

A szóhasonlóság mértékének tesztelése
CVCVC szerkezetű hangkivető főnevekkel

Rung András

BME Fizikai Intézet

Alapelvek

- ◆ **Produkción és megértés analógiás alapon szabályok helyett**
- ◆ Az analógiás források kiválasztásához pontos **hasonlósági mérce** kell
- ◆ A pontos mérce számítógépesen **modellezhető**
- ◆ A hasonlósági mércének **mentálisan reálisnak** kell lennie
- ◆ „az anyanyelvi beszélők a fonológiailag hasonlóknak érzett/tartott elemeket morfofonológiailag hasonlóknak hajlamosak venni” Rebrus és Törkenczy 2008: 757
- ◆ Minél hasonlóbb egy szó az analógia forrásául szolgáló szóhoz, annál hasonlóbban fog viselkedni

Célok

- ◆ Megmutatni, hogy a hasonlóság toldalékolás szempontjából a **jobb széleken** fontosabb
- ◆ Nem csak a **legszélső** fonéma számít
- ◆ Megmutatni, hogy analógiás algoritmikus modellünkkel **modellezhető** a viselkedés

Előzmények

- ◆ Lukács Ágnes: Alaktanilag kivételes tövek vizsgálata a magyarban (2002)
- ◆ **Hangok hasonlóságának hatása a szóalakok hasonlóságára a **pozíció** függvényében**
- ◆ **Nem csak hangkivetőket** vizsgált
- ◆ Nem a pozíciók fontosságára volt kiélezve a teszt
- ◆ **Alacsony elemszám, nem szisztematikus eltérések**, nem tökéletes illeszkedés a sémára
- ◆ Nem volt megfelelő korpuszháttér (gyakoriság számítása)

Módszer

- ◆ Online teszt (30-40 perc), 110 kérdés (4 változatban)
- ◆ Első mondat tesztelő
- ◆ Másodikban kiegészítést vártam
- ◆ Öt/három választási lehetőség
- ◆ Nem közöltem, hogy álszavak

A szopor vállon hordható
prém volt. A főleg
nemesasszonyok
engedhették meg
maguknak.

- szoprat
- szoprot
- szoport
- szoporot
- szoporat



Tesztanyag

- ◆ Valódi CVCVC szerkezetű hangkivető főnevekre hasonlító CVCVC szerkezetű álszavak (364 db), mint pl. *böcök, purok*
- ◆ Fontos, hogy a **tesztiszavak valódi szavakat is felidézzenek** és megfeleljenek a $\sim VC_{\alpha}(o/e/ö)C_{\beta}\#$ **sémának**
- ◆ Minden valódi szónak négy változatát készítettem el: pl. *kapor: gapor, kopor, kabor, kapol*
- ◆ **Lecserélésre használt fonéma a lehető legközelebbi legyen a lecserélthez**
- ◆ A tesztben szereplő álszavakat úgy hoztam létre, hogy a valódi szavak **1., 2., 3., illetve 5. fonémáját lecseréltem egy másikra.**

Fonémák a pozíciókban

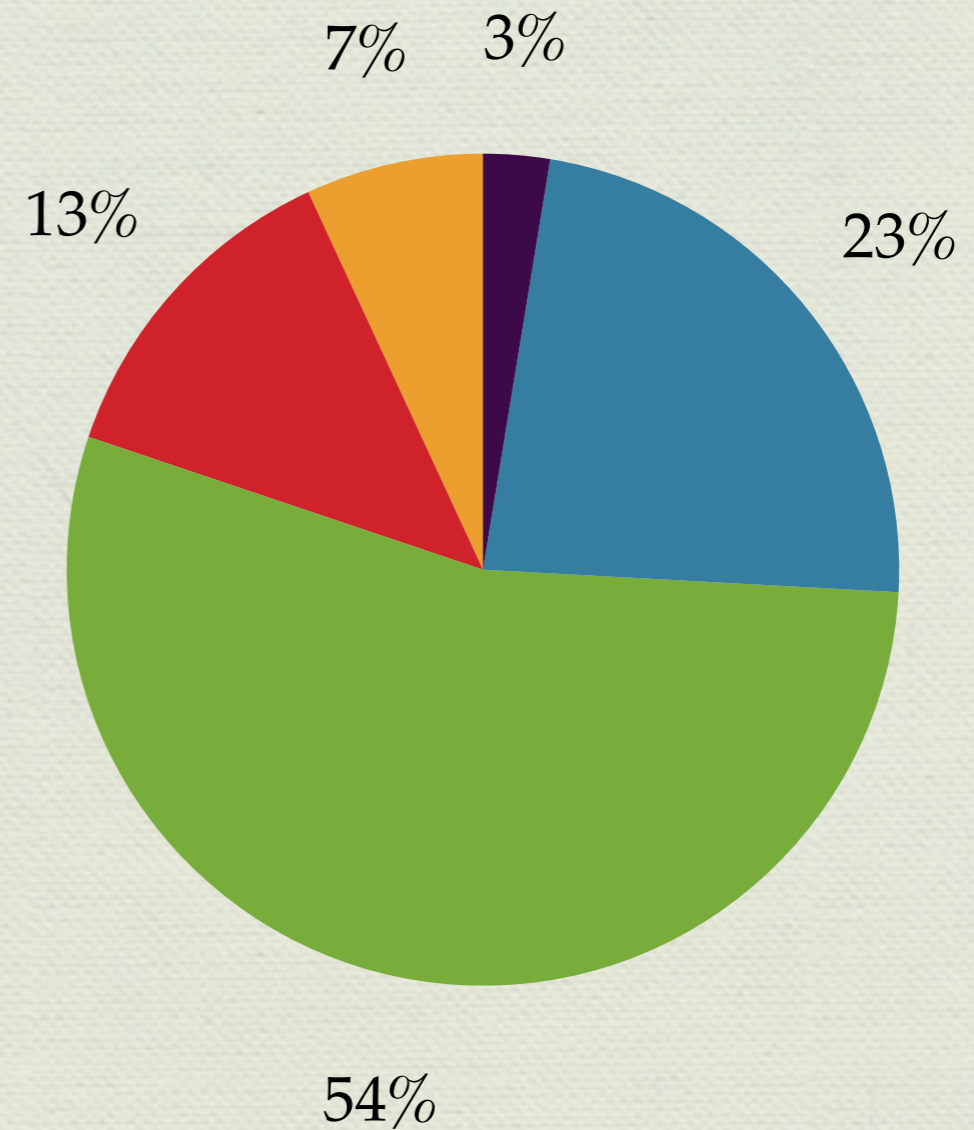
| | |
|--------------------|---|
| 1. pozíció fonémái | <i>b, c, cs, d, f, g, gy, h, k, l, m, ny, p, r, s, sz, t, v</i> |
| 2. pozíció fonémái | <i>e, i, ö, ü, o, u, a</i> |
| 3. pozíció fonémái | <i>b, c, cs, d, g, gy, j, k, l, m, p, r, s, sz, t, ty, z</i> |
| 5. pozíció fonémái | <i>cs, g, j, k, l, m, n, ny, r, sz, z</i> |

Résztvevők

116 magyar anyanyelvű beszélő

75 nő, 51 férfi

36,8%-os hangkivetés



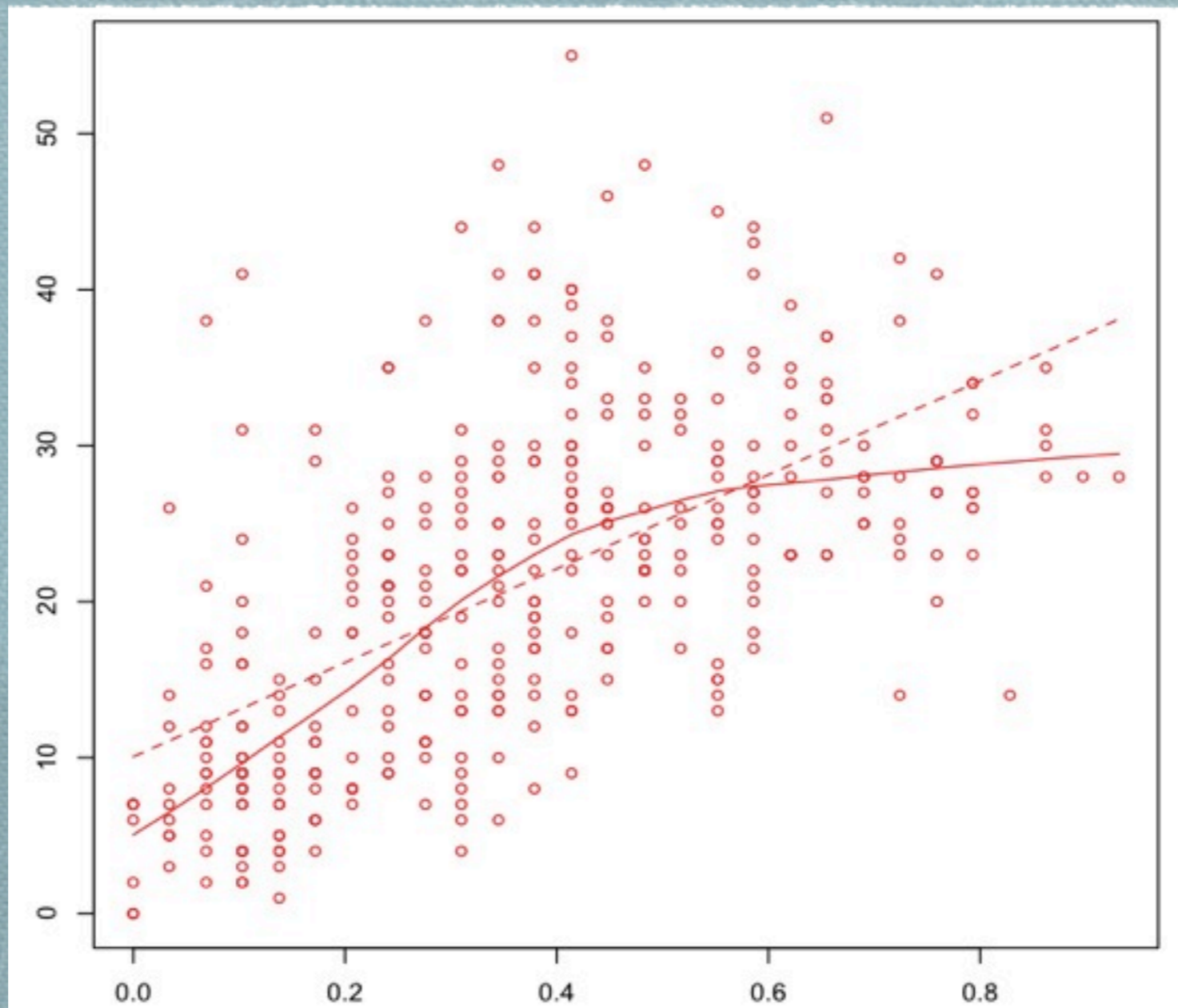
● 14-17 ● 18-24 ● 25-39 ● 40-59 ● 60+

GLM ismételt mérések

| Változtatás helye | hangkivetés mértéke | Szignifikáns eltérések | szórás | min. | max. | leginkább hangkivető szavak |
|-------------------|---------------------|------------------------------------|--------|------|-------|---------------------------------------|
| 1. fonéma | 44,2% | > 3. fonéma *** > 5. fonéma *** | 21,9 | 0% | 93,1% | lücsök, rucsok, böcök, pürök, dücsök |
| 2. fonéma | 41,5% | > 5. fonéma *** | 20,1 | 0% | 86,2% | pücsök, vücsök, vocok |
| 3. fonéma | 39,2% | > 5. fonéma *** | 22,1 | 0% | 79,3% | vöcsök, rücsök, tücsök, tüşzök, surom |
| 5. fonéma | 22,5% | | 13,5 | 0% | 58,6% | sulyog, hurocs, bögül |

Pozíció és a hangkivetés mértékének korrelációja

| | 1. fonéma változtatás | 2. fonéma változtatás | 3. fonéma változtatás | 5. fonéma változtatás |
|----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Szószablya tárgyeset | 0,320** | 0,263* | 0,345*** | 0,205 (n.sz.) |
| Google tárgyeset | 0,279** | 0,207* | 0,227* | 0,185 (n.sz.) |



Az algoritmikusan meghatározott leghasonlóbb hangkivető szavak számának (100 leghasonlóbból) és a tesztszavak hangkivetési mértékének együttjárása

A tesztszavak hangkivetését befolyásoló változók (lépésenkénti lineáris regresszió)

| | Béta | Szignifikancia |
|--|--------|----------------|
| Legközelebbi hangkivető szavak száma (komplex tengelymérték) | 0,459 | $p < 0,001$ |
| Eredeti szavak hasonlósága a gyakorisági prototípusokhoz (komplex tengelymérték) | -0,463 | $p < 0,001$ |
| Hangkivetési mérték (Szószablya) | 0,250 | $p < 0,001$ |
| Legközelebbi szó hangkivető-e (komplex jegymérték) | 0,174 | $p < 0,001$ |
| Eredeti szó hasonlósága a 0,9-es küszöbérték mellett kiválasztott prototípusokból és a leggyakoribb szavakból létrehozott halmaz szavaihoz | 0,268 | $p < 0,01$ |
| Eredeti szó hasonlósága a 0,7-es küszöbérték mellett kiválasztott prototípusokhoz (komplex tengelymérték) | -0,120 | $p < 0,01$ |
| Tárgyesetű alak hangkivetési mértéke (Google) | -0,129 | $p < 0,05$ |
| Legközelebbi hangkivető szavak száma (komplex jegymérték) | 0,126 | $p < 0,05$ |

Köszönöm a figyelmet!
rungandras@gmail.com