



# Introducing MorphoLogic to LIRICS

**Gábor Prószéky**  
proszeky@morphologic.hu

**MorphoLogic**  
<http://www.morphologic.hu>

**Pázmány Péter Catholic University**  
**Faculty of Information Technology**  
<http://www.itk.ppke.hu>

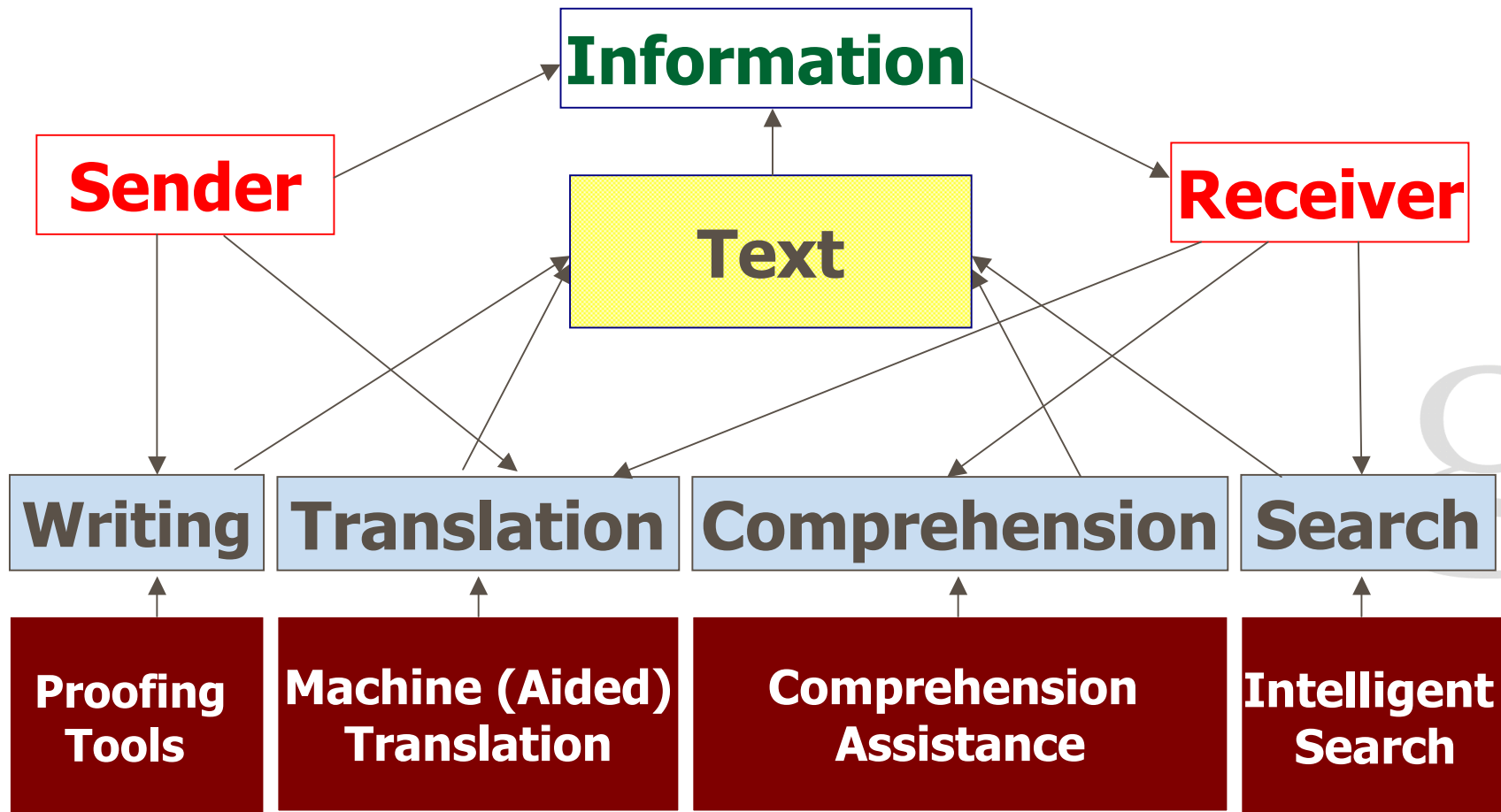
a  
b  
c  
d  
f  
h  
i  
j  
k  
l  
m

# MorphoLogic

- **Established in 1991**
- **Private enterprise**
- **14 years in HLT: from basic research to applications**
- **Personnel:  
started with 3 (1991);  
30 in HLT + 10 in Localization (2005)**

a  
b  
c  
d  
f  
h  
i  
k  
l  
m

# Basic Fields of R&D Activity at MorphoLogic



a  
b  
c  
d  
f  
h  
i  
k  
l  
m

## Basic HLT Resources at MorphoLogic

- **Morphologies:** Hungarian, Polish, Czech, Romanian, Slovak, German, English, Spanish, French, Italian, Zuryen, Nganasan, ...
- **Syntactic Descriptions:** English, Hungarian, ...
- **Lexical databases:** Hungarian explanatory dictionaries, synonym/antonym dictionaries, dictionaries of idioms, collocations and other MWEs, bilingual databases (English–Hungarian, German–Hungarian, French–Hungarian), terminological databases
- **Ontology:** building the Hungarian WordNet
- **Building Corpora:** monolingual (Hungarian) morphologically annotated corpus, Hungarian treebank, bilingual corpora (English–Hungarian)
- **Everything in XML**

## MorphoLogic's Industrial Tools

- **Proofing Tools** (end-user products and corporate solutions)  
Spelling Checkers,  
Grammar Checkers,  
Inflectional Thesauri,  
Hyphenators  
[licenced by Microsoft, Lotus, Xerox, Franklin, ...;  
award: Compfair-93]
- **Linguistic Search Support** for Various Languages  
(**MorphoStem**)  
[lic: Microsoft]
- **Dictionaries** for Various Language Pairs (local, intranet, internet)  
(**MoBiDic, MoBiWeb**)  
[aw: Compfair-96]
- **Comprehension Assistants** for Various Languages  
(**MoBiMouse Plus, MoBiCAT**)  
[aw: EU IST-Prize-99]
- **English-Hungarian Machine Translation**  
(**MetaMorpho: MorphoWeb, MorphoWord**)

## MorphoLogic's Other Industrial Activities

- **Localization for various companies (biggest partners: SAP, IBM, ...)**
- **Translation of international registration of trade marks (from Czech, Polish, Slovak, Lithuanian, Latvian and Hungarian into English)**
- **„XML-ization“ for various partners**

## Some Recently Finished MorphoLogic R&D Projects

- **Spotting and Analyzing Multi-Word Expressions (with Universiteit Groningen)**
- **Linguistic Feedback to Recognition Systems**
- **Machine Learning of Hungarian Syntax**
- **Psychological Text Processing**
- **Morphological Description of Small Endangered Uralic Languages**
- **Information Extraction from Political and Economical Texts (with Gallup)**
- **English–Hungarian Machine Translation**
- **Phonological Converter**
- **Automatic Rule Builder to Translation Applications**

## Running MorphoLogic R&D Projects

- **Building multilingual terminology: EuroTermBank (DK, D, PL, EE, LT, LV, HU)**
- **Intelligent Translation Memory: organic combination of TM with MT**
- **Controlled language applications**
- **„Digital Terminologist“**
- **Intelligent Translation Groupware**
- **Information Extraction in Medical Applications**
- **Building and Applying Ontology in Information Extraction**
- **A new guesser**
- **Hungarian–English Machine Translation**



## **Some needs and expectations – standards from LIRICS**

- **Standards for description of variant and invariant parts of MWEs**
- **Standardized treatment of special categories: „inflected inflections“, postpositions (with frames)**
- **Standards for marking neutral syntactic positions**
- **Standardization of special forms in human readable dictionaries (e.g. tilde with capitalization)**



```

>Barcelona
Analysis of "Barcelona":
└Barcelona[FN]—[NOM]

>barcelonai
Analysis of "barcelonai":
└Barcelona[FN]=barcelona—i [IKEP]—[NOM]

>barcelonaias
Analysis of "barcelonaias":
└Barcelona[FN]=barcelona—i [IKEP]└as [SFN]—[NOM]
                                   └as [SKEP]—[NOM]

>barcelonaiasodik
Analysis of "barcelonaiasodik":
└Barcelona[FN]=barcelona—i [IKEP]—as [SKEP]—od=ik [MI]=od—ik [e3]

>elbarcelonaiasodik
Analysis of "elbarcelonaiasodik":
└el [IK]—Barcelona[FN]=barcelona—i [IKEP]—as [SKEP]—od=ik [MI]=od—ik [e3]

>elbarcelonaiasodott
Analysis of "elbarcelonaiasodott":
└el [IK]—Barcelona[FN]=barcelona—i [IKEP]—as [SKEP]—od=ik [MI]=od└ott [MIB]—[NOM]
                                                                    └ott [Me3]

>legelbarcelonaiasodottabb
Analysis of "legelbarcelonaiasodottabb":
└leg [FF]—el [IK]—Barcelona[FN]=barcelona—i [IKEP]—as [SKEP]—od=ik [MI]=od—ott [MIB]—abb [FOK]—[NOM]

>legelbarcelonaiasodottabbak
Analysis of "legelbarcelonaiasodottabbak":
└leg [FF]—el [IK]—Barcelona[FN]=barcelona—i [IKEP]—as [SKEP]—od=ik [MI]=od—ott [MIB]—abb [FOK]—ak [PL]+[NOM]

>legelbarcelonaiasodottabbaké
Analysis of "legelbarcelonaiasodottabbaké":
└leg [FF]—el [IK]—Barcelona[FN]=barcelona—i [IKEP]—as [SKEP]—od=ik [MI]=od—ott [MIB]—abb [FOK]—ak [PL]+é [POS]+[NOM]

>legelbarcelonaiasodottabbakéi
Analysis of "legelbarcelonaiasodottabbakéi":
└leg [FF]—el [IK]—Barcelona[FN]=barcelona—i [IKEP]—as [SKEP]—od=ik [MI]=od—ott [MIB]—abb [FOK]—ak [PL]+éi [POSi]+[NOM]

>legelbarcelonaiasodottabbakéival
Analysis of "legelbarcelonaiasodottabbakéival":
└leg [FF]—el [IK]—Barcelona[FN]=barcelona—i [IKEP]—as [SKEP]—od=ik [MI]=od—ott [MIB]—abb [FOK]—ak [PL]+éi [POSi]+val [INS]

```

**meaning: “with the belongings of those who have been becoming most similar to the inhabitants of Barcelona”**