

*István Micsinai – Csaba Molnár:*

# Analysing Parliamentary Data in Hungarian

# The Hungarian Comparative Agendas Project

Participant of international Comparative Agendas Project

Datasets on:

Laws (1949-2014)

Decrees (1990-2014)

Budgets (1868-2014)

Interpellations (1949-2014)

Daily newspapers' front pages (1990-2014), etc.

Collected manually and automatically

Dates, details of introduction, policy content etc.

Both in English and Hungarian

Website: [cap.tk.mta.hu/en](http://cap.tk.mta.hu/en)

# Hungarian Grammar

Agglutinative language

Numerous suffixes, 18 grammatical cases, no grammatical gender

Example: elszámoltathatóságáról ('on his/her accountability') = el  
+ **számol** + tat + hat + ó + ság + á +ról

Diacritical marks (á, é, í, ó, ö, ő, ú, ü, ű)

Sometimes changes in the root word

Example: darázs ('wasp') → darazs + ak ('wasp + s')

# Geographical names in parliamentary questions

Miklós Sebők – Bálint Kubik – Csaba Molnár – Gábor Szendi (*under publication*) Exercising Control and Gathering Information: The Functions of Interpellations in Hungary (1990-2014)

5<sup>th</sup> Hypothesis: Interpellations of MPs elected in single-member district contains more geographical (settlement) names than the ones' elected in regional or national lists.

Since the end of the 19th century each town and village has unique name (e.g.: Vári → Mezővári and Gyulavári)

# Geographical names in parliamentary questions

## Sources:

Database of National Electoral Office on settlement names (N=3162) + districts of Budapest (N=23) + names of counties (N=19) + name of Lake Balaton (N=1)

Corpus: text of interpellations (1990-2014)

## Methodology:

Combination of every settlement names and every possible suffixes

Removal of diacritical marks (á, é, í, ó, ö, ő, ú, ü, ű)

Adding irregular root changes (e.g. **Eger** – **egri**)

Automatic dictionary-based method

Manual validation

# Geographical names in parliamentary questions

915 geographical names appeared in 1597 interpellations (32.5%)

Limitations:

Geographical names have usually other meanings too (e.g. Polgár – ‘civil’, Levelek – ‘Letters’)

Many geographical names are derived from personal names (e.g. **Gyula** Horn, Péter **Medgyessy**, Gordon **Bajnai**)

Does not contain nicknames of settlements (e.g. Csaba for Békéscsaba) and sub-settlement-level geographical names (e.g. Kisterenye in Bátorterenye)

# Type of connections among laws

Miklós Sebők – Bálint Kubik – Csaba Molnár (under publication) The Formal Quality of Legislation: An Empirical Analysis

Research question: on which variables does the quality (stability) of accepted laws depend ?

Source:

Previous named-entity recognition of Optijus Legal Database (119,600 references of 3,357 other laws)

Decree 61/2009 (XII.14) of the Ministry of Justice and Law Enforcement on the Drafting of Legal Statutes (keyword list of law amendment) + some modifications

Type	Keywords
Reregulation	“fejezet helyébe/helyére következő fejezet lép”
	“alcím helyébe/helyére következő alcím lép”
	“§ helyébe/helyére következő rendelkezés lép”
	“bekezdés helyébe/helyére következő rendelkezés lép”
	“pont helyébe/helyére következő rendelkezés lép”
	“alpont helyébe/helyére következő rendelkezés lép”
Supplementary amendment	“következő fejezet egészül ki”
	“következő alcím egészül ki”
	“a következő § egészül ki”
	“következő bekezdés egészül ki”
	“bekezdés a következő pont egészül ki”
	“következő alpont egészül ki”
Textual specification	“szövegrész helyébe/helyére szöveg lép”
Amendment concerning both text and promulgation	“szövegrész helyett szöveggel lép hatályba”
Amendment concerning promulgation	“nem lép hatályba”
	“hatály vesz”



# Type of connections among laws

## Methodology:

Automatic dictionary-based method

N=107,407 amendment-type connection

Several rounds of manual validation on a random sample of laws

Randomly selected sample of 500 observations

5% error rate → margin of error between 3.09%-6.91% for the entire database

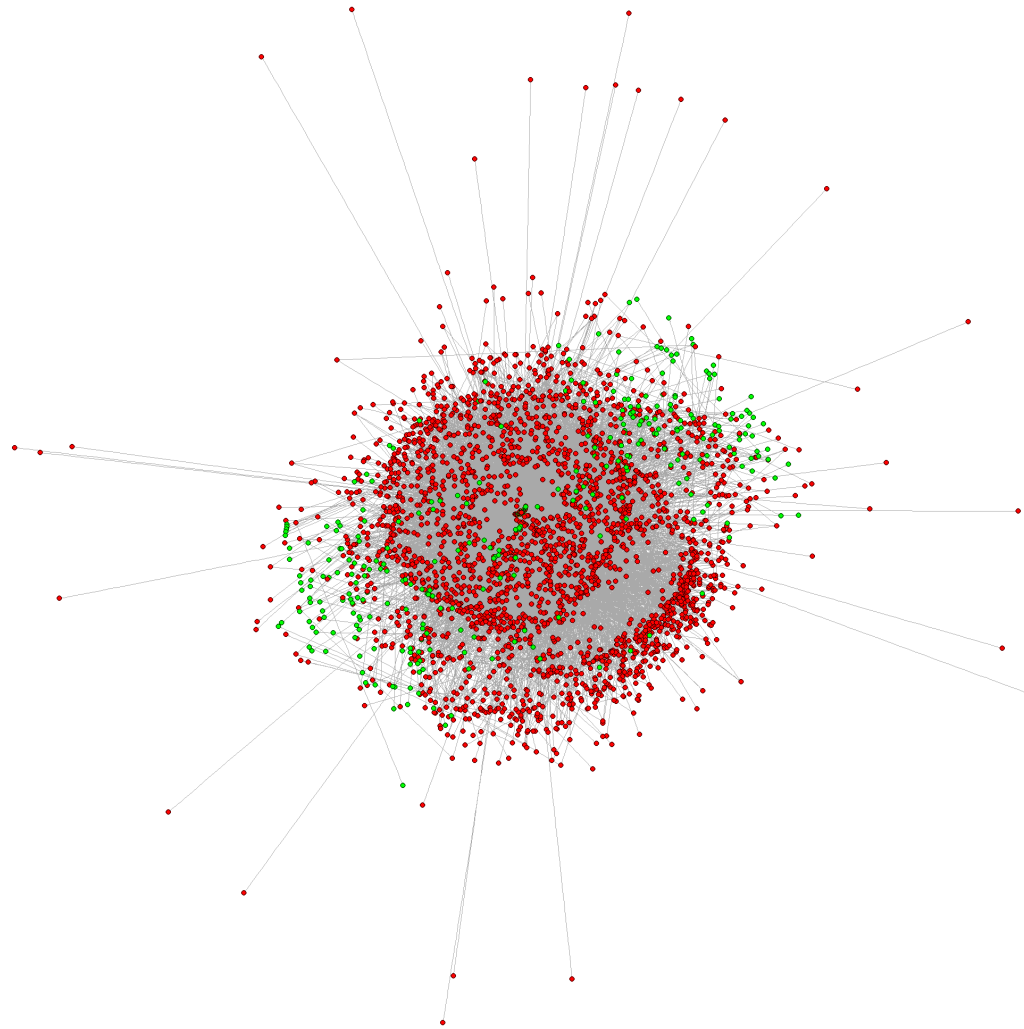
## Difficulties:

Defining keywords

Defining the length of examined text around the named-entities:

Too short: may not contain keyword

Too long: may contain keyword of other named-entity



Thank you for your attention!

[molnar.csaba@tk.mta.hu](mailto:molnar.csaba@tk.mta.hu)  
[micsinai.istvan@tk.mta.hu](mailto:micsinai.istvan@tk.mta.hu)  
[cap.tk.mta.hu/en](http://cap.tk.mta.hu/en)