

Pakete lexikoen (lexical bundles) azterketa euskarazko corpus akademiko batean finkatutako errutina diskurtsiboak deskribatzeari begira

Proposer(s) / Proposatzailea(k)

Igone Zabala

Contact / Kontaktua:

igone.zabala@ehu.eus

Description / Deskribapena

Corpus akademikoak erabilienetakoak dira corpus-hizkuntzalaritzako lanetan aukera ugari ematen baitituzte erregistro horien bereizgarriak deskribatzeko askotariko aplikazioei begira: terminologia, fraseologia edota errutina diskurtsibo espezializatuak biltzen dituzten datu-baseak, edota erregistro espezializatuen irakaskuntzarako materialak prestatzeari begira (Biber, 1988, 1994, 2006; Parodi 2005, 2007, Flowerdew 2004) beste askoren artean). Edozein hizkuntzak komunikazio-tresna egokia izateko garatuta izan behar ditu askotariko erregistroak.

Corpusek gidatutako corpus akademikoetan gehien aztertutako elementu bat dira pakete lexikoak (lexical bundles) (Biber et al. 1999, Groom, 2005, Hayland 2008; Salazar 2011, Laso eta Salazar 2013). Izan ere, testu-corpusen azterketarako motibazio garrantzitsu bat izan zen ondo garatutako hizkuntzen hitzunez etengabe hitzen konbinazio berriak egin beharrean, hitzen pakete prefabrikatu ugari erabiltzen dituztela erdiautomatikoki. Euskara bezala garapenean duden hizkuntzetan, aldiz, ez da denborarik egon errutina horiek garatzeko eta finkatzeko. Erregistro akademikoen irakaskuntzan laguntzeko aplikazioak egiteari begira, funtsezkoa da honelako errutina diskurtsiboak deskribatzea eta aplikazioetan integratzea. Euskarak ere horrelako baliabideen deskripzioa eta aplikazioa behar ditu, baina zenbait zailtasun ditu deskripzioak hainbat bidetatik datozenak: euskara hizkuntza eranskaria eta hitz-hurrenkera librekoa izatea, erregistro akademikoak oraindik garapenean egotea eta hizkuntza guztietan geratzen den pakete lexiko bereizgarrien aldakortasun morfosintaktiko eta lexikoa (Zabala et al. 2018).

Goals / Helburuak

Lan honen helburua da corpus akademikoetan erabiltzen diren pakete lexikoen zerrenda lortzea, eta horietako bakoitzak diskurtso akademikoetan betetzen dituen funtzioen karakterizazioa egitea.

Requirements / Betebeharrak

Lan hau aurrera eramateko euskara jakitea ezinbesteko baldintza da, automatikoki lortuko diren datuak garbitzeko ezinbestekoa baita adibide zuzenak eta okerrak bereizi ahal izatea, bai eta horietako

bakoitzak diskurtsoan betetzen dituen funtzioak aztertzeke gauza izatea ere.

Framework / Esparrua

Corpus-hizkuntzalaritza: erregistro akademikoen garapena

Tasks and plan / Atzak eta plana

1. Garaterm corpusaren barruan azterketarako erabiliko diren azpicorpusak definitzea
2. N-gramak, maiztasunak, egile-kopurua eta agerraldiak erauzteke estrategiak eta baliabideak definitzea
3. Erauzitako n-gramen artean pakete lexikoak definitzeko parametroen (maiztasuna eta dispertsioa) balioak definitzea
4. Estatistikak eta agerraldiak lantzea ondorioak lortzeari begira