerkesztőtől független, önálló interaktív javításokat is lehetővé tevő változata is megjelenik. Az önálló verzió ára 15 000 forint + áfa körül lesz. Önálló változatát valószínűleg másolásvédelemmel tesztik, de a védelem semmiképpen nem kulcselemmel megoldású lesz.

A távolabbi tervekről Prószéky Gábor elmondta, hogy mivel az alapkonceptió nyelvfüggetlen, a vegyes nyelvű dokumentumok ellenőrizhetősége érdekében tervezik idegen nyelvek szóanyagának beépítését is. A szakszótárak integrálására pedig egy — ingyen terjesztett — program szolgál, amellyel a szóanyag a Helyes-e kívánalmainak megfelelően rögzíthető. Így rögzített szóanyagot a fejlesztőcsoport — külön megállapodás keretében — integrálja a Helyes-e törzsszókészletébe.

Még nem választottak

A NyelvÉSZ sajtótájékoztatóján elhangzott, hogy a Rolitronnal aláírás előtt áll a helyesírás-ellenőrző program Rosytextbe való beépítéséről szóló szerződés.

Németh László, a Kollonon fejlesztőcsoportjának vezetője, valamint a beépítés helyzetéről beszéltek. Németh úr elmondta, hogy nemcsak a NyelvÉSZ-t, hanem konkurenciáját, a Helyes-et is tesztelik, és a tesztfolyamat során szerzett eddigi tapasztalatok — a sebességviszonyok és egyéb tényezők — jelenleg a Helyes-e beillesztése mellett szólnak. Véleménye szerint a NyelvÉSZ szabadságfoka eltűzött, ezért ha mégis megkötődne, akkor maximálisan a cseréjét, beépítését kell engedni meg a javításra kijelölt szóban, valamint letiltják az összetett szavak külön szóként való vizsgálatát is. Nem kívánják felhasználni a NyelvÉSZ elválasztó modulját sem, mivel szövegszerkesztőjük beépített elválasztórendszerrel rendelkezik.

A döntés egy-két héten belül mindenképp megtörténik, a beépített helyesírás-kezelővel kiegészített Rosytext ára körülbelül 20 000 forinttal nő. Tervek szerint a programot védelem nélkül bocsátják piacra.

Nyelv körüli

Lejegyezte: Révész Gábor, átfésülte: Mester Sándor

kerekasztal

Lapunk 1991/21. számában megjelent egy összeállítás a két hazai fejlesztésű magyarnyelv-helyesség-ellenőrző programról, a NyelvÉszről és a Helyes-eről. Az összeállítás vitákat kavart. Szerkesztőségünk az ügyben érdekelte a *Softing*-vel együttműködő

■ Mester Sándor: A szélesebb értelemben vett szakma is örömmel fogadná, ha egy igazi magyar helyesírás-ellenőrző jelenne meg a piacon. Bizonyára minden elektronikus rendszerrel dolgozó magyarországi szerkesztőség hasonlóan vélekedik. Hogy mi a neve annak a bizonyos

■ M. S.: Említette, hogy az összeállítás megjelenése óta történt egy s más.

■ D. P.: Igen, a Computer Media esetében annyi, hogy az Írnok hamarosan meg fog jelenni, ezért e héten szerződést kell kötnünk; pillanatnyilag a Helyes-et teszteljük, amelynek elkészült az Írnokba beépített, tökéletesen integrált változata. Ez a NyelvÉsszel nem történt meg — ez tény. Nincsen az Írnokba integrált, tökéletes változata.

■ M. S.: Nem mintha az Írnok a világ leghíresebb szövegszerkesztője lenne.

■ D. P.: Nem, sőt azt hiszem, hogy a világ legkevésbé híres szövegszerkesztője — kevesen ismerik.

■ M. S.: Tisztázzunk egy tényt! A Helyes-e termék-e? Kapható-e a boltban, mennyibe kerül?

■ D. P.: Igen, a Helyes-e termék, kapható, 15 000 forint + áfáért megvásárolható, jelenlegi forgalmazója a Domine Kft.,



ar
dk
zá
—
ez

■ M. S.: Kizárja a NyelvÉszből való kiindulást?

■ P. G.: Azt hiszem, hogy ők is.

■ B. T.: Nem állítottuk.

■ P. G.: Mindannyian...

Az első vevők...

sorsvevő	cím	telef
1 CWI IDG	Bp. Rákóczi 16	1117-917
6 Softinvest	Bp. Jászai tér 3	112-4873
31 Vén Adatbank	Győr Arpád ut 55	25019
2 DEC Hungary	Bp. Vahot u	1668-011/222
13 Műszertechnika RT	Bp. Venyige u. 3	1476-590
12 Fényszedd Közp.	Bp. Váci u. 78	1182-622
8 Recognita Rt	Bp. Iskola u 12	
18 Mac-Martoni		
33 Európa Alapítvány	1395 Bp. Pf 414	1294516
29 Multi Kiszöv.	Bp. Erzsébet krt. 40	1223913
35 Közműv. Inf. Váll.	Bp. Miklós téri	1803594
27 Akceptor Kft	Bp. Sallai I. u. 52	1498-287
16 Hemingway		
25 Gödöllői Egyetem		28-10-975
14 Hemingway	Bp. Szakasits 68	1853-011
15 Hemingway		
17 HEMINGWAY		
37 Multilingua	Bp. Csalogány u 13.	12020817
11 JATE Ped. Tansz.	Szeged, Petőfi 30	62-21611-84
20 MH REK		1664-611
19 Cedrus Media		

HEJES-E HELYES-E? 2.0

Magyar helyesírás-ellenőrző
program

© MorphoLogic

Prószéky Gábor,
Tihanyi László, Pál Miklós,
Doór Ferenc, 1991

Memóriaigénye: szöveg-
szerkesztőbe beépítve kb. 90
kByte, az önálló változaté kb.
200 kByte

Háttérigénye: kb. 600 kByte

Sebessége: 5-10 mp/A 4
oldal, a gép sebességétől és a
hibák számától függően
(RAM-diszkkal ez még gyorsítható)

Alapszókészlete: az Értel-
mező Kéziszótárban szereplő
minden szó, továbbá mintegy
10 000 szó, beleértve a leg-
gyakoribb tulajdonneveket is.

Ára: 15 000 Ft + ÁFA, min-
den további munkahelyre
5 000 Ft + ÁFA

Érdeklődés: MorphoLogic
Gmk.

1023 Budapest, Komjádi B. u. 1.
Tel.: 115-68-21

Minden ötödik felnőtt használ

A legfrissebb adatok a magyarországi sz

(CMH) A Cédrus Média megbízásából 1991 szeptember 12. és 15. között a GALLUP Közvéleménykutató Intézet felmérésében a magyarországi számítógép-használatra is rákérdezt. A vizsgálat során 1500 embert kerestek meg, akiket úgy választottak ki, hogy reprezentálják az ország teljes felnőtt (18 éves és idősebb) lakosságát.

A felnőtt magyar lakosság közel egyötöde, 18% használ számítógépet munkahelyén, otthonában vagy mindkét helyen. Az emberek 9%-ának van otthonában számítógépe, 13%-uk pedig munkahelyén kerül közvetlen kapcsolatba a számítástechnikával. Ezek közül 4%-nak otthon is van számítógépe. Az, hogy a magyar felnőttek 18%-a valamilyen módon kapcsolatban áll a számítástechnikával, igen jelentős tény a munkakultúra szempontjából is, hiszen (mint alább látni fogjuk) az aktív korosztályokban ez az arány még magasabb.

A munkahelyükön számítógépet használók 30%-ának otthon is van gépe (a munkahelyükön számítógéphez nem kapcsolódók körében ez az arány csak 6%). Az otthoni számítógép-tulajdonosok 42%-a munkahelyén is használ számítástechnikát (az otthoni számítógéppel nem rendelkezőknek csak 10%-a találkozik munka közben számítógéppel). A munkahelyi számítógép-használat és az otthoni számítógép-tulajdonosság között tehát igen szoros a kapcsolat. Érdekes lenne megtudni

lehetősé
zésére?
felelni
használa
kapcsol
összefü
A szár
munkah
biakban
(a telje
Ugyana
munkah
feltehet
tástechn
telben,
területe
zottak
tartozn
mítógép
kőn szá
adminis
az össz
csak 28
tő, hog
számít
egy-egy
mításte
túl nir
többsé
tógépp
63%-a
teli-szá
felhasz

Morpho Logic

Hibés-e?

helyellenőrző
programok.
makrók
ítés

MAGYAR SZINONIMASZÓTÁR

MOR morfológiai elemző programok
előállító program
szótár-építés
felvezetési programok

1137 Radnóti M. u. 2.
Tel.: 112-7490/279

Hibának
helye

hogy ez
anel sose
lthasson elő



Radnóti Miklós utca, 1992



20 év 10 percben

Pál Miklós

palmiklos@gmail.com

Ki vagyok én? (1)

- **MorphoLogic** alapító tag, $1/3^*$ részben tulajdonos, alkalmazásban 1991-től 2011-ig
- Munkakörök: **programozó**, szoftverfejlesztő, szoftvertervező, termékfejlesztő, minőségbiztosító hibaelhárító, termékmenedzser, terméktámogató (felhasználók és partnerek felé)

Ki vagyok én? (2)

- Területek: **Nyelvi eszközök:** (analyser* speller hyphenator thesaurus grammar lemmatizer) önálló termékek, külső beépítések szöveg- és kiadványszerkesztőkbe, beépítőkkel való együttműködések, kiemelten a Microsoft Office beépítések, minőségbiztosítás, a partner követelményeinek való megfelelés elérése
Nyelvek: nem csak magyar, hanem pl. lengyel és román nyelvek feldolgozása

Ki vagyok én? (3)

- **Cseh nyelvtani ellenőrző:** cseh partner Linux-alapú kutatói alkalmazásból Office-beépítés
Szótárak: Rich Text tooltip, Acrobat Reader (PDF) szövegkiolvasó plugin, Markup Wizard (Word RTF formázással strukturált szöveg + formális nyelvtani leírás -> XML és DTD), Lexicographer's Workbench v1: továbbfejlesztés, v2: (web-alapú) tervezés és konzultáció

Amire büszke vagyok (és több közöm volt hozzá)

- Elsőségek, és magyarországi piacvezető szerep
- „Garázscégből” nemzetközileg ismert „műhely” alakult ki
- A „műhely” sok szakembert indított el a nyelvi és szoftverpályán
- Sikeres, hosszú ideig piacon lévő termékbeépítések (pl. Microsoft, Adobe)
- Sok szótár dokumentumból (doc) strukturált, számítógéppel jobban kezelhető (xml) formátumra való átalakítása (MarkupWizard)
- A termékek stabilitása (a különböző termékekre különböző mértékben volt hatásom)

Amire büszke vagyok (de kevesebb közöm volt hozzá)

- Egy kutatás kidobott eredményeinek megmentése pusztuló adathordozókról, ami alapján gépi szóalaktani rendszer alapjainak lerakása (Prószéky Gábor; Papp Ferenc: A magyar nyelv szóvégmutató szótára)
- A MoBiMouse sikere (Földes András)
- A Fordító program újszerű adatábrázolásának és működésének alapötlete, a MetaMorpho (Tihanyi László)
- Fordító motor (Ugray Gábor; az alapok lerakásánál segítettem)
- Web-alapú Lexicographer's Workbench (Rácz András; a tervezésében vettem részt)
- A szóalaktan újszerű feldolgozása (Novák Attila)
- A Lázár-varga szótárak „élővé” tétele (Merényi Csaba...)

Amire még szerettem volna büszke lenni

- Piacon maradás a nyelvi eszközökkel, elsősorban a beépítésekkel
- A MoBiMouse piacon maradása a Windows Vistas nehézségek után
- A MoBiMouse nemzetközi elterjedése
- A MoBiMouse bevonása szótárkiadókkal történő együttműködésekbe (volt rá esély)
- Korrekt, pénzért eladható, az ingyenes, internetes eszközökét jelentősen meghaladó fordítási minőség elérése a fordító programmal
- A MetaMorpho alapelv technológiává válása, nemzetközi elterjedése és elismerése
- A fordítómemória projekt cégen belül maradása (szép munka volt, Kilgray!)
- A szoftvertrendek követése, pl. mobilalkalmazásokban való megjelenés
- Technológiai céggé válás: elsősorban beépítőkészletek és (web-)szolgáltatások értékesítése
- A csapat megtartásához megfelelő üzleti működés

A SzTAKI-tól A MorphoLogicig

Naszódi Máttyás



- A nyelvi csoport alakulása
- Mondatelemző
- Morfológiai rendszer
- Eredmények, tanulságok
- A csoport feloszlása
- Csatlakozás a MorphoLogichoz
- És mi a hozadék...

a
b
c
d
e
f
g
h
i
j
k
l
m

A nyelvi csoport alakulása



- Programnyelvek elmélete, elemzése, implementálása 1980-ig
- 1970. Fortran
- 1973. Lisp
- 1976. Basic
- 1980-ig készülés az ADÁ-ra
- Afgán válság

a
b
c
d
e
f
g
h
i
j
k
l
m

Váltás természetes nyelvekre



- A programnyelveken kialakultak a monopóliumok
- Az OMFB beindította az 5. generációs programot
- Nagy kedvünk volt hozzá
- A kis csoport tagjai:
 - Bach Iván
 - Farkas Ernő
 - Naszódi Mátyás

a
b
c
d
e
f
g
h
i
j
k
l
m

Bach Iván

- 1927-2006



a
b
c
d
e
f
g
h
i
j
k
l
m

mo

Farkas Ernő



- 1944-1994



a
b
c
d
e
f
g
h
i
j
k
l
m



Mondatelemző



- 1985-1987. Előtanulmányok, tervek készítése.
 - Magyar mondatok gépi elemzése
 - Felhasználás természetes nyelvű interfész céljából
- 1988-1990. Elkészül az általános célú mondatelemző.
 - Szótani elemző - véges automata
 - Vonzatokra építő szabad szórendű - attribútumgrammatika

Morfológiai rendszer



- Igény: szükség van nagy tömegű adatra az adatbázishoz
- Kellő sebesség szpellerekhez
- Lehetőség: Elekfi László ragozási szótára
- Véges automata alapján működő helyesírás-ellenőrző
- Kísérleti morfológiai elemző, generátor

Eredmények, tanulságok



- Lehet elemezni a magyart is!
 - A jó szóelemzés fél győzelem
 - Vonzatok felhasználásával lehet használható generatív modellt építeni
 - Nem gond a szabad szórend, legfeljebb nem a szokványos eszközöket kell használni

mo

a
b
c
d
e
f
g
h
i
j
k
l
m

Eredmények, tanulságok



- Számos publikáció
- Megerősítettük a nyelvészek konstruktív szárnyát
- Magyarországon először oktatni kezdtünk nyelvtechnológiát
- Az ígért támogatást soha sem kapjuk meg



A csoport feloszlása



- Elfogyott a muníció
- A SzTAKI nem tartotta elég súlyosnak a projektet
- Nem volt olyan akadémiai cég, amely átvállalt volna
- Nem találtunk ipari támogatóra
- Leépítés volt a SzTAKI-ban

a
b
c
d
e
f
g
h
i
j
k
l
m

Csatlakozás a MorphoLogichoz



- Miért?
 - Nekünk nem volt hol folytatni
 - Egy éve már prosperált a MorphoLogic
 - Ismertük egymást a szakmából
 - A MorphoLogic tovább akart lépni
- Bach Iván az egyetemre ment át nyugdíjasként
- Ernő egy év múltán távozott.
 - Az Elekfi rendszert átvittük



És mi a hozadék...

- Közvetlen eredmény:
 - Javult a Humor rendszer - főleg Ernőnek köszönhetően
 - Elkészült az első magyar mondatellenőrző
- Közvetett eredmény:
 - A tapasztalat nem vész el
 - A tapasztalatok beépültek későbbi termékekbe
 - Számos diákom vált a szakma tagjává - néhányan a MorphoLogicon keresztül



**Szünet
után
folytatjuk**



a
b
c
d
e
f
g
h
i
j
k
l
m

**Szünet
után
folytatjuk**



mo

a
b
c
d
f
h
i
k
l
m

Prószéky Gábor

**A MorphoLogic további
sikeres termékei és a
MorphoLogic Lokalizáció
létrejötté**

Hócipő,
1995. január

vassunk
uk a Jú-

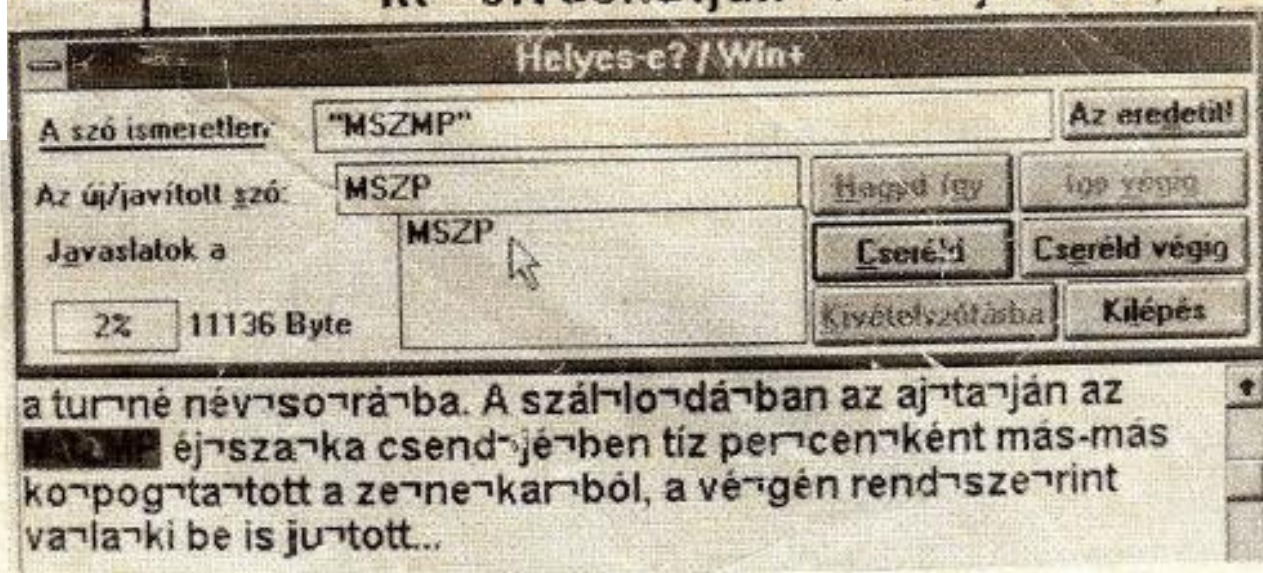
rendben
t: Ham-
lia, Ro-
Gertrud,

LY
albérel
arba ke-
larab.

kollégá-
gáról hí-
szerelmé-
ce) a po-

FILELAND

Itt olvashatjuk a folytatást.



A SZOCIK MÁR A TÁRBAN VANNAK

Iszonyatos tendenciákra világít rá a Titánsoft helyesírás-ellenőrző programja, a Helyes-e? 2.0, melynek segítségével azt ellenőrizhetjük minden Windows alatt futó szöve-
günkben: hejesen-e irrtauk-e a mondat?

A Helyes-e? 2.0 roppant udvarias. Az általa helytelennek, hibásnak, azaz ismeret-
lennek ítélt szavakat kiírja az ablak egyik sorába, az alatta lévő sorban pedig az általa
javasolt szót tünteti fel. A mi dolgunk csak annyi, hogy egyetlen kattintással utasítást
adjunk, belenyugszunk-e a megváltoztathatóba vagy sem.

Nos – és itt az iszonyat – ez a hasznos kis program nem ismeri az MSZMP-t. Helyette
folyamatosan az MSZP-t javasolja. Lehet, hogy nemsokára az MDF, a Fidesz, az
SZDSZ, az FKGP, a KDNP helyett is az MSZP-t fogja javasolni?

Esetleg az is elképzelhető, hogy még májusban ez a szoftver befurakodott a BM adat-
bázisába, s emiatt győzött a rózsaszín gőzhenger?

SOFTWARE COMPANIES

Ranked by 1993 revenue

Rank	Company Address	Revenue 1993 Revenue 1st half 1994	Number of employees Year founded in Hungary	Main types of business			Types of software sold	Brands of software sold	Services	Major clients	Non-Hungarian ownership (%)		Top local executive Title Phone, Fax
				trading	consulting	development					Hungarian ownership (%)	Foreign ownership (%)	
9	Morphologic 1011 Budapest Ró u. 56-58.	15 8	6 1991	X		X	Linguistic software				Individuals, 100	Peter Dominis General Manager 201-8355,201-8355	
10	Graphisoft 1145 Budapest Kolumbusz u. 29.	6 4.5	100 1982			X	high-end software for AEC/CAD market on Macintosh, PowerPC and Windows platforms	Autodesk, TopCAD	Autodesk	architectural firms	Associates, 31.1 Graphisoft Holding Kft, 60.1, Employees, 8.8	President 251-1000,251-1890	
NR	Ablaksoft Kft 1051 Budapest Oktober 6. u. 7.	WND	32 1991			X	Business planning, accounting	Ablaksoft	Software engineering, localization, system integration	Microsoft, Beszámolótervező AB	Scala ECE, 100	Michael Simon Managing Director 266-5777,266-5703	
NR	Scala ECE Ltd 1051 Budapest Oktober 6. u. 7.	WND	1994	X		X	International Financial Management, Accounting	Scala	Implementation, demand evaluation, consulting, training	ABB, Burger King, Resti, Astris, Alcatel, Pannon GSM	Scala Sweden, 100	Andreas Kerni Managing Director 266-5777,266-5703	
NR	Coda Ltd 1051 Budapest Révay u. 10. (ESZ)	NA	3 1993	X		X	Financial, accounting	Coda IAS, OAS	Implementation, training, hot-line	Westel, Herkules Rt, AmWay	Coda Ltd, 100	Richard Wilms Managing Director 269-1114,269-1030	
NR	Microsoft Hungary 1075 Budapest Madách I. u. 13-14.	NA 600*	13 1992			X	Business and personal, word processor, database management, graphic	Microsoft	Hot-line, consulting	Prime Minister's Office, Matáv, OTP, Budapest Bank, MOL, Paks Power Plant, Medical University	Microsoft, 100	Andrić Gábor Managing Director 268-1668,268-1558	

Notes: WND=Would Not Disclose NA=Not Available NR=Not Ranked

- Footnotes:
1. Revenue figures are only estimated.
 2. Coda Ltd also reported Tibor Szente as managing director.
 3. Figure represents the period of July 1993 - June 1994.

The information in The List was provided by the companies in response to a survey completed during the week of November 22, 1994. Only those responding to the questionnaire were included. To the best of the Busi-

ness Journal's knowledge, the information is accurate as of press time. While every effort is made to ensure accuracy and thoroughness, omissions and typographical errors may occur. Corrections or additions to The List should be sent on company letterhead to Beatrix Balogh, Research Coordinator, Budapest Business Journal, Ferenciek tere 7-8, I em. 4, 1053 Budapest; or faxed to 116-0215 © 1994 Budapest Business Journal. The List may not be reprinted or reproduced in whole or in part without prior written permission from the Editor. Reprints are available. To order, contact Beatrix Balogh, Research Coordinator.

— Researched and compiled by Beatrix Balogh

Bálna a szavak tengerén

Az utolérhetetlen szótáróriás

Hogy kerül a csizma az asztalra? Azaz, kerül egy számítógépes nyelvészettel foglalkozó cég a Természet magazinba? Nos, túl hamar remélhetőleg minden nyelvész és nyelvész-számítógépes szereti a természetet, egy eszközt szeretnénk itt bemutatni, mely segítségével olyan szótárak készíthetők, amelyek eddig csak a professzionális szoft-

(A természetes nyelvek számítógépes feldolgozásával foglalkozó MorphoLogic cég munkatársai nagy fogékonyságot mutatnak a szójátékok iránt.

Megmutatkozik ez névadási gyakorlatukban is: legelterjedtebb programjuk a Helyes-e? névre hallgat, kétnyelvű szótárunk nevét pedig a híres fehér bálnára rimelgetve választották MoBIDIC-nek (= MorphoLogic Bilingual Dictionary).

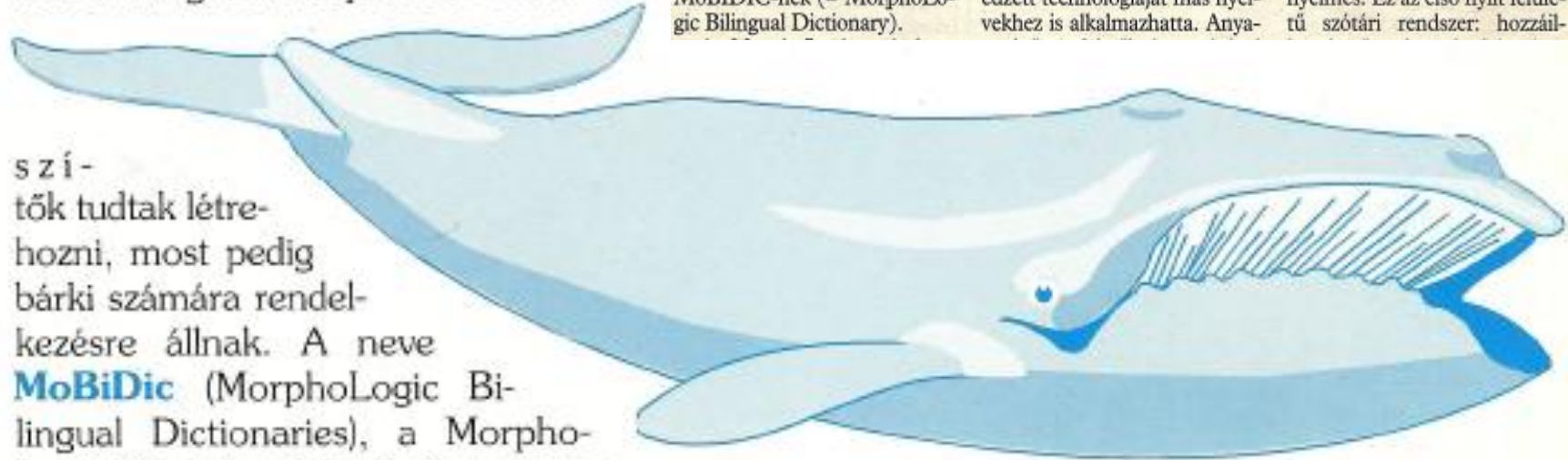
Microsoft-programok mind alkalmazzák. Újdonság azonban a macintoshos QuarkXpress-hez, a PageMaker 6.0-hoz és a Lotus WordPro 96-hoz készített változat.

Mivel a magyar nyelvben lényegesen több szóalak képezhető, mint más, elsősorban indoeurópai nyelvben, a MorphoLogic a magyar nyelven edzett technológiáját más nyelvekhez is alkalmazhatta. Anya-

ram nem javít, csak roppant udvariasan figyelmeztet a lehetséges hibákra – a döntés és a javítás a felhasználó dolga.

A MorphoLogic Ifabó-programjának középpontjában a MoBIDIC, a szótárprogram áll. A két irányban használható, a toldalékolt szavakat is gond nélkül megtaláló program felhasználói feltétele rendkívül kényelmes. Ez az első nyílt felületű szótári rendszer: hozzáil-

szí-
tők tudtak létrehozni, most pedig bárki számára rendelkezésre állnak. A neve **MoBiDic** (MorphoLogic Bilingual Dictionaries), a MorphoLogic Kft. kétnyelvű szótárprogram-családja. A Windows alatt futó program akár száz szótár egyidejű megjelenítésére szolgál. Akár vásároljuk a hozzá való szótárakat, akár magunk készítjük őket, az okos eszköz egyszerre keres ezekben, lehetővé téve azt, ami oly so-



Mindebben segít a MorphoLogic: aki ennek a felhívásnak a szövegével eljön 1996. február 2-án a MorphoLogic irodájába (1126 Németvölgyi út 25.), az ingyen kaphat egy **MoBiDic** alapprogramot (kereskedelmi

Ismét előtérbe lép a gépi fordítás

Géppel támogatott fordítás

Magyarországon a 60-a fordítással foglalkozó akadémikusok — a világban — a megfelelően — felszámolták — együtt más irányú számításokkal — viszonylag kis támogatással — kutatóhely nem feltétlenül igényével végzett, sokszor csak kisebb volumenű kutatásai

Műhelyvitába csöppentem a MorphoLogicnál; a résztvevők (Prószéky Gábor, Pál Miklós és névrokonom, Tihanyi László) egy távol-keleti eredetű, csapnivaló minőségű „25 nyelvű fordító” CD-t ostromoltak. Ilyenek is megjelennek, mert a

rülethez. Ezután EuroLangot, a céljával, a piac integrálása jegyében éledt újjá a Met Az EuroLang

MorphoLogic, avagy a titkárnők álma

Régebben ha egy titkárnő felvételre jelentkezett, az első kérdés a helyesírására vonatkozott. A számítógép korában elég, ha azt mondja, ért a komputerhez. Ha azonban azt a mondatot, hogy „A víz csendesen fojt tovább.” pontos jével írja, még most is rosszalló pillantást vet rá a főnöke. Illetve csak velett egészen mostanáig. A MorphoLogic cég ugyanis kidolgozott egy olyan nyelvi programot, amely többek között nemcsak különálló szavak értelmezésére és helyesírási ellenőrzésére képes, hanem szövegek környezetében, az egész mondat jelentését vizsgálva korrigálja a hibát.

– Rájöttünk, hogy képesek vagyunk a különböző nyelveket tisztességesen leírni a számítógépek számára. Célunk alapvetően nem egy termék, vagy eszköz előállítása volt, hanem a számítógéppel való nyelvkutatás – mondja Prószéky Gábor, az 1991-ben alakult cég vezetője.

tárunk is, amelynek érdekessége, hogy a kért szó helyére felajánlott szinonimákat nem adott tári alakjukban, hanem az adott mondatba illesztve, megfelelően ragozva adja meg. Az új modulokkal lehetővé tesszük eddig lehetetlennek tűnő dolgokat, például a beszédfelismerést is.

MorphoLogic

Új székházban

Helyesebben a MorphoLogic-al

Új központjában tart előtti sajtótájékoztatót phoLogic Kft. Amin székly Gábor ügyvezet

Gépi fordítás?

Prószéky Gábor (40) kandidátus, programtervező matematikus, nyelvész, a mindössze tízfős MorphoLogic Kft. alapító főnöke. Ő és munkatársai a nyelvészet eredményeit ültetik be a számítástechnikai eszközökbe, hogy szoftverek

MorphoLogic. Ez a név ma egy átlagos számítógépes

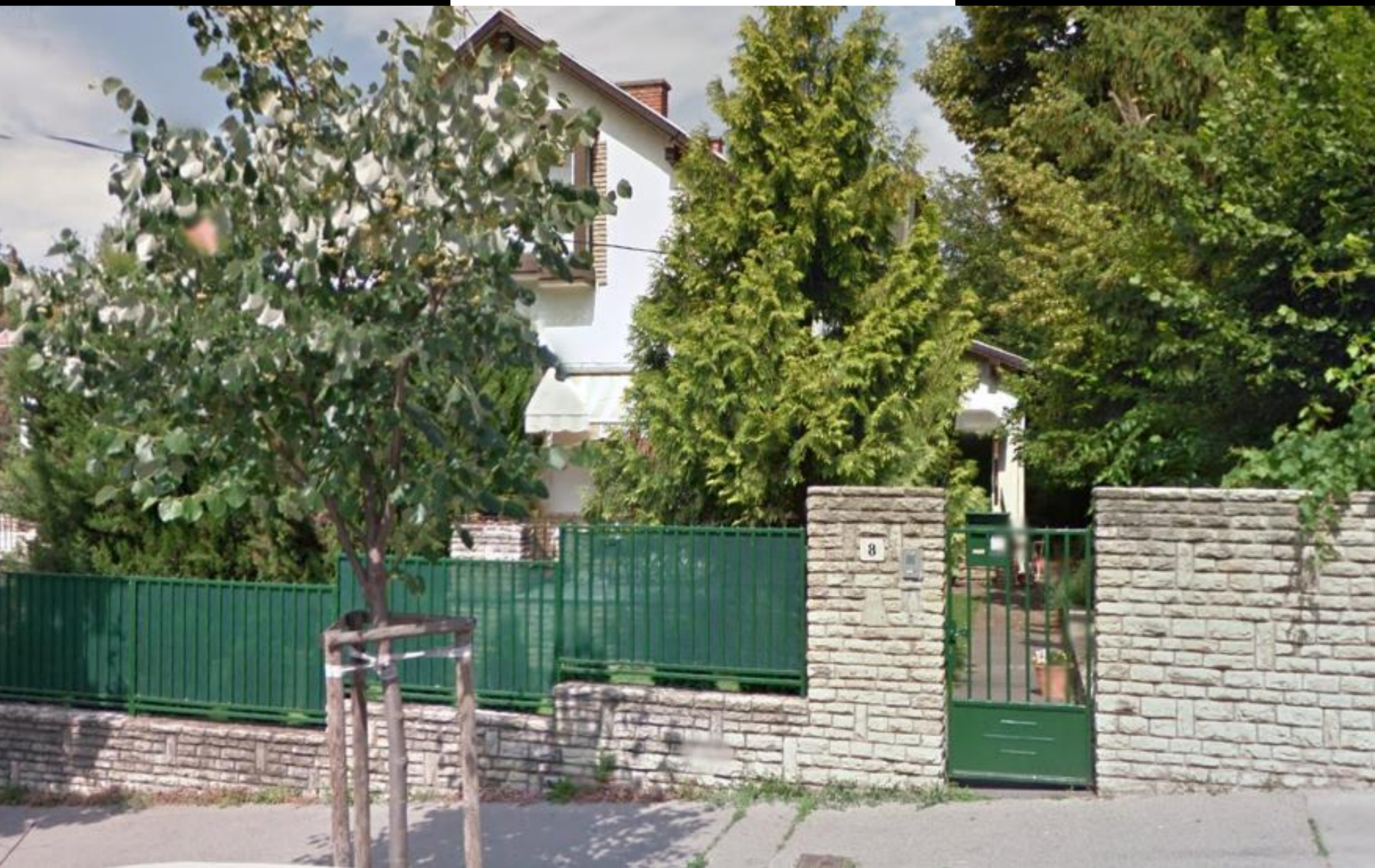
bet, de egy játékokat futtató, iskolai tanulmányait végző embernek szinte semmit sem. E cikk célja, hogy egy kicsit közelebb vigyük ehhez a réteghez azokat a termékeket, amiket valóban jól használ-

területen érdeme egyikünknek sem nyelvtanóra két

Nyelvészek és matematikusok informatikai üzlete

Bill Gates is a MorphoLogictól vásárol nyelvi programot

Késmárki utca, 1999



MoBiMouse, 1999



MoBiMouse



Világszenzáció!

A MoBiMouse, a fordító egér
gombnyomás nélkül fordít!

- o Minden Windows alkalmazáshoz
- o MoBiMouse for Acrobat
Acrobathoz Acrobat Readerhez

Eloivassa a képernyőn lévő szöveget,
sőt le is fordítja!



1467

6

MoBi Mouse
Gombnyomás nélkül fordít

MoBi Mouse
Gombnyomás nélkül fordít





**Megszületik a MorphoLogic Lokalizáció
2001 végén**

Éke(zete)s anyanyelvünk

MIKROVILÁG: Beszélgetésünk résztvevői három különböző korosztályt képviselnek. A közös vonás: a számítógépes nyelvészet szeretete. Honnan ered ez az érdeklődés?

PAPP F.: Közel harminc éve, első generációs — tehát elektroncsöves — gépen kezdtem el a munkát. Nagyon lelkesek voltunk. Emlékszem, hogy Kalmár László is előadást tartott számítógépes nyelvészetről, sőt még szociológusok is érdeklődtek a téma iránt.

PRÖSZÉKY G.: 1979-ben, egyetemistaként csöppentem bele a számítógépes nyelvészet izgalmas világába. Akkoriban alkalmazták először a PROLOG-ot, s az MTA Nyelvtudományi Intézet és a SZTA-KI közös vállalkozásaként magyar morfológiai elemző rendszert dolgoztunk ki.

BESSENYEI G.: Szerencsés csillagzat alatt születtem, hiszen gimnazistaként én már személyi számítógéppel dolgozhatok. Olyan programokat írok, amelyek a nyelvtan tanítását segíthetik.

Szavak, szavak,
szavak

MIKROVILÁG: Meglepő, de tény, hogy a nyelvészek elsők között kezdtek el foglalkozni a



DR. PAPP FERENC
akadémikus
**ELTE Általános
és Alkalmazott
Nyelvészeti Tanszék**
*„Örömmel javítottam a hi-
bákat”*



PAJZS JÚLIA
MTA Nyelvtudományi
Intézet
*„A számítógépes
mondatgenerálás
rendkívül izgalmas”*



PRÖSZÉKY GÁBOR
Országos Pedagógiai
Könyvtár és Múzeum
*„Az UNESCO is támogatja
a nemzeti törekvéseket”*



BESSENYEI GÁBOR
Berzsenyi Dániel
Gimnázium
*„Szerencsés csillagzat
alatt születtem”*

val szemléltetném. Előfordulhat, hogy azt a szót: *csónak*, a számítógép a „csón” szó többes számaként értelmezi. Az sem lehetetlen, hogy egy ige többes szám harmadik személyű alakjának tekinti, mint például a „futnak” esetében. De az is lehet, hogy egy főnév birtokos személynévragjának értelmezi. Ami egy iskolás gyerek számára is természetes, az komoly nehézséget jelenthet a számítógépnek.

PAJZS J.: Egy másik izgalmas terület a számítógépes mondatgenerálás. A gép képes arra, hogy nyelvtanilag tökéletes mondatokat hozzon létre véletlenszerűen a szótárából. Ilyeneket például: a kutyája jól ugat a szobában este. Ez azonban csak formális szerkezet, a mondatnak nincs „lelke”, gondolati tartalma. Meg tudjuk határozni, hogyan épüljön fel (szintaxis), de azt nem, hogy mit mondjon (szemantika).

PRÖSZÉKY G.: Az ember — beleérző képessége révén — képes arra, hogy nyelvtanilag nem tökéletes mondat alapján is megértse a beszélő szándékát. A gép ilyen szempontból tökéletlenebb. A számítógépes nyelvészek nagy álma, hogy olyan rendszert fejlesszenek ki, amellyel társaloghatunk. Ez nagyon hasznos

MIKROVILÁG

A VSZM Közösségi Házban
a számítógépes nyelvészettel foglalkozó
szakembereket kérdeztük kutatási
eredményeikről.

1987. június

MorphoLogic Localisation

- Budapest és Köln (testvérvállalat)
- Üzletágak
 - Szoftverlokalizáció
 - Tanácsadás (SAP-ügyfelek globális roll-out projektjeinek támogatása)
 - Nyelvtechnológia és fejlesztés (gépi fordítás, SAP)



Globalese

Machine Translation System

Powered by MorphoLogic Localisation

Globalese: áttekintés



- Böngészőből elérhető statisztikai és neurális gépi fordítási rendszer
- Egyéni gépi fordítói motorok (statisztikai és neurális) készítésére alkalmas
- Könnyen integrálható a fordítási munkafolyamatba
- Automatikus „in-domain adaptation” funkció (kis korpuszból is készíthető jó motor)

Some of the companies using Globalese



METAMORPHO

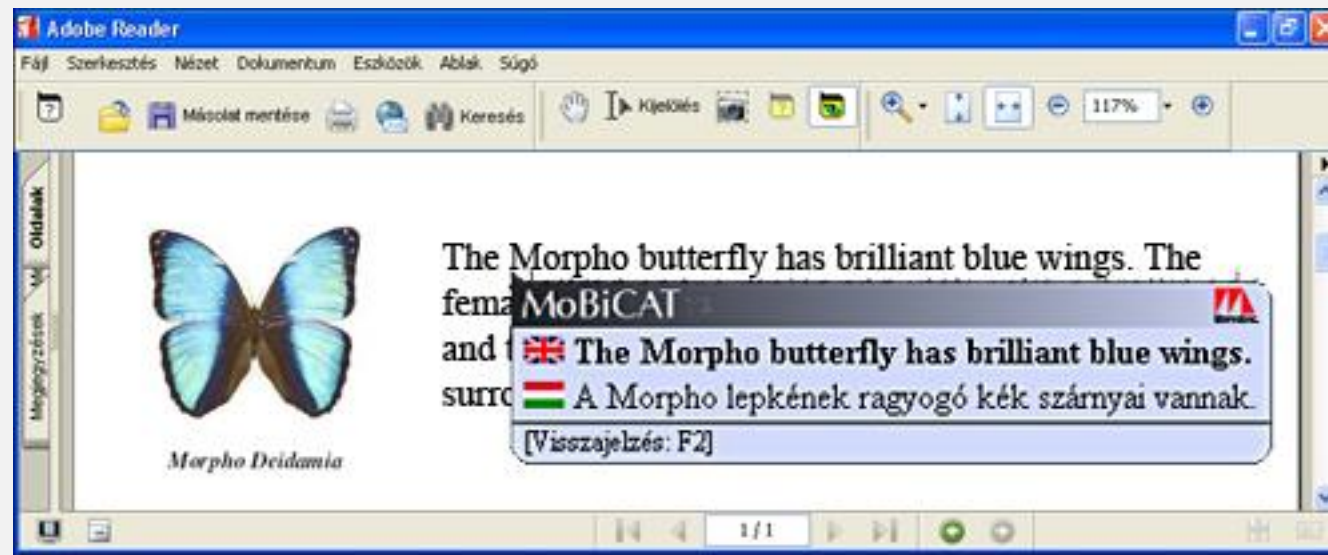
A MorphoLogic gépi fordítórendszere

MI A METAMORPHO?

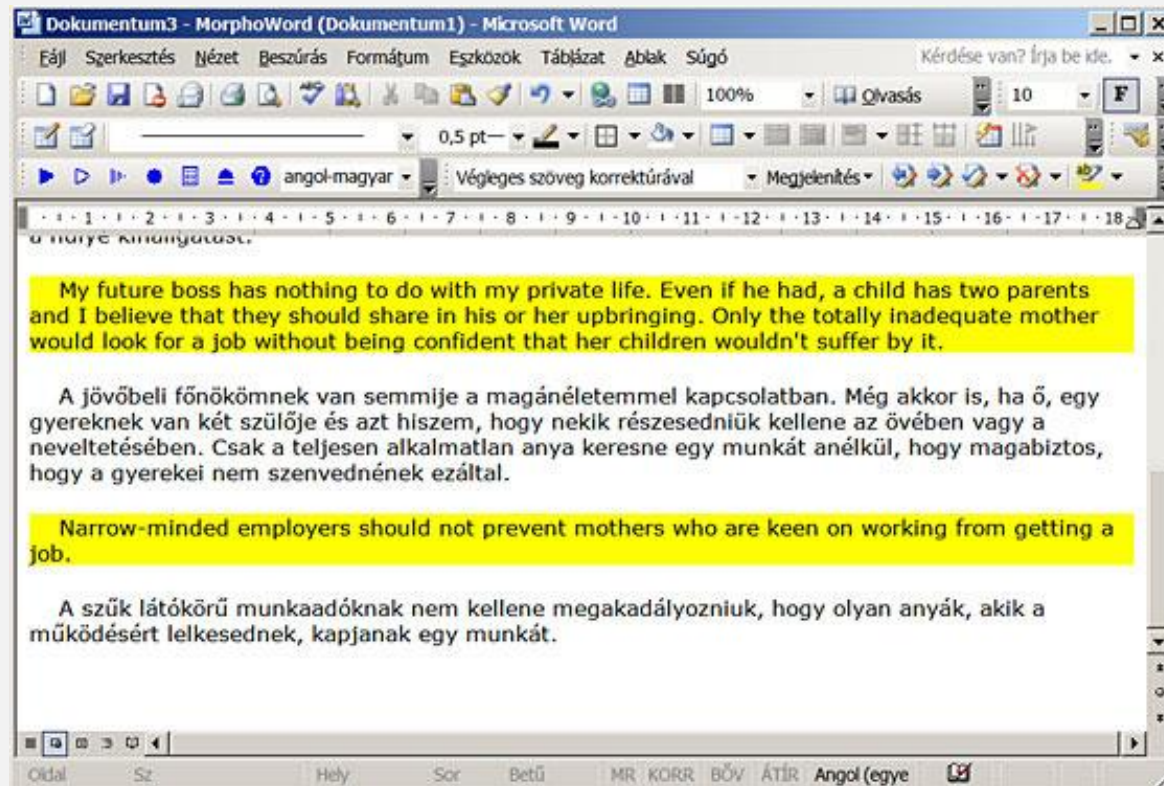
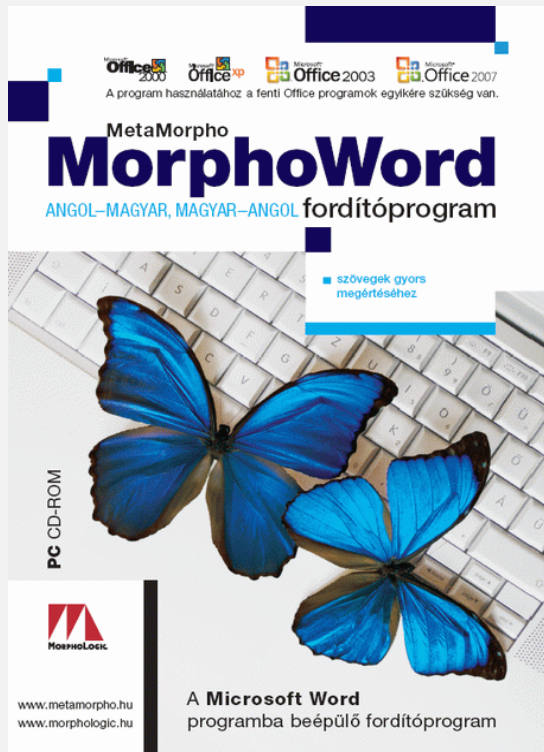
- Szabályalapú gépi fordítórendszer
- Angol-magyar 2001-2005
- Magyar-angol 2005-2008

METAMORPHO TERMÉKEK

MOBICAT



MORPHOWORD



WEBFORDITAS.HU



A MorphoLogic ingyenes nyelvészeti szolgáltatásokat nyújt: szövegfordítás, weblapfordítás, szótárak, hangos beszéd, nyelvfelismerés, helyesírás, többnyelvű keresés és mondatelemzés. Fordítóprogram a magyar és 12 másik nyelv között!

weblapfordító szövegfordító szótár kereső elemző helyesírás

[English version](#) | [Belépés](#)

fordít angol magyar javaslat visszajelzés [hirdetés]

Can you translate this sentence to Hungarian? Le tudod fordítani ezt a mondatot magyarra?

felolvas [F] nyelvfelismerés helyesírás felolvas [F] angol-magyar © MorphoLogic

45/500 karakter

angol francia magyar német orosz spanyol

[angol](#), [bolgár](#), [dán](#), [francia](#), [lengyel](#), [lett](#), [magyar](#), [német](#), [norvég](#), [orosz](#), [portugál](#), [spanyol](#), [ukrán](#)

[beépítés](#) | [elérhetőség](#) | [közreműködők](#) | [hírek](#) | [médiaajánlat](#) | [GYIK](#)

A METAMORPHO PROJEKT SZÁMOKBAN

A KÉSZÍTŐK

Projektvezető
Tihanyi László

14 fő

Szoftver

- Ugray Gábor
- Gröbler Tamás
- Hegedűs Balázs
- Kundráth Péter
- Kiss Balázs
- Vajda Kristóf

Nyelvtan

- Ugray Gábor
- Vlaskovits Dóra
- Újvárosi Gábor
- Merényi Csaba

Morfológia

- Novák Attila
- Kiss Gabriella

QA

- Vancsa László
- Szécsi Katalin

PARTNEREINK

2 külső intézményi partner

- **NKFP projekt - Magyar-angol lexikális erőforrások**
- **Szegedi Tudományegyetem Informatikai Tanszékcsoport**
- **Magyar Tudományos Akadémia Nyelvtudományi Intézet**

FEJLESZTÉSI IDŐ

~ 8 év

Angol-magyar: 2001-2005

Magyar-angol: 2005-2008

A NYELVTANOK MÉRETEI

ANGOL-MAGYAR

- 277.416 (mmo) – 149.984 (mmd)
- Core: 4.289 (mmd)
- Lexicon: 114.868 (mmd)
- VP: 23.318 (mmd) + VV: 1.158 (mmd)

- 186 class
- 785 property

MAGYAR-ANGOL

- 193254 (mmo) – 181509 (mmd)
- Core: 5741 (mmd)
- Lexicon: 132157 (mmd)
- VP: 34309 (mmd)

- 196 class
- 3364 property

A FORRÁSKÓD

- A program legnagyobb része C++ nyelven íródott.
- Több mint kétezer saját forrásfájlból és 250 projektből áll.

A TECHNOLÓGIA

Az MMD és MMO nyelvek
Fejlesztőeszközök

MMD ÉS MMO

*NP=DET+NX:12345678-1

HU.NP[det=YES, def=DET.def] = DET(art=YES) + NX(allowdet=YES)

EN.NP(HU.DET.def=YES) = DET[lex="the"] + NX

EN.NP(HU.DET.def=NO) = DET[lex="a"] + NX

!12345678-2, 12345678-3

*VP=elő|áll:200607120956

HU.VP = SUBJ(human=YES) + TV(:lex="elő|áll") + COMPL#1(pos=N, case=INS)

EN.VP = SUBJ + TV[lex="come"] + PART[lex="up"] + COMPL#1[prep="with"]

VALÓDI MMO ELRETTENTÉSKÉNT

*VP=abba|hagy:101-3-0-12-0000-00

HU.VP[V]PQ.Cauclink, VPQ.Cauplink, VPQ.Caurlink, VPQ.Causer, VPQ.Clink1, VPQ.Clink2, VPQ.Clink3, VPQ.Clink4, VPQ.Clink5, VPQ.Clink6, VPQ.Const1, VPQ.Const2, VPQ.Const3, VPQ.Const4, VPQ.Const5, VPQ.Const6, Encauser=VPQ.Causer, Encauser_clink=VPQ.Cauclink, Encauser_plink=VPQ.Cauplink, Encauser_rlink=VPQ.Caurlink, VPQ.Plink1, VPQ.Plink2, VPQ.Plink3, VPQ.Plink4, VPQ.Plink5, VPQ.Plink6, VPQ.Rlink1, VPQ.Rlink2, VPQ.Rlink3, VPQ.Rlink4, VPQ.Rlink5, VPQ.Rlink6, VPQ.Vauy, VPQ.Vy, adv_goal=NO, adv_source=NO, cau_clinktarget=VPQ.cauclinktarget, cau_ht=VPQ.cauht, cau_num=VPQ.caunum, cau_pers=VPQ.caupers, cau_ref=VPQ.cauref, cau_rel=VPQ.caurel, cau_rlinkrel_conj=VPQ.caurlinkrel_conj, cau_rlinkrel_msg=VPQ.caurlinkrel_msg, cau_rlinktarget=VPQ.caurlinktarget, cau_rlinktype=VPQ.caurlinktype, cau_xrl_msg=VPQ.cauxrl_msg, VPQ.cauclinkf, caup_ref=VPQ.cauplinkref, caup_rel=VPQ.cauplinkrel, VPQ.cauplinkf, VPQ.caurlinkf, VPQ.causerf, VPQ.clink1f, VPQ.clink1target, VPQ.clink2f, VPQ.clink2target, VPQ.clink3f, VPQ.clink3target, VPQ.clink4f, VPQ.clink4target, VPQ.clink5f, VPQ.clink5target, VPQ.clink6f, VPQ.clink6target, VPQ.cond, VPQ.const1clause, VPQ.const1f, VPQ.const1focus, VPQ.const1ref, VPQ.const1rel, VPQ.const1subargf, VPQ.const1uid, VPQ.const1video, VPQ.const1wh, VPQ.const1xrl_msg, VPQ.const2clause, VPQ.const2f, VPQ.const2focus, VPQ.const2ref, VPQ.const2rel, VPQ.const2subargf, VPQ.const2uid, VPQ.const2video, VPQ.const2wh, VPQ.const2xrl_msg, VPQ.const3clause, VPQ.const3f, VPQ.const3focus, VPQ.const3ref, VPQ.const3rel, VPQ.const3subargf, VPQ.const3uid, VPQ.const3video, VPQ.const3wh, VPQ.const3xrl_msg, VPQ.const4clause, VPQ.const4f, VPQ.const4focus, VPQ.const4ref, VPQ.const4rel, VPQ.const4subargf, VPQ.const4uid, VPQ.const4video, VPQ.const4wh, VPQ.const4xrl_msg, VPQ.const5clause, VPQ.const5f, VPQ.const5focus, VPQ.const5ref, VPQ.const5rel, VPQ.const5subargf, VPQ.const5uid, VPQ.const5video, VPQ.const5wh, VPQ.const5xrl_msg, VPQ.const6clause, VPQ.const6f, VPQ.const6focus, VPQ.const6ref, VPQ.const6rel, VPQ.const6subargf, VPQ.const6uid, VPQ.const6video, VPQ.const6wh, VPQ.const6xrl_msg, VPQ.constnum, VPQ.e, encauser_clink=VPQ.cauclinkf, encauser_clinktarget=VPQ.cauclinktarget, encauser_huarg=CAU, encauser_plink=VPQ.cauplinkf, encauser_rel=VPQ.caurel, encauser_rlinkf=VPQ.caurlinkf, encauser_rlinkrel_conj=VPQ.caurlinkrel_conj, encauser_rlinkrel_msg=VPQ.caurlinkrel_msg, encauser_rlinktarget=VPQ.caurlinktarget, encauser_rlinktype=VPQ.caurlinktype, encauser_wh=VPQ.cauwh, encauser_xrl_msg=VPQ.cauxrl_msg, encauserf=VPQ.causerf, enpstrgt_genmode=NORMAL, enpstrgt_huarg=OBJ, ensubj_genmode=NORMAL, ensubj_huarg=SUBJ, envpntype=TRANS, VPQ.focus, VPQ.frontdet, VPQ.gerund, VPQ.hat, VPQ.headlast, VPQ.ik, VPQ.is_arg, VPQ.is_v, VPQ.lexb, mayfact=VPQ.muv, maypass=YES, VPQ.neg, VPQ.negagr, VPQ.negp, VPQ.num, VPQ.ocase, VPQ.opers, VPQ.ownernum, VPQ.ownerpers, VPQ.participle, VPQ.pass, VPQ.pers, VPQ.plink1enum, VPQ.plink1f, VPQ.plink1ht, VPQ.plink1lex, VPQ.plink1num, VPQ.plink1pers, VPQ.plink1ref, VPQ.plink1rel, VPQ.plink1uid, VPQ.plink2enum, VPQ.plink2f, VPQ.plink2ht, VPQ.plink2lex, VPQ.plink2num, VPQ.plink2pers, VPQ.plink2ref, VPQ.plink2rel, VPQ.plink2uid, VPQ.plink3enum, VPQ.plink3f, VPQ.plink3ht, VPQ.plink3lex, VPQ.plink3num, VPQ.plink3pers, VPQ.plink3ref, VPQ.plink3rel, VPQ.plink3uid, VPQ.plink4enum, VPQ.plink4f, VPQ.plink4ht, VPQ.plink4lex, VPQ.plink4num, VPQ.plink4pers, VPQ.plink4ref, VPQ.plink4rel, VPQ.plink4uid, VPQ.plink5enum, VPQ.plink5f, VPQ.plink5ht, VPQ.plink5lex, VPQ.plink5num, VPQ.plink5pers, VPQ.plink5ref, VPQ.plink5rel, VPQ.plink5uid, VPQ.plink6enum, VPQ.plink6f, VPQ.plink6ht, VPQ.plink6lex, VPQ.plink6num, VPQ.plink6pers, VPQ.plink6ref, VPQ.plink6rel, VPQ.plink6uid, VPQ.possnum, VPQ.prereqcomma, VPQ.ref, VPQ.rel, VPQ.remodifier, VPQ.relprnht, VPQ.relprnlex, VPQ.relprnnum, VPQ.reqcomma, VPQ.rlink1f, VPQ.rlink1rel_conj, VPQ.rlink1rel_msg, VPQ.rlink1target, VPQ.rlink1type, VPQ.rlink2f, VPQ.rlink2rel_conj, VPQ.rlink2rel_msg, VPQ.rlink2target, VPQ.rlink2type, VPQ.rlink3f, VPQ.rlink3rel_conj, VPQ.rlink3rel_msg, VPQ.rlink3target, VPQ.rlink3type, VPQ.rlink4f, VPQ.rlink4rel_conj, VPQ.rlink4rel_msg, VPQ.rlink4target, VPQ.rlink4type, VPQ.rlink5f, VPQ.rlink5rel_conj, VPQ.rlink5rel_msg, VPQ.rlink5target, VPQ.rlink5type, VPQ.rlink6f, VPQ.rlink6rel_conj, VPQ.rlink6rel_msg, VPQ.rlink6target, VPQ.rlink6type, VPQ.scase, VPQ.spadv, VPQ.spostp, VPQ.tense, VPQ.trans, VPQ.vauf, vpid="101-3", vpsign="SVO", VPQ.vti, VPQ.vyf, VPQ.wh, xo=YES, xo_argtype=NP, xo_case=VPQ.ocase, xo_padv="", xo_postp="", xo_uid="0", xs=YES, xs_argtype=NP, xs_case=VPQ.scase, xs_padv="", xs_postp=VPQ.spostp, xs_uid="0", VPQ.zerocopula] = VPQ.(Vy->med=NO, Vy->gyak=NO, am_accf=AM_NO, am_nomf=AM_NO, ik="abba", lexb="hagy")

EN.VP = VPP#I [HU.VPQ.Cauclink, HU.VPQ.Cauplink, HU.VPQ.Caurlink, HU.VPQ.Causer, VP.Mod1, VP.Mod1clink, VP.Mod1ensubj, VP.Mod1plink, VP.Mod1rlink, VP.Mod2, VP.Mod2clink, VP.Mod2ensubj, VP.Mod2plink, VP.Mod2rlink, VP.Mod3, VP.Mod3clink, VP.Mod3ensubj, VP.Mod3plink, VP.Mod3rlink, VP.Mod4, VP.Mod4clink, VP.Mod4ensubj, VP.Mod4plink, VP.Mod4rlink, VP.Mod5, VP.Mod5clink, VP.Mod5ensubj, VP.Mod5plink, VP.Mod5rlink, VP.Obj, VP.Oblref, VP.Oblrefclink, VP.Oblreflink, VP.Oclink, VP.Oensubj, VP.Oplink, HU.VPQ.Relclink, VP.Sclink, VP.Sensubj, VP.Splink, VP.Srlink, VP.Subj, VP.Supoblref, VP.Supoblrefclink, VP.Supoblreflink, VP.Supoblrefrlink, VP.Vauy, HU.VPQ.cauclinkf, HU.VPQ.cauclinktarget, cauhuarg=CAU, HU.VPQ.cauplinkf, HU.VPQ.caurlinkf, HU.VPQ.caurlinkrel_conj, HU.VPQ.caurlinkrel_msg, HU.VPQ.caurlinktarget, HU.VPQ.caurlinktype, HU.VPQ.causerf, HU.VPQ.cauwh, HU.VPQ.cauxrl_msg, VP.cont, VP.dirpart, VP.embed, VP.hucond, VP.huhat, VP.hutense, VP.inv, lex="stop", VP.mod1adjarg, VP.mod1agnum, VP.mod1case, VP.mod1clausetype, VP.mod1clinkf, VP.mod1clinktarget, VP.mod1deg, VP.mod1det, VP.mod1ensubj, VP.mod1f, VP.mod1huarg, VP.mod1lexf, VP.mod1mod, VP.mod1mov, VP.mod1nosubj, VP.mod1num, VP.mod1plinkf, VP.mod1poo, VP.mod1pos, VP.mod1prep, VP.mod1prepr, VP.mod1prn, VP.mod1prontype, VP.mod1rel, VP.mod1rlink, VP.mod1rlinkrel_conj, VP.mod1rlinkrel_msg, VP.mod1rlinktype, VP.mod1shiva, VP.mod1shivaposs, VP.mod1tense, VP.mod1wh, VP.mod1xrl_msg, VP.mod2adjarg, VP.mod2agnum, VP.mod2case, VP.mod2clausetype, VP.mod2clinkf, VP.mod2clinktarget, VP.mod2deg, VP.mod2det, VP.mod2ensubj, VP.mod2f, VP.mod2huarg, VP.mod2lex, VP.mod2mod, VP.mod2nosubj, VP.mod2num, VP.mod2plinkf, VP.mod2poo, VP.mod2pos, VP.mod2prep, VP.mod2prepr, VP.mod2prn, VP.mod2prontype, VP.mod2rel, VP.mod2rlink, VP.mod2rlinkrel_conj, VP.mod2rlinkrel_msg, VP.mod2rlinktype, VP.mod2shiva, VP.mod2shivaposs, VP.mod2tense, VP.mod2wh, VP.mod2xrl_msg, VP.mod3adjarg, VP.mod3agnum, VP.mod3case, VP.mod3clausetype, VP.mod3clinkf, VP.mod3clinktarget, VP.mod3deg, VP.mod3det, VP.mod3ensubj, VP.mod3f, VP.mod3huarg, VP.mod3lex, VP.mod3lexf, VP.mod3mod, VP.mod3mov, VP.mod3nosubj, VP.mod3num, VP.mod3plinkf, VP.mod3poo, VP.mod3pos, VP.mod3prep, VP.mod3prepr, VP.mod3prn, VP.mod3prontype, VP.mod3rel, VP.mod3rlink, VP.mod3rlinkrel_conj, VP.mod3rlinkrel_msg, VP.mod3rlinktype, VP.mod3shiva, VP.mod3shivaposs, VP.mod3tense, VP.mod3wh, VP.mod3xrl_msg, VP.mod4adjarg, VP.mod4agnum, VP.mod4case, VP.mod4clausetype, VP.mod4clinkf, VP.mod4clinktarget, VP.mod4deg, VP.mod4det, VP.mod4ensubj, VP.mod4f, VP.mod4huarg, VP.mod4lex, VP.mod4lexf, VP.mod4mod, VP.mod4mov, VP.mod4nosubj, VP.mod4num, VP.mod4plinkf, VP.mod4poo, VP.mod4pos, VP.mod4prep, VP.mod4prepr, VP.mod4prn, VP.mod4prontype, VP.mod4rel, VP.mod4rlink, VP.mod4rlinkrel_conj, VP.mod4rlinkrel_msg, VP.mod4rlinktype, VP.mod4shiva, VP.mod4shivaposs, VP.mod4tense, VP.mod4wh, VP.mod4xrl_msg, VP.mod5adjarg, VP.mod5agnum, VP.mod5case, VP.mod5clausetype, VP.mod5clinkf, VP.mod5clinktarget, VP.mod5deg, VP.mod5det, VP.mod5ensubj, VP.mod5f, VP.mod5huarg, VP.mod5lex, VP.mod5lexf, VP.mod5mod, VP.mod5mov, VP.mod5nosubj, VP.mod5num, VP.mod5plinkf, VP.mod5poo, VP.mod5pos, VP.mod5prep, VP.mod5prepr, VP.mod5prn, VP.mod5prontype, VP.mod5rel, VP.mod5rlink, VP.mod5rlinkrel_conj, VP.mod5rlinkrel_msg, VP.mod5rlinktype, VP.mod5shiva, VP.mod5shivaposs, VP.mod5tense, VP.mod5wh, VP.mod5xrl_msg, VP.modal, VP.neg, VP.noargs, VP.nosubj, VP.num, VP.oadjarg, VP.oagnum, VP.obj, VP.oblrefarg, VP.oblrefclinkf, VP.oblrefclinktarget, VP.oblreff, VP.oblrefht, VP.oblrefmov, VP.oblrefnum, VP.oblrefpers, VP.oblrefplinkf, VP.oblrefrlinkf, VP.oblrefrlinkrel_conj, VP.oblrefrlinkrel_msg, VP.oblrefrlinktype, VP.oblrefrlinkrel_msg, VP.ocase, VP.oclausetype, VP.oclinkf, VP.oclinktarget, VP.odet, VP.odet, VP.oensubj, VP.oahuarg, VP.oalex, VP.oolexf, VP.omod, VP.omov, VP.onosubj, VP.onum, VP.oplinkf, VP.opoo, VP.opos, VP.oprep, VP.oprepstr, VP.oprn, VP.oprontype, VP.orel, VP.orlinkf, VP.orlinkrel_conj, VP.orlinkrel_msg, VP.orlinktarget, VP.orlinktype, VP.oshiva, VP.oshivaposs, VP.otense, VP.owh, VP.oxrl_msg, VP.pass, passtarget=OBJ, VP.passtype, VP.perf, VP.pers, HU.VPQ.rel, VP.rel_conj, VP.rel_msg, HU.VPQ.relclink_conj, HU.VPQ.relclink_msg, HU.VPQ.relclinkrel_conj, VP.relenum, VP.relht, VP.sadjarg, VP.sagnum, VP.scase, VP.sclausetype, VP.sclinkf, VP.sclinktarget, VP.sdeg, VP.sdet, VP.sensubj, VP.sshuarg, VP.slex, VP.slexf, VP.smod, VP.smov, VP.snosubj, VP.snum, VP.splinkf, VP.spool, VP.spos, VP.sprep, VP.sprepr, VP.sprn, VP.sprontype, VP.srel, VP.srlinkf, VP.srlinkrel_conj, VP.srlinkrel_msg, VP.srlinktarget, VP.srlinktype, VP.sshiva, VP.sshivaposs, VP.stense, VP.subj, VP.supoblrefarg, VP.supoblrefclinkf, VP.supoblrefclinktarget, VP.supoblreff, VP.supoblrefht, VP.supoblrefmov, VP.supoblrefnum, VP.supoblrefpers, VP.supoblrefplinkf, VP.supoblrefrlinkf, VP.supoblrefrlinkrel_conj, VP.supoblrefrlinkrel_msg, VP.supoblrefrlinktarget, VP.supoblrefrlinktype, VP.supoblrefxrl_msg, VP.swh, VP.sxrl_msg, VP.tense, VP.trans, VP.vauf, HU.VPQ.wh, VP.xmod1f, VP.xmod1ht, VP.xmod1num, VP.xmod1pers, VP.xmod2f, VP.xmod2ht, VP.xmod2num, VP.xmod2pers, VP.xmod3f, VP.xmod3ht, VP.xmod3num, VP.xmod3pers, VP.xmod4f, VP.xmod4ht, VP.xmod4num, VP.xmod4pers, VP.xmod5f, VP.xmod5ht, VP.xmod5num, VP.xmod5pers, VP.xof, VP.xoht, VP.xonum, VP.xopers, VP.xsf, VP.xst, VP.xsnum, VP.xsopers, VP.zerorel]

MOOSE

```
d:\MetaMorpho\Data>MmoMooseServer.exe MmoServerEnHuHuEn.cfg  
Hello, this is Moose [1.1.0.3] (Fri Apr 23 15:43:33 2010)
```

```
Initializing factories... OK  
Database type: SQLite3
```

```
Language pair: Hu-En  
Creating sessions... OK  
Grammar timestamp: Thu Nov 23 16:39:33 2017
```

```
Language pair: En-Hu  
Creating sessions... OK  
Grammar timestamp: Thu Nov 23 16:16:36 2017
```

```
Starting up the service thread OK  
The Moose is at your service after 9.161 sec
```

```
When lost, type !help
```

```
Moose [Hu-En]>!help
```

FORDÍTÁSI PÉLDA

```
Moose[Hu-En]>minden kutya ugat.  
1: [all dogs bark.]
```


ELEMZÉSI FA

```
Moose[Hu-En]>!trs
HU.S_ROOT 312
  S_FULL 309
    CP 268
      VPP 166
        CONST 63
          NPX 53
            Q 21
              ADJ lex="minden"
            NN 51
              N lex="kutya", num=SG, pers=P3, case=NOM
          VPP 162
            V ik="", lexb="ugat", num=SG, pers=P3
        PUNCT lex="period"
```

GENERÁLÁSI FA

```
EN.S_ROOT 314{312}
  S_FULL 315{309}
    CP 316{268}
      VPP 335{????}
        TV 336{????}
          VXT 339{????}
            SUBJ 340{????}
              DP 349{55}
                Q 350{21}
                  DET lex="all"
                NP 351{54}
                  N lex="dog"
              CONST 341{????}
                N lex=""
                V lex="bark", form=F1
            VPP 337{????}
              N lex=""
          PUNCT lex="period"
```

MAGYAR NYELVI SAJÁTOSSÁGOK

- Zéró elemek

Moose[Hu-En]>beletettem.

1: [I put it into it.]

- “Szabad szórend”

- Szétszakított összetevők

Moose[Hu-En]>azt a könyvet olvastam el, amit te ajánlottál.

1: [I read the book that you recommended.]

- Pólusharmónia

Moose[Hu-En]>Ezt a könyvet nem olvasta el senki sem.

1: [Nobody read this book.]

Prószéky Gábor

**A MorphoLogic pályázati
tevékenysége és
nemzetközi jelenléte**

**Bejelentjük a MorphoWordöt,
2005. október**



... jönnek is megnézni 😊 (2005. december)

Pályázati projektjeink

- 2008–2010 *Nyelv- és Beszédtechnológiai Platform (Nemzeti Technológiai Platformok)*
- 2005–2007 *Magyar WordNet: általános célú fogalomtár (GVOP)*
- 2005–2007 *Orvosi szakszövegek interaktív tartalomelemzése elektronikus kórlapok kitöltésére (GVOP)*
- 2005–2007 *A fordítási csoportmunka hatékonyságának maximalizálása számítógépes eszközökkel (GVOP)*
- 2005–2007 *Magyar egységes ontológia (NKFP)*
- 2005–2007 *Magyar–angol gépi fordítórendszer (NKFP)*
- 2004–2006 *A számítógépes nyelvészeti infrastruktúra egységesítése és fejlesztése (GVOP)*
- 2004–2006 *Kontrollált nyelvi szolgáltatás kifejlesztése (INFRA GVOP)*
- 2004–2006 *Digitális terminológus (IKTA)*
- 2004–2004 *Intelligens többnyelvű dokumentumkezelés Eurovoc rendszerben (ITEM)*
- 2003–2005 *Intelligens fordítómemória (IKTA)*
- 2003–2003 *Üzleti metakereső rendszer nyelvi támogatással (ITEM)*
- 2002–2005 *Komplex uráli nyelvészeti adatbázis (NKFP)*
- 2002–2004 *Mondatszintaxis gépi tanulása (IKTA)*
- 2002–2005 *MMIR: egy- és többnyelvű információ-visszakeresés (CIMO)*
- 2001–2003 *Automatikus információszerzés rövid (politikai, üzleti, piaci) hírekből (NKFP)*
- 2001–2004 *Morfológiai-lexikai szint feletti magyar nyelvű tartalomelemző programok kifejlesztése (NKFP)*
- 2001–2003 *Alkalmazott beszédinformációs rendszerek (NKFP)*
- 2001–2003 *Információ- és tudás-tárház (IKTA)*
- 2001–2002 *Helyesen – a szabad szoftverek világában is (SZT)*
- 2000–2002 *Nyelvi visszacsatolás karakter- és beszédfelismerő rendszerek számára (IKTA)*
- 2000–2002 *Magyar nyelvi szófaji egyértelműsítő módszer fejlesztése (IKTA)*

Kardhegy, 2008



Ráth György utca, 2012



**THE IRON CURTAIN HAS BEEN DEMOLISHED,
POLITICAL WALLS EVAPORATED,
BUT THE LANGUAGE WALLS STILL STAND...
LET'S OVERCOME THESE OBSTACLES TOGETHER!**

YOUR PARTNER, MORPHOLOGIC, THE SPECIALIST OF LINGUISTIC TECHNOLOGY, IS THE LARGEST SOFTWARE HOUSE COMMITTED TO THE LANGUAGE INDUSTRY IN CENTRAL & EASTERN EUROPE.

MorphoLogic's present technology supports

- multilingual text retrieval
- applications for machine-aided translation
- building intelligent mono-, bi- and multi-lingual dictionaries
- spelling checkers, grammar checkers, hyphenators
- inflectional thesaurus
- tools for highly inflectional and other morphologically complex languages
- software modules for OEM and distributors



MORPHOLOGIC

**THE BUDAPEST EXPO'96 AND THE 21ST CENTURY ARE FAST
APPROACHING, YOU HAVE TO COMMUNICATE - YOU WILL NEED
THE MORPHOLOGIC TECHNOLOGY**

Fő u. 56-58. I/3., H-1011 Budapest, Hungary, Phone/fax: +36-1-2018-355 E-mail: h6109pro@ella.hu

Dublin, Microsoft, 1995





PHOTO BY MIRIEL BOUQUET

Gábor Prózéký of Morphologic: "Commercialization is good, but the main focus will be on research."

Language whiz soars to success on spell check

BY JIM KHAROUF

So don't you want to be rich?

Gábor Prózéký, director and founder of Hun-

company's big break came when Microsoft's Hungarian version of Word 2.0 was about to be released in 1992. Prózéký called a Microsoft contact, who forwarded his spell checker and

Minding The Ps And Qs

ness director and
Zoner.

Gabor Proszeky may have a doctorate in mathematics, but his real specialty is words.

The founder and director of **Morphologic**, a small Budapest firm with a staff of 12, has carved out a curious niche for his company as a leader in software dictionaries, spell-checkers, hyphenation tools and thesauruses. Its products have brought giants such as Microsoft and IBM knocking on the door for licenses. Morphologic has become the common fixture in most of these company's Eastern European products, as well as a must-have for the region's software companies.

Already, 1,000 colleges program have the company is re-enter the German market. Callisto 3 is also on-line software

*Marek Goscho
a sales company
by a U.S. firm*

Wall Street Journal Europe, 1998



MorphoMouse, 2013





NYELVTUDOMÁNYI INTÉZET
MAGYAR TUDOMÁNYOS AKADÉMIA

Együttműködés a Nyelvtudományi Intézet és a MorphoLogic Kft. között

Váradi Tamás

MTA Nyelvtudományi Intézet
Nyelvtechnológiai és alkalmazott nyelvészet osztály
varadi.tamas@nytud.mta.hu



MorphoLogic

a nyelvészeti programok specialistája



Közös projektek, fejlesztések

- Magyar Nemzeti Szövegtár (1998 – 2017)
- [Automatikus információszerzés rövid \(politikai, üzleti, piaci\) hírekből \(NewsPro\) 2001–2003](#)
- [Magyar-angol gépi fordítórendszer 2004–2007](#)
- [Nyelv- és Beszédtechnológiai Platform 2008–2010](#)
- [Central and South-East European Resources \(CESAR\) 2011–2013](#)
- Magyar WordNet
- Szeged Treebank
- Finnugor nyelvek támogatása



MorphoLogic

a nyelvészeti programok specialistája

Nemzeti és etnikai identitás vizsgálata történelmi
eseményekre vonatkozó szövegek narratív
alakzatainak számítógépes tartalomelemzése
révén (2006 – 2008)



NYELVTUDOMÁNYI INTÉZET
MAGYAR TUDOMÁNYOS AKADÉMIA



László János



Miháltz Márton



MorphoLogic

a nyelvészeti programok specialistája

itranslate4.eu



NYELVTUDOMÁNYI INTÉZET
MAGYAR TUDOMÁNYOS AKADÉMIA

2010 -- 2012



MorphoLogic

a nyelvészeti programok specialistája

Tihanyi László



NYELVTUDOMÁNYI INTÉZET
MAGYAR TUDOMÁNYOS AKADÉMIA

e-magyar.hu
DIGITÁLIS NYELVFELDOLGOZÓ RENDSZER

BEMUTATKOZÁS SZÖVEGFELDOLGOZÁS + BESZÉDFELDOLGOZÁS + SZÖVEGELEMZŐ

e-magyar.hu
BELÉPÉS



MorphoLogic

a nyelvészeti programok specialistája



NYELVTUDOMÁNYI INTÉZET
MAGYAR TUDOMÁNYOS AKADÉMIA



MorphoLogic

a nyelvészeti programok specialistája

A MORPHOLOGIC - BME EGYÜTTMŰKÖDÉSEK



Olaszy Gábor, Németh Géza

Távközlési és Médiainformatika Tanszék,
Budapesti Műszaki és Gazdaságtudományi Egyetem

2017. 11. 24.

Írás, beszéd és számítógép

- ❖ 1980-as évek (Szalay utca 10.)
- ❖ 1999 BME Profivox szövegfelolvasó
- ❖ 2000- 2017 Profivox + Morphologic
- ❖ 2010 Morphologic + BME
- ❖ **KIZÁRÓLAG KUTATÁSI EGYÜTTMŰKÖDÉSEK**

Webfordítás gépi felolvasással



MorphoLogic provides free translation services: text translation, web page multilingual search and sentence parser. Now we translate between Hungarian

web translation

text translation

dictionary

search

parser

speller

translate

English

Hungarian

suggest

send feedback

This is the Hungarian voice of the Morphologic. Really? Yes, the Profivox speech synthesizer reads the text in Hungarian.

Ez a Morphologic magyar hangja. Igazán? Igen, a Profivox beszéd-szintetizátor olvassa a magyar szöveget.



MobiVoice

A MobiVoice modul segítségével a Morphologic hangot adott a MoBiDic rendszernek.

Profivox szöveg-beszéd átalakító ?

KI EMLÉKSZIK ERRE ?

Speech-to-speech translation, STS

A Morphologic STS kísérletében
(angolról magyarra) a
Profivox szöveg-beszéd átalakító volt
a magyar hang

KI EMLÉKSZIK MIKOR VOLT ?

BME-TMIT Elektronikus Magyar Kiejtési Szótár

A Morphologic segít a BME-nek 1,5 millió magyar szóalak osztályozásában.

Feladat: a tulajdonnevek kiválogatása

SIKERES VOLT!

Magyar szavak elektronikus kiejtési szótára - 2010 (1,5 millió szóalak)

» Keresés hangok alapján

A keresett hangsor:

A hangok beviteléhez használjuk a '« Hang' gombot.

Különleges karakterek: - * (csillag) karakter, tetszőleges hango(ka)t helyettesít,
- # (kettőskereszt) hangsorkezdő és hangsor befejező jel.

» Keresés betűk alapján

A keresett betűsor:

Használható karakterek: - a magyar ábécé betűi,
- * (csillag) karakter, tetszőleges betű(ke)t helyettesít,
- # (kettőskereszt) szókezdő és szó befejező jel.

» Keresési beállítások

Keresési terület: teljes szótár családnevek települések
Hangszimbólumok: IPA magyar betűk TMIT kódjelölés SAMPA
Megjelenítés: db Meghallgathatók

[Bővebb információ »](#)

» A keresés eredménye

Találatok száma: 25 Megjelenített tételek: 1-25

	Miskolc	[miʃkolʈs]
	Miskolccal	[miʃkolʈsɔl]
	Miskolcé	[miʃkolʈse:]
	Miskolcéhoz	[miʃkolʈse:hoz]
	Miskolchoz	[miʃkolʈshoz]
	miskolci	[miʃkolʈsi]
	miskolciak	[miʃkolʈsiɔk]
	miskolciaké	[miʃkolʈsiɔke:]
	miskolciakért	[miʃkolʈsiɔke:rt]
	miskolciakéval	[miʃkolʈsiɔke:vɛl]

<http://magyarbeszed.tmit.bme.hu/mksz>

Köszönöm



A FIGYELMET!