

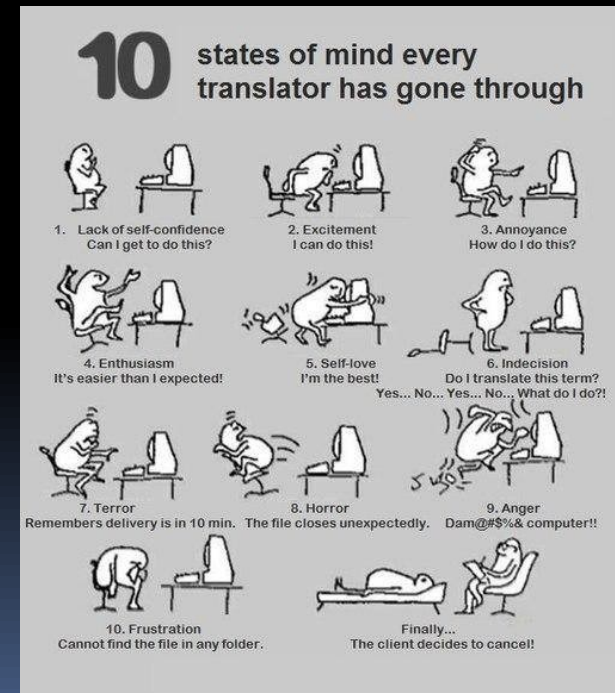
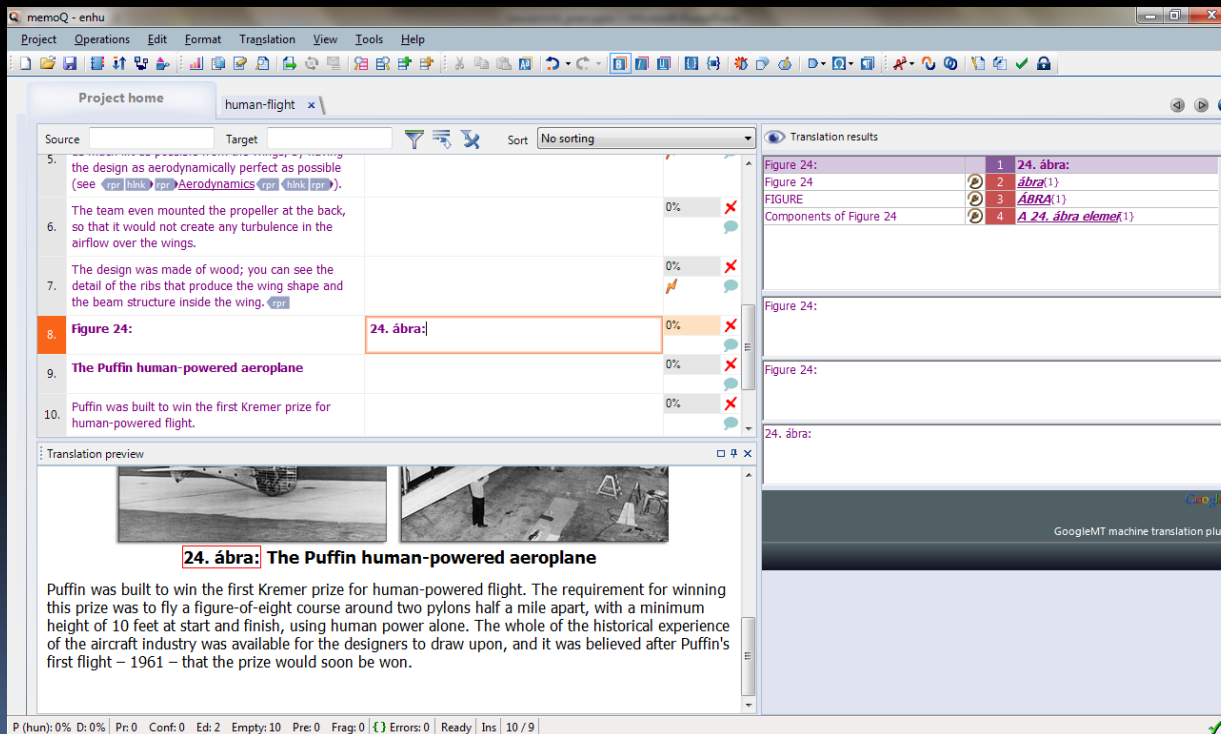
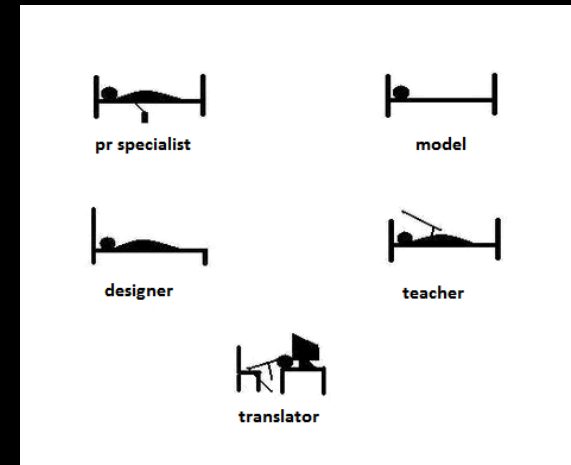
STATISZTIKAI GÉPI FORDÍTÁS ALAPJAI

Készítette: Szabó Gábor


Konzulens: Dr. Juhász Sándor

Kontextus

- Fordítói (translator) munka megkönnyítése



Automatikus kiegészítés, prediktív gépelés



PREDICTIVE TEXT

*It's like Marmite.
You either love it or hate it.
There is no in between!*

I'll be flying to Toronto next

Sent from my Windows Phone

to week year time

q w e r t y u i



Egyszer mikor|



egyszer mikor **micimackónak semmi dolga nem akadt**

egyszer mikor **micimackó álló répával felébredt**

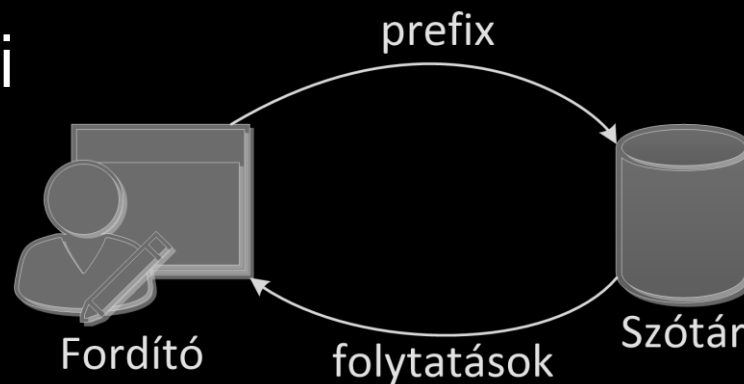
egyszer mikor **micimackónak semmi dolga**

egyszer mikor **micimackónak semmi**

Nyomja meg az Enter gombot a kereséshez.

Feladat

- **Szótár alapú predikci**



- Többszavas kifejezések képzési módjának összehasonlítása
- Jóság mértéke: **spórolás** (+ **szótár mérete**)

Prediktív szótár

■ Ké

■ Ny

(19

■ Kifejezések együttes előfordulása alapján

score ($\geq \text{min}$)

■ Szé

□ (

| Angol | Magyar |
|--|---|
| Her mood was totally great, after all . | Végül is nagyszerű hangulata volt. |
| ... | ... |
| After all , we only have one life. | Végül is egyszer élünk. |
| ... | ... |

tése

| Forrás kif. | Lehetséges fordítások |
|-------------|--|
| I would | (je; 0.58), (voudrais; 0.44), (je voudrais; 0.31) |
| new | (nouvelle; 0.40) |
| Parliament | (Parlement; 0.89), (Parlement européen; 0.47) |
| and | (et; 0.89), (de; 0.64), ... |

Mérés

- Bemeneti szöveg:



Proceedings of the European Parliament

- Angol-francia nyelvpár



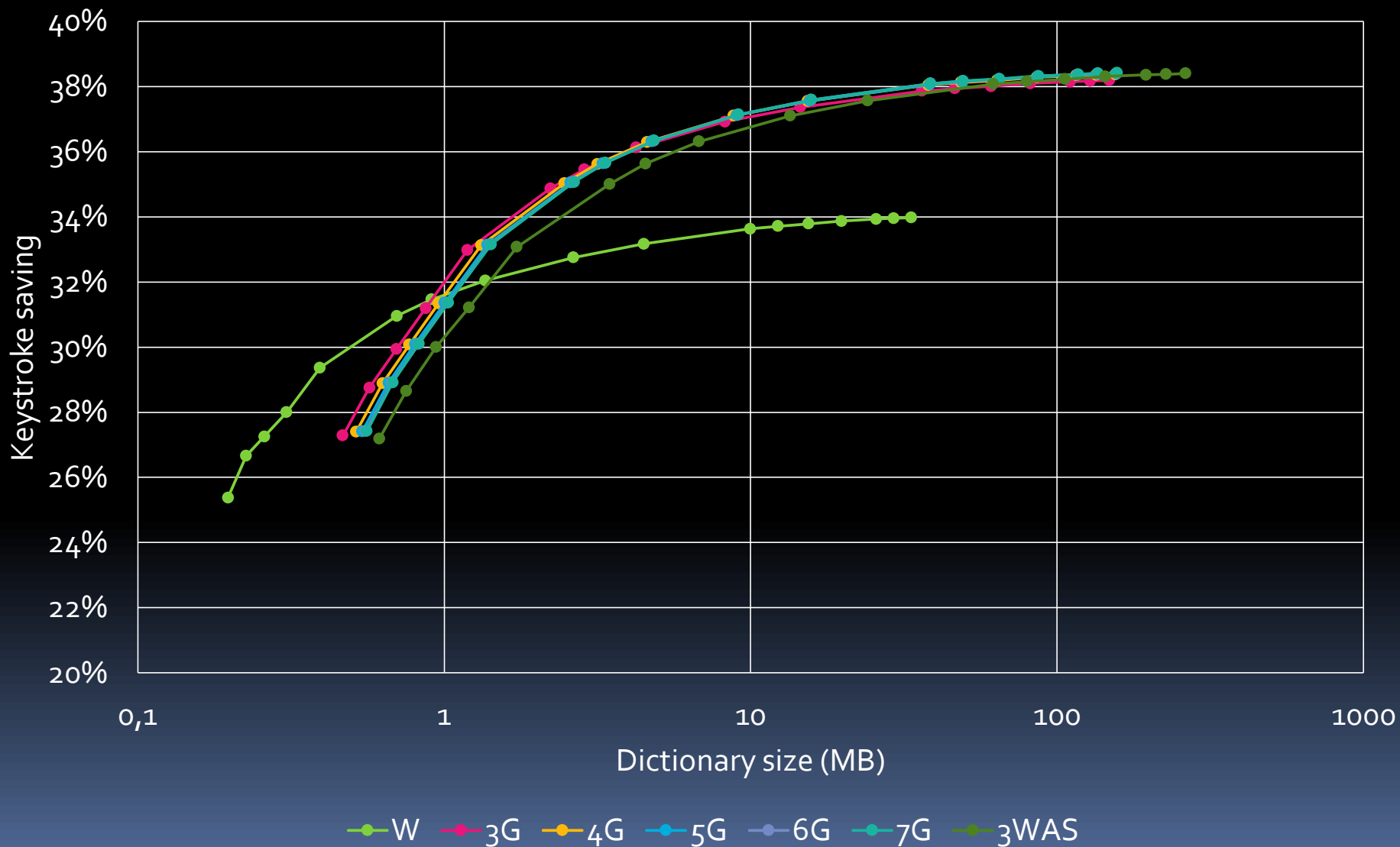
- Kifejezés képzési módok:

- 1, 2, 3, ..., 7-gram

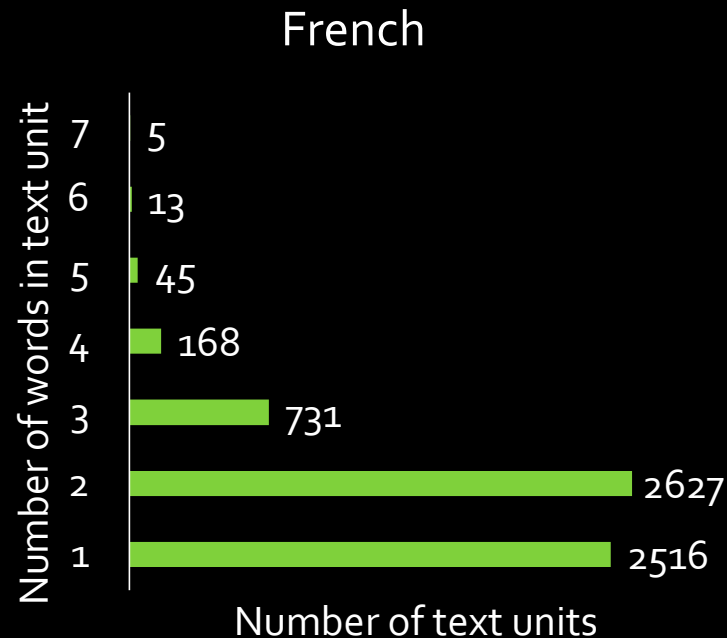
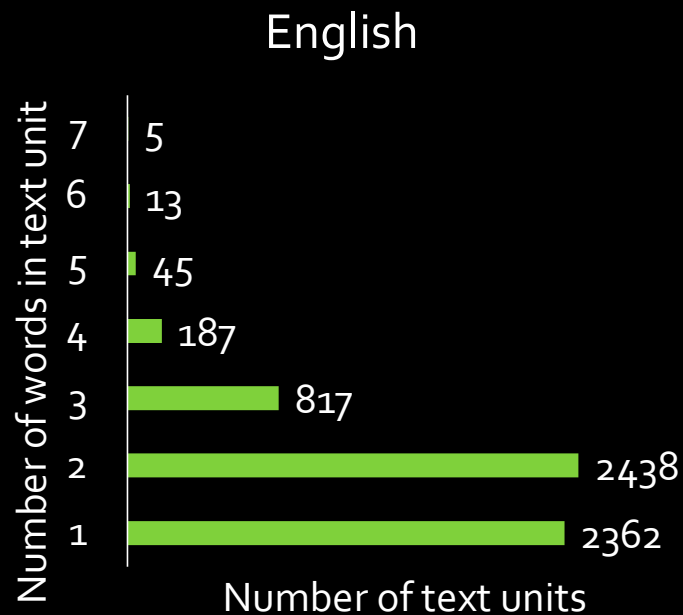
- n szavas csúszóablak összes részhalmaza ($n = 3$)

- 200k mondat tanítás, 20k ellenőrzés
- Lényeg: $spórolás = f(szótárméret)$

Mérési eredmények

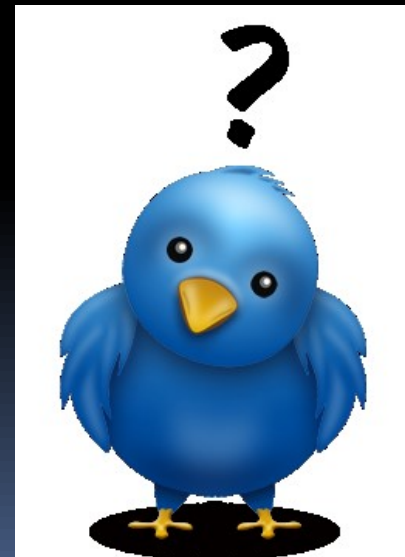
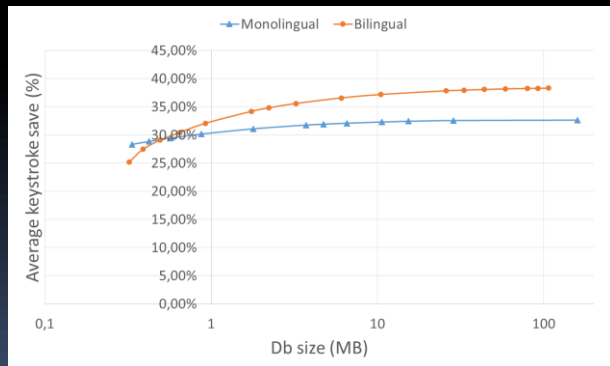
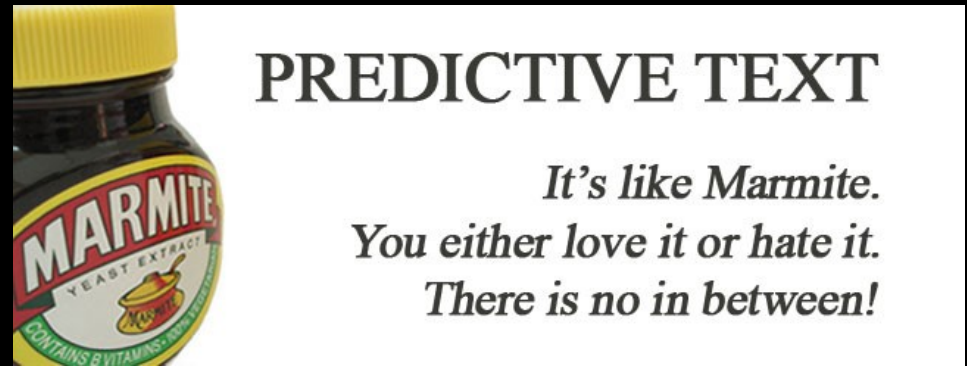
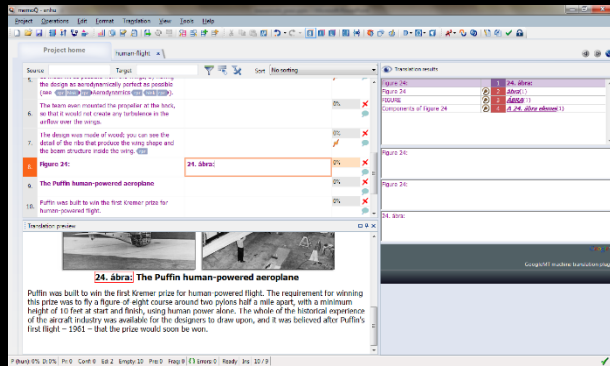


Mérési eredmények



- 7 szavas kifejezések:
 - The vote will take place tomorrow at
 - the Committee on Legal Affairs and the
 - Committee on Legal Affairs and the Internal
 - on Legal Affairs and the Internal Market
 - on behalf of the Group of the

Kérdések ?



Hivatkozások

- OPUS - <http://opus.lingfil.uu.se/>
- Kilgray - <http://www.kilgray.com/>
- Google Image Search