Hizkuntza eta elebitasuna zahartze osasuntsuan: hiztegi mentalaren biltegiratzeari buruzko azterketa (AGINGLEXICON)

Miren Arantzeta [1] eta Josu Goikoetxea [2]

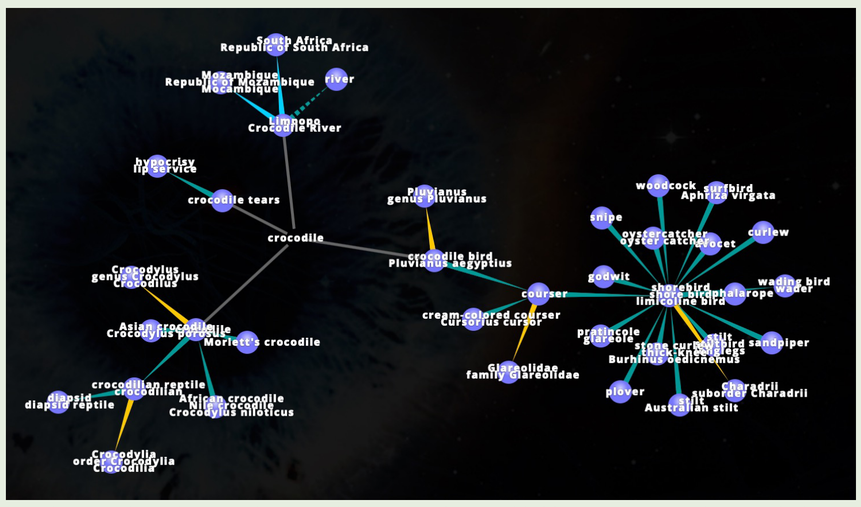
[1] Gogo Elebiduna Psikolinguistika ikerketa taldea. Micaela Portilla ikerketa zentroa. Vitoria-Gasteiz.
[2] HiTZ Hizkuntza Teknologiako Zentroa. Donostiako Informatika Fakultatea.

Funts teorikoa:

- Zahartze osasuntsuan, zailtasunak agertzen dira hiztegiaren sarbidean, eta zehazki jariakortasun semantikoan (‘*Esan ahalik eta animalia izen gehien 2 minutuko tartean*’)[1].

HIZTEGIAREN ANTOLAKETA SEMANTIKOA:

- Kontzeptuak sare batean lotutako nodo gisa irudikatzen dira memoria semantikoan, hurbiltasun semantikoaren printzipioari jarraituz (Difusio Bidezko Aktibazio Eredua [2]) . Alegia, semantikoki lotutako kontzeptuak elkarrengandik hurbilago daude, eta lotura sendoagoak dituzte.



1. Irudia: Ingelesezko Wordnet 3.0 bertsioan ‘Crocodile’ hitzaren inguruko auzokide semantikoak eta beraien arteko erlazioak.

- Kenett et al. (2017)-ek distantzia semantikoaren funtzioa aztertu dute, sare kontzeptualeko bi nodoren arteko biderik laburrena bezala neurtuta [3]. Ondorioztatu dute:
- Bi kontzepturen arteko erlazioa identifikatzeko probabilitatea murriztu egiten da bi elementuren arteko distantziaren arabera, lau nodoko atalaseraino.
- Harreman semantikoa zehazten duen atalase berak kontzeptu-bikote bakoitza memoriatik berreskuratzeko gaitasuna baldintzatzen du.

ZAHARTZE OSASUNTSUA ETA ELEBITASUNA:

- Prozesamendu semantikoari loturiko hainbat zailtasun gaixotasun neurodegeneratiboen adierazgarri goiztiarrak dira [4].
- Hainbat ikerketek erakutsi dute elebitasunak funtzio neurobabeslea duela. Urte luzez hizkuntza bat baino gehiago hitz egiteak erreserba kognitiboa hobetzen du eta honek aztarna positiboa uzten du zahartzaroan. Adibidez; zenbait dementzia mota 4 urtetik 5 urtera atzeratzen dira hiztun elebidunetan [5].

Ikerketaren helburuak eta aurreikuspenak:

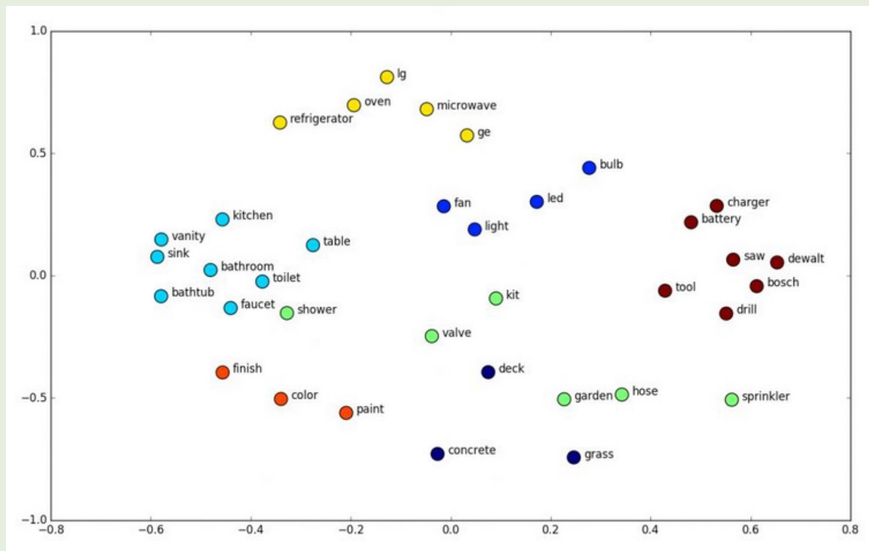
- Bizitzan zehar prozesamendu semantikoak daukan gainbehera aztertzea eta sare semantikoaren narriadura ez-patologikoa bultzatzen duten aldagai esplikatzaileak identifikatzea.
- Elebitasunak sare semantikoen narriadura progresiboan izan dezakeen balio neurobabeslea aztertzea.
- Aurreikuspenak: a) adinarekin bi elementuren arteko harreman semantikoa zehazten duen distantzia gutxitu egingo da; b) hiztun elebidunek adin bereko hiztun elebakarrek baino atalasa altuagoa erakutsiko dute zahartzaroan.

Metodoa:

- PARTE-HARTZAILEAK:
  30 eta 90 urte bitarteko 225 pertsona (150 elebidunak eta 75 elebakarrak)
- ARIKETA:
  1. Parte-hartzaileek hitz bikoteak aurkeztuko zaizkie eta hauek semantikoki erlazionatuta dauden edo ez esan beharko dute botoi bat sakatuz.
  2. Ondoren, hitz bikoteak memoriatik berreskuratzeko eskatuko zaie.
- NEURKETAK: Erantzuna, erreakzio denbora eta pupilaren diametroaren neurketa zailtasun kognitiboaren adierazle bezala.
- BALIABIDEAK: E-Prime Software esperimentala eta Tobii X120 eye-tracker-a.

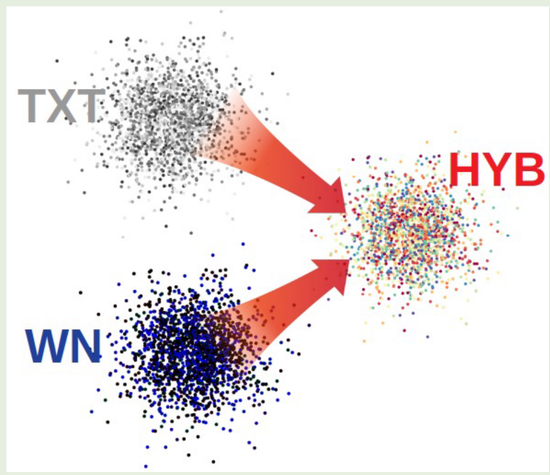
Material esperimentalen sorkuntza:

- Esperimentuaren material linguistikoa osatuko duten hitz bikoteak izenak izango dira; luzera, zehaztasun maila, auzokide fonologiko, auzokide semantiko eta maiztasunaren arabera parekatuak.
- Gainera, aurreko irizpideekin lortutako hitz bikoteen arteko antzekotasun semantikoa ere kalkulatuko da bektore hibridoen bidez [6]. Azken horiek kalkulatzeko, hurrengo pausuak jarraituko dira:
  a) Alde batetik, open Multilingual Wordnet bidez, gaztelerazko eta euskarazko pseudo corpusak sortu [7] eta beren errepresentazio distribuzionalak kalkulatu. Bestetik, bi ele horietan testu corpusen errepresentazioak lortu. 2. irudian adierazten duen legez, errepresentazio horiek Hipotesi Distribuzionala beteko dute.



2. Irudia: Testu corpus baten errepresentazio distribuzionalak bi dimentsiotan adierazita. Auzokideak partekatzen dituzten hitzak espazio semantikoan hurbil daude.

b) Ele bakoitzerako, bere WordNet espazioko (WN) eta testu corpuseko (TXT) hitzen errepresentazioak konbinatu [6][8], eta errepresentazio semantiko osoagoko espazio hibridoa (HYB) sortu (3. irudia).



3. Irudia: Ele baten testu corpuseko espazioa (TXT) eta Wordnetekoa (WN) conbinatuz gero, espazio hibridoa (HYB) sortzen da.

c) Lehenengo puntuko irizpideei jarraiki lortutako hitz pareen antzekotasuna kalkulatuko da bektore hibridoen bidez, eta, horren arabera eskuz aukeratuko dira esperimenturako datu multzo finala.

Erreferentziak:

[1] Brickman, et al. (2005). Category and letter verbal fluency across the adult lifespan: relationship to EEG theta power. *Archives of Clinical Neuropsychology, 20*, 561- 573.
[2] Collins & Loftus (1975). A spreading-activation theory of semantic processing. *Psychol. Review, 82*, 407-28.
[3] Kenett, et al. (2017). The semantic distance task: quantifying semantic distance with semantic network path length. *Journal of Psychology Learning, Memory and Cognition, 43*, 1470-1489.
[4] Boller, et al. (1991). Predictors of Decline in Alzheimer’s Disease. *Cortex, 27*, 9-17.
[5] Alladi, et al. (2013). Bilingualism delays age at onset of dementia, independent of education and immigration status. *Neurology, 81*, 1938-1944.
[6] Goikoetxea et al. (2016). Single or multiple? combining word representations independently learned from text and wordnet. In *Thirtieth AAAI Conference on Artificial Intelligence*.
[7] Goikoetxea et al. (2015). Randomwalks and neural network language models on knowledge bases. Proceedings the 2015 conference of the NACL-HLT, pages 1434–1439.
[8] García et al. (2020). A common semantic space for monolingual and cross-lingual meta-embeddings. arXiv preprint arXiv:2001.06381

Esker onak: BBVA Foundation