

Megastudy paradigma i procesiranje jezika

Izlaganje za Znanstveni kolokvij

Denis Vlašiček

6. lipnja 2020.

Megastudy paradigm

Uvod

- tradicionalni način ispitivanja — faktorijalni eksperimenti

Uvod

- tradicionalni način ispitivanja — faktorijalni eksperimenti
- *megastudy*: velika istraživanja procesiranja jezika

Uvod

- tradicionalni način ispitivanja — faktorijalni eksperimenti
- *megastudy*: velika istraživanja procesiranja jezika
- prednosti pred faktorijalnim nacrtima (Balota i sur., 2012; Keuleers i Balota, 2015)
 - ▷ izjednačavanje podražaja na relevantnim dimenzijama
 - ▷ korištenje kontinuiranih umjesto kategorijalnih mjera

- ishod su velike, (često) otvorene baze podataka
 - ▷ leksičke karakteristike (duljina riječi, učestalost u jezičnim korpusima)
 - ▷ psiholingvističke karakteristike (apstraktnost, polisemija)
 - ▷ mjere procesiranja (vrijeme reakcije, trajanje fiksacija)

- ishod su velike, (često) otvorene baze podataka
 - ▷ leksičke karakteristike (duljina riječi, učestalost u jezičnim korpusima)
 - ▷ psiholingvističke karakteristike (apstraktnost, polisemija)
 - ▷ mjere procesiranja (vrijeme reakcije, trajanje fiksacija)
- omogućavaju provođenje virtualnih eksperimenata (Kuperman, 2015)

- ishod su velike, (često) otvorene baze podataka
 - ▷ leksičke karakteristike (duljina riječi, učestalost u jezičnim korpusima)
 - ▷ psiholingvističke karakteristike (apstraktnost, polisemija)
 - ▷ mjere procesiranja (vrijeme reakcije, trajanje fiksacija)
- omogućavaju provođenje virtualnih eksperimenata (Kuperman, 2015)
- olakšavaju testiranje novih leksičkih i psiholingvističkih varijabli (Yarkoni i sur., 2008)

Što nam megastudije govore o jeziku?

- varijabla koja objašnjava uvjerljivo najviše varijance je učestalost riječi u jeziku (Balota, Yap i Cortese, 2006; Harley, 2014, ch. 6)

Što nam megastudije govore o jeziku?

- varijabla koja objašnjava uvjerljivo najviše varijance je učestalost riječi u jeziku (Balota, Yap i Cortese, 2006; Harley, 2014, ch. 6)
 - ▷ kako definirati i mjeriti učestalost?

Što nam megastudije govore o jeziku?

- varijabla koja objašnjava uvjerljivo najviše varijance je učestalost riječi u jeziku (Balota, Yap i Cortese, 2006; Harley, 2014, ch. 6)
 - ▷ kako definirati i mjeriti učestalost?
 - ▷ subjektivna učestalost

Što nam megastudije govore o jeziku?

- varijabla koja objašnjava uvjerljivo najviše varijance je učestalost riječi u jeziku (Balota, Yap i Cortese, 2006; Harley, 2014, ch. 6)
 - ▷ kako definirati i mjeriti učestalost?
 - ▷ subjektivna učestalost
- duljina riječi (Brysbaert i sur., 2016; Ferrand i sur., 2018)

Što nam megastudije govore o jeziku?

- varijabla koja objašnjava uvjerljivo najviše varijance je učestalost riječi u jeziku (Balota, Yap i Cortese, 2006; Harley, 2014, ch. 6)
 - ▷ kako definirati i mjeriti učestalost?
 - ▷ subjektivna učestalost
- duljina riječi (Brysbaert i sur., 2016; Ferrand i sur., 2018)
- ortografska sličnost ili gustoća ortografskog susjedstva (Coltheart i sur., 1977; Yarkoni i sur., 2008)

- vrijedni resursi za proučavanje jezika
 - ▷ generalizacija nalaza na cjelokupni jezik (Yarkoni, 2019)

- vrijedni resursi za proučavanje jezika
 - ▷ generalizacija nalaza na cjelokupni jezik (Yarkoni, 2019)
 - ▷ „prototipno“ testiranje hipoteza

- vrijedni resursi za proučavanje jezika
 - ▷ generalizacija nalaza na cjelokupni jezik (Yarkoni, 2019)
 - ▷ „prototipno“ testiranje hipoteza
 - ▷ preciznija procjena parametara i prijatelji (Yarkoni i Westfall, 2017)

Što nam megastudije ne govore o jeziku?

„In certain circles there is an almost religious faith that we can find the answers to [causal] questions in the data itself, if only we are sufficiently clever at data mining. However, [...] causal questions can never be answered from data alone.“ Pearl i Mackenzie (2018, str. 351)

- veliki setovi podatka ne mogu nadomjestiti reprezentativnost uzorka (Meng, 2018)

- veliki setovi podatka ne mogu nadomjestiti reprezentativnost uzorka (Meng, 2018)
- standardni statistički modeli ne mogu zamijeniti znanstvenu ekspertizu (McElreath, 2020; Navarro, 2019, 2020)

- veliki setovi podatka ne mogu nadomjestiti reprezentativnost uzorka (Meng, 2018)
- standardni statistički modeli ne mogu zamijeniti znanstvenu ekspertizu (McElreath, 2020; Navarro, 2019, 2020)
- važnost razvoja formalnih teorija (Fiedler, 2004)

Modeli procesiranja jezika

Drift-diffusion model (Ratcliff, 1978)

- zamišljen kao okvir koji može obuhvatiti različite paradigme u kognitivnoj psihologiji

Drift-diffusion model (Ratcliff, 1978)

- zamišljen kao okvir koji može obuhvatiti različite paradigme u kognitivnoj psihologiji
- primijenjen i na zadatak leksičke odluke
 - ▷ do odluke se dolazi kroz proces koji tijekom vremena prikuplja nejasne (*noisy*) informacije
 - ▷ informacije se prikupljaju sve dok se ne dosegne jedan od kriterija za odlučivanje — niz znakova je ili nije riječ

Drift-diffusion model (Ratcliff, 1978)

- zamišljen kao okvir koji može obuhvatiti različite paradigme u kognitivnoj psihologiji
- primijenjen i na zadatak leksičke odluke
 - ▷ do odluke se dolazi kroz proces koji tijekom vremena prikuplja nejasne (*noisy*) informacije
 - ▷ informacije se prikupljaju sve dok se ne dosegne jedan od kriterija za odlučivanje — niz znakova je ili nije riječ
- uspješno modelira i prosječna vremena reakcija i njihove distribucije
- Ratcliff i sur. (2004) pokazali su osjetljivost modela na frekvencije riječi

Bayesian reader (Norris, 2006)

- specifično za čitanje

Bayesian reader (Norris, 2006)

- specifično za čitanje
- kreće od pretpostavke o idealnom opažaču koji donosi optimalne odluke

Bayesian reader (Norris, 2006)

- specifično za čitanje
- kreće od pretpostavke o idealnom opažaču koji donosi optimalne odluke
- optimalno ponašanje ovisi o konkretnom zadatku koji opažač treba obaviti
- podaci koje opažač dobiva u sebi imaju šum, zbog čega mogu nastati pogreške u izvršavanju zadatka

Bayesian reader (Norris, 2006)

- specifično za čitanje
- kreće od pretpostavke o idealnom opažaču koji donosi optimalne odluke
- optimalno ponašanje ovisi o konkretnom zadatku koji opažač treba obaviti
- podaci koje opažač dobiva u sebi imaju šum, zbog čega mogu nastati pogreške u izvršavanju zadatka
- uspješno lovi učinak frekvencije riječi, te učinak ortografske sličnosti (?)

Modeli procesiranja jezika i megastudije

- obilje otvorenih podataka dobivenih u megastudijama pogodno za testiranje različitih modela
 - ▷ kros-validacija
 - ▷ procjena parametara
 - ▷ brzo testiranje prototipova i modifikacija modela
 - ▷ testiranje modela na različitim jezicima

Multiverse analiza

Od QRP do multiverzuma

- I) QRP — eng. *questionable research practices* (John, Loewenstein i Prelec, 2012)

Od QRP do multiverzuma

I) QRP — eng. *questionable research practices* (John i sur., 2012)

- ▷ nisu naveli sve zavisne varijable
- ▷ nisu naveli sve eksperimentalne situacije
- ▷ isključivanje podataka nakon što su pogledali kako to utječe na rezultate

Od QRP do multiverzuma

I) QRP — eng. *questionable research practices* (John i sur., 2012)

- ▷ nisu naveli sve zavisne varijable
- ▷ nisu naveli sve eksperimentalne situacije
- ▷ isključivanje podataka nakon što su pogledali kako to utječe na rezultate

II) mnogo različitih načina na koje se podaci *moгу analizirati* (Gelman i Loken, 2013)

Od QRP do multiverzuma

- I) QRP — eng. *questionable research practices* (John i sur., 2012)
 - ▷ nisu naveli sve zavisne varijable
 - ▷ nisu naveli sve eksperimentalne situacije
 - ▷ isključivanje podataka nakon što su pogledali kako to utječe na rezultate
- II) mnogo različitih načina na koje se podaci *moгу analizirati* (Gelman i Loken, 2013)
- III) mnogo različitih načina na koje se podaci *analiziraju* (Silberzahn i sur., 2018)

IV) *Multiverse* analiza

Reference I

- Balota, D. A., Yap, M. J. i Cortese, M. J. (2006). Visual Word Recognition: The Journey from Features to Meaning (A Travel Update). U: *Handbook of Psycholinguistics* (Second izdanje, str. 285–375). Academic Press.
- Balota, D. A., Yap, M. J., Hutchison, K. A. i Cortese, M. J. (2012). Megastudies: What do Millions (or so) of Trials Tell us About Lexical Processing? U: J. S. Adelman (Ur.), *Visual Word Recognition (Models and Methods, Orthography and Phonology)* (svezak 1, str. 90–115). Hove: Psychology Press.
- Brysbaert, M., Stevens, M., Mandera, P. i Keuleers, E. (2016). The Impact of Word Prevalence on Lexical Decision Times: Evidence from the Dutch Lexicon Project 2. *Journal of Experimental Psychology: Human Perception and Performance*, 42(3), 441–458.
- Coltheart, M., Davelaar, E., Jonasson, J. T. i Besner, D. (1977). Access to the Internal Lexicon. U: S. Dornick (Ur.), *Attention and Performance* (svezak VI, str. 535–556). Erlbaum.
- Ferrand, L., Méot, A., Spinelli, E., New, B., Pallier, C., Bonin, P., ... Grainger, J. (2018). MEGALEX: A Megastudy of Visual and Auditory Word Recognition. *Behavior research methods*, 50(3), 1285–1307.
- Fiedler, K. (2004, May). Tools, Toys, Truisms, and Theories: Some Thoughts on the Creative Cycle of Theory Formation. *Personality and Social Psychology Review*, 8(2), 123–131. doi: 10.1207/s15327957pspr0802_5

Reference II

- Gelman, A. i Loken, E. (2013). The Garden of Forking Paths: Why Multiple Comparisons can be a Problem, Even When There is no “Fishing Expedition” or “p-hacking” and the Research Hypothesis was Posited Ahead of Time. *Department of Statistics, Columbia University*.
- Harley, T. A. (2014). *The Psychology of Language: From Data to Theory* (Fourth izdanje). London, UK: Psychology Press,.
- John, L. K., Loewenstein, G. i Prelec, D. (2012, May). Measuring the Prevalence of Questionable Research Practices With Incentives for Truth Telling. *Psychological Science*, 23(5), 524–532. doi: 10.1177/0956797611430953
- Keuleers, E. i Balota, D. A. (2015, August). Megastudies, Crowdsourcing, and Large Datasets in Psycholinguistics: An Overview of Recent Developments. *Quarterly Journal of Experimental Psychology*, 68(8), 1457–1468. doi: 10.1080/17470218.2015.1051065
- Kuperman, V. (2015, August). Virtual Experiments in Megastudies: A Case Study of Language and Emotion. *Quarterly Journal of Experimental Psychology*, 68(8), 1693–1710. doi: 10.1080/17470218.2014.989865
- McElreath, R. (2020). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. CRC press.
- Meng, X.-L. (2018, June). Statistical paradises and paradoxes in big data (I): Law of large populations, big data paradox, and the 2016 US presidential election. *The Annals of Applied Statistics*, 12(2), 685–726. doi: 10.1214/18-AOAS1161SF

Reference III

- Navarro, D. J. (2019, March). Between the Devil and the Deep Blue Sea: Tensions Between Scientific Judgement and Statistical Model Selection. *Computational Brain & Behavior*, 2(1), 28–34. doi: 10.1007/s42113-018-0019-z
- Navarro, D. J. (2020, March). *If mathematical psychology did not exist we would need to invent it: A case study in cumulative theoretical development* (Preprint). PsyArXiv. doi: 10.31234/osf.io/ygbjp
- Norris, D. (2006). The Bayesian Reader: Explaining Word Recognition as an Optimal Bayesian Decision Process. *Psychological Review*, 113(2), 327–357. doi: 10.1037/0033-295X.113.2.327
- Pearl, J. i Mackenzie, D. (2018). *The Book of Why: The New Science of Cause and Effect*. Basic Books.
- Ratcliff, R. (1978). A Theory of Memory Retrieval. *Psychological Review*, 85(2), 59–108. doi: 10.1037/0033-295X.85.2.59
- Ratcliff, R., Gomez, P. i McKoon, G. (2004). A Diffusion Model Account of the Lexical Decision Task. *Psychological Review*, 111(1), 159–182. doi: 10.1037/0033-295X.111.1.159
- Silberzahn, R., Uhlmann, E. L., Martin, D. P., Anselmi, P., Aust, F., Awtrey, E., ... Nosek, B. A. (2018, September). Many Analysts, One Data Set: Making Transparent How Variations in Analytic Choices Affect Results. *Advances in Methods and Practices in Psychological Science*, 1(3), 337–356. doi: 10.1177/2515245917747646

Reference IV

- Yarkoni, T. (2019, November). *The Generalizability Crisis* (Preprint). PsyArXiv. doi: 10.31234/osf.io/jqw35
- Yarkoni, T., Balota, D. i Yap, M. (2008). Moving Beyond Coltheart's N: A New Measure of Orthographic Similarity. *Psychonomic Bulletin & Review*, 15(5), 971–979.
- Yarkoni, T. i Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science*, 12(6), 1100–1122.