

# 4

## Estandar klinikoaren itzulpen automatikoa

**Olatz Perez-de-Viñaspre**

*Informatikan Doktorea. EHU/UPV*  
olatz.perezdevinaspre@ehu.es

**Maite Oronoz**

*Informatikan Doktorea. EHU/UPV*  
maite.oronoz@ehu.es

Sarrera-data: 2018-03-26 / Onartze-data: 2018-05-07

**Laburpena.** Lan honetan, osasun-zientzietan euskarak duen lehentasunezko beharretako bati heldu diogu: osasun-terminoen euskaratzea. Osasun-terminoak automatikoki euskaratu ditugu lau urratseko algoritmo bat erabiliz eta abiapuntutzat SNOMED CT estandar kliniko hartuz. Lehenengo urratsak baliabide lexikalak erabiltzen ditu SNOMED CTren terminoei euskarazko ordainak zuzenean esleitzeko. Bigarren urratsean, ingelesezko termino neoklasikoak euskaratzeko NeoTerm sistema garatu dugu. Sistema horrek, afixu neoklasikoaren baliokidetzak eta transliterazio-erregelak erabiltzen ditu osasun-termino baliokideak lortzeko. Hirugarrenean, ingelesezko termino konplexuak euskaratzen dituen KabiTerm sistema garatu dugu zeinak termino konplexuetan agertzen diren habiaratutako terminoen egiturak erabiltzen ditu euskarazko egiturak sortzeko. Azken urratsean, erregeletan oinarritzen den Matxin itzultzaile automatikoa osasun-zientzien domeinura egokitu dugu, MatxinMed sortuz. • **Hitz gakoak:** Estandar klinikoak, itzulpen automatikoa, terminologia kliniko euskaraz, SNOMED CT.

**Abstract.** In this work, we deal with one of the greatest priorities for Basque with regard to health sciences: obtaining a Basque Health Terminology. Health care terms of the SNOMED CT clinical standard have been automatically translated from English into Basque by using a four-step algorithm. The first step uses lexical resources to directly allocate SNOMED CT terms to Basque. In the second step, we developed the NeoTerm system for translating Neoclassical English terms. This system uses translations of neoclassical affixes and transliteration rules in order to come up with health term equivalents. In the third, we developed the KabiTerm system that translates complex terms by means of the structure of its nested terms. In the final step, the rule based automatic translator called Matxin, was adapted to the healthcare domain by creating MatxinMed. • **Key words:** Clinical standards, machine translation, clinical terminology in Basque, SNOMED CT

## 1. SARRERA ETA MOTIBAZIOA

**I** Euskal Autonomia Erkidegoan euskara eta gaztelania hizkuntzak dira ofizialak. Maiz, ordea, osasun-langileekin tratua euskaraz izaten dugun arren, haiek osasun-txostenak gaztelaniaz idazten dituzte. Euskal Herriko gainerako herrialdeetan egoera ez da hobe, eta erdarak dira osasun-txostenak idazteko hizkuntza nagusiak. Batzuetan arrazoi desberdinak medio, erosoago sentitzen direlako egiten dute hori (hezkuntzarako erabilitako hizkuntza, osasunarekin lotutako terminoen ezagutza eza...); besteetan, berriz, euskaraz osasunaren inguruan idazteko gai izan arren, beraien burua erdi behartuta sentitzen dute gaztelaniaz idaztera. Adibidez, Osakidetza sistema zentralizatua du eta osasun-langile batek idazten duen txostena, gaixoa tratatzen duten gainerako osasun-langileek ere irakur dezakete. Euskaraz idatzi duenaren ondoren datozen irakurleak euskara ulertzeko gai ez badira, gaixoaren segurtasuna kolokan egon daiteke.

Kanadan, elebitasuna ofiziala den zonaldeetan, gaixoaren eta osasun-langilearen arteko harremanerako hizkuntza gaixoa erabakitzen du (Desjardins, 2003) eta osasun-txostenak hizkuntza horretan idazten dira. Belgikan, aldiz, hizkuntza-komunitate bakoitzak bere osasun-zerbitzua kudeatzen duen sistema ez-zentralizatua dauka, eta Bruselaren kasuan, zonalde elebiduna izanik, bi osasun-zerbitzu sistema eskaintzen dira: frantziar komunitatearena eta flandiarrena (Gerkens eta Merkur, 2010). Luxenburgon, herrialde txiki bezain anitza izanik, nahiz eta alemana, frantsesa, italiarra, ingelesa eta portugesa hizkuntza zabalduak diren, *lingua franca* moduan frantsesa erabiltzen dute (European Observatory on Health Care Systems, 1999).

Gurean aldiz, hizkuntza komunitateak nahastuta daude eta *lingua franca* gisa gaztelania erabiltzen dugu Hego Euskal Herrian. Horrek, ordea, ez ditu herritar guztien hizkuntza-eskubideak bermatzen. Euskara hizkuntza ofiziala izan arren, osasun-langile guztiek ez dute euskara ezagutzen, eta euskaldunak diren osasun-langileek ahozko komunikazioa euskaraz egiten dute eta berehalako itzulpena egin beharra dute oharrak gaztelaniaz idazteko. Nolanahi ere, gaixoa oraindik ez du sistematikoki bere arreta euskaraz lortzeko aukerarik eta horretarako bidea ireki nahiko genuke.

Osasun-langileen euskaraz aritzeko ohiturak eta arazoak ezagutzeko asmoz, 2014 urteko otsailean osasun alorreko 45 langileri galdeketa bat pasatu genien honako inguruan: lan-harremanetan egiten duten eus-

---

**Kanadan, elebitasuna ofiziala den zonaldeetan, gaixoaren eta osasun-langilearen arteko harremanerako hizkuntza gaixoa erabakitzen du (Desjardins, 2003) eta osasun-txostenak hizkuntza horretan idazten dira.**

**Osasun-  
zientzietako  
kontzeptuak  
deskribatzeko  
euskarazko  
terminoak nahi  
baditugu,  
dagoeneko beste  
hizkuntzetarako  
sortutako  
sistema  
terminologiko  
bat oinarri  
hartzea iruditu  
zaigu egokiena.**

kararen erabilera ezagutu nahi genuen eta, osasun-txostenei dagokienez, txostenak edozein hizkuntzatan idazten dituztenean zein zailtasun topatzen dituzten ezagutu nahi genuen; eta txostenak euskaraz idazteko aukera izanez gero, tresnak zein behar ase beharko lituzkeen zerrendatu nahi genuen (ortografia-zuzentzailea, osasun-arloko terminologia euskaraz, txantiloia etab.).

Galdeketa emaitzetatik ateratako ondorioak zerrendatuko ditugu jarraian. Hizkuntzaren erabilerari dagokionez, 41 osasun-langilek gaitxoekin komunikazioa euskaraz egiten dute (beste 4ek ez dute erabiltzen) baina 5ek soilik idazten dituzte txostenak euskaraz. Galdetegian, osasun-txosten bat euskaraz idaztera gonbidatu genituen inkestatuak eta honako zailtasunak aurkitu zituzten: euskarazko terminologiaren falta (45etik 32 pertsonak), eredu falta (24 pertsonak) eta ohitura falta (15 pertsonak). Euskaraz idazteko beharrezkoak iruditzen zaizkien baliabideen inguruan galdetuta, lehentasun handiena eman zieten euskaraz osasun-zientzietako terminologia bateratu eta osatu bati (45etik 39 pertsonak), medikuntzara egokitutako ortografia-zuzentzaile bati (30 pertsonak) eta idazketarako eredu edo txantiloia izateari (25 pertsonak). Galdetutako % 89 prest agertu zen tresna sortuko balitz erabiltzeko; gainerakoek, berriz, bere zalantzak agertu zituzten.

Beraz, inkestaren emaitzetatik ondoriozta dezakegu osasun-zientzietako terminologia bateratu eta osatu bat euskaraz izatea zein garrantzitsua den, gaixo eta osasun-langileen artean euskarazko harremana normalizatzeko, baita osasunaren munduan egunerokoan euskararen erabilera sustatzeko ere. Hori izan dugu lan honen motibazio nagusia: *osasun-zientzietako terminologia euskaraz izateko urratsak ematea, osasunaren alorrean euskararen normalizazioan aurrera pauso bat egiteko*. Artikulu honetan deskribatzen den lanak, Olatz Perez-de-Viñasprek Maite Oronozen zuzendaritzapean garatutako “Osasun-alorreko termino-sorkuntza automatikoa: SNOMED CTren eduki terminologikoaren euskaratzea” tesi-proiektuaren zati bat laburtzen du<sup>1</sup>.

Osasun-zientzietako kontzeptuak deskribatzeko euskarazko terminoak nahi baditugu, dagoeneko beste hizkuntzetarako sortutako sistema terminologiko bat oinarri hartzea iruditu zaigu egokiena. Sistema horien errebaso bat egin ondoren, *Systematized Nomenclature Of Medicine – Clinical Terms* edo SNOMED CT estandar kliniko euskarara ekartzea erabaki genuen. SNOMED CT da ingelesez, espainieraz eta beste hainbat hizkuntzatan, osasun alorreko terminologia jasotzen duen baliabide zabalduena, eta, era berean, orain arte sortu den terminologia eleaniztun

osatuentzat hartzen da. Nolabait esateko, hizkuntza eta sistema desberdinen arteko osasun-txostenen adierazpen eta interpretazio automatikoa eta anbiguotasunik gabea ahalbidetuko duen hiztegi normalizatua da, eta hiztegi-sarreren arteko harremanak zehaztuta ditu. Osasun-txostenetan agertzen diren kontzeptuak, deskribapenak eta erlazioak barnebiltzen ditu. Ingelesa lantzeko sortu zen hasieran, eta 300.000 kontzeptu baino gehiago ditu definituak, baita horiek izendatzeko ingelesezko 1.000.000 termino baino gehiago ere.

SNOMED CTren indargune nagusia estaldura da. Domeinu klinikoan erabiltzen diren kontzeptuak barnebiltzen ditu, eta espezialitate gehienetako terminologia ere jasotzen du. Horretaz gain, kontzeptuen artean harremanak zehazten ditu: alde batetik, egitura hierarkikoa ematen dioten harremanak ditu, eta, beste aldetik, informazio semantikoa gehitzen duten harremanak, hala nola, agente eragilea (*causative-agent*), aurkikuntzaren tokia (*finding site*), etab. SNOMED CTren barruko harremanez gain, gaur egunean beste baliabide lexikal askoren kontzeptuekin parekatu dituzte SNOMED CTren kontzeptuak. Izan ere, *Unified Medical Language Systemen* (UMLS) metatesauroaren parte da, eta horri esker UMLSren metatesauro horretan dauden gainerako baliabide terminologikoekin parekatuta dago. Parekatze horri esker, SNOMED CT euskaraz izango bagenu, baliabide terminologiko horiek erabiltzen dituzten teknologiak eskura izango genituzke euskararentzat ere.

Hurrengo atalean, SNOMED CTren euskarazko bertsioa automatikoki sortzeko diseinatu eta inplementatu dugun algoritmoa eta honen urrats nagusiak aurkeztuko ditugu.

## 2. OSASUN TERMINOEN SORKUNTZA AUTOMATIKOA

### I

### 2.1 Itzulpenerako algoritmo orokorra

EuSnomed, SNOMED CT sare semantikoa euskaratzeko garatu dugun sistema da eta lau urratsetako algoritmo bat dauka oinarrian. Algoritmo horrek jatorri-helburu hizkuntzetarako eskura dauden baliabide lexiko elebidunak berrerabiltzen ditu eta horien hutsuneak betetzeko baliabide berriak sortzen ditu. Ordainak modu inkrementalean lortzen ditugu, lortutako ordain berriak berrerabiliz.

Honakoak dira algoritmoaren lau urratsak:

- **Baliabide lexikaletatik jasotako terminoak:** osasun-zientzien domeinuko hainbat hiztegi elebidunetatik zein eleaniztunetatik

---

**SNOMED CT sare semantikoa euskaratzeko garatu dugun sistema da eta lau urratsetako algoritmo bat dauka oinarrian. Algoritmo horrek jatorri-helburu hizkuntzetarako eskura dauden baliabide lexiko elebidunak berrerabiltzen ditu eta horien hutsuneak betetzeko baliabide berriak sortzen ditu.**

erauzitako pareak erabili ditugu. Gure beharretarako, ingelesa-euskara zein espainiera-euskara hiztegi eleaniztunak erabili ditugu, hala nola, Euskalterm, ZT Hiztegia, eta abar. Adibidez, 1. irudian urrats honetan kokatzen den adibide bat ikus dezakegu. “*Deoxyribonucleic acid*” ingelesezko jatorri-terminoarentzat bi ordain jasotzen ditugu hiztegietatik, “azido desoxirribonukleiko” eta “DNA”. Jatorria jasotzeaz gain (ikus irudia), datu gehiago sortu eta gordetzen ditugu parekatzean, iturriaren konfiantza maila, kasua, eta abar.

**1. irudia: Baliabide lexikalek emandako ordainen adibidea.**

<b>Jatorri-terminoa:</b>	<i>Deoxyribonucleic acid</i>
<b>Ordainak:</b>	azido desoxirribonukleiko (ZT Hiztegia, Euskalterm) DNA (ZT Hiztegia, Euskalterm)

— **Termino neoklasikoak:** urrats honetan termino neoklasikoen afixu semantikotaz baliatzen gara ordainak sortzeko. Hurrengo atalean, urrats honetarako garatu dugun NeoTerm sistema sakonki aurkeztu badugu ere, 2. irudian honen adibide bat erakusten dugu. “*Photodermatitis*” ingelesezko jatorri-terminoaren euskaratze-prozesua urratsez urrats erakusten dugu.

**2. irudia: Termino neoklasikoen sorkuntza-erregelak lortutako ordainen adibidea.**

<b>Jatorri-terminoa:</b>	<i>Photodermatitis</i>
<b>Sorkuntza prozesua:</b>	
<u>Identifikatutako afixuak:</u>	photo+dermat+itis
<u>Itzulitako afixuak:</u>	foto +dermat+itis
<b>Sortutako ordaina:</b>	fotodermatitis

— **Termino habiaratuak:** termino konplexuetarako (hitz batez baino gehiagoz osatutako terminoak) termino habiaratuetan oinarritutako sorkuntza-erregelak erabili ditugu. Termino habiaratuak termino baten barruan dauden beste termino batzuk dira. Adibidez, *fracture of elbow* terminoak, bi termino habiaratu ditu: *fracture* eta *elbow*. Kasu honetan, *fracture* terminoa gaixotasun bat da SNOMED CTn eta *elbow* gorputz-egitura bat da. Informazio horretaz baliatu gara sorkuntza-erregelak idazteko. Aparteko programa bat sortu dugu urrats honetarako ere, KabiTerm deiturikoa. Aurreko adibidearekin jarraituz, 3. irudian termino horren euskaratzea erakusten dugu urratsez urrats.

**3. irudia: Termino habiaratuen sorkuntza-erregelen lortutako ordainen adibidea.**

**Jatorri-terminoa:** *Fracture of elbow*  
**Sorkuntza prozesua:**  
**Identifikatutako egitura:** NAHASMENDU + of + GORPUTZ\_EGITURA  
**Egitura baliokidea:** GORPUTZ\_EGITURA+ren + NAHASMENDU  
**Sortutako ordaina:** Ukondoaren haustura

— **Itzultzaile automatikoak sortutako termino konplexuak:** azkeneko, aurreko urratsetan ordainak lortu ez ditugun terminoetarako itzultzaile automatiko bat egokitu dugu. Horretarako, Matxin erabilera orokorreko itzultzaile automatikoa erabili dugu. Medikuntzaren domeinura egokitu dugu eta MatxinMed deitu diogu egokitzapen honi. MatxinMed gainerako urratsetan sortu ditugun termino-ordain pare berriekin elikatu dugu.

**4. irudia: MatxinMeden bidez lortutako ordainen adibidea.**

**Jatorri-terminoa:** *Lymphoma of lower esophagus*  
**Sortutako ordaina:** Beheko esofagoko linfoma

Hurrengo lerroetan, NeoTerm eta KabiTerm programen azalpenean sakonduko dugu, izan ere bi programa horiek aurkezten dute ikuspegi linguistikotik interes handiena. Alderdi informatikotik transduktoreak erabiltzen dituzte bi programek. Transduktoreak sarrera bat hartu eta aldatzen duten erregelak gauzatzeko software-makinak dira.

## **2.2 Termino neoklasikoaren euskaratzea: NeoTerm**

NeoTerm ingelesezko termino neoklasiko bat jaso eta euskal ordainak sortzen dituen programa da eta ezagutza adierazten duten erregelatan du oinarria. Termino neoklasikoak, jatorri grekoa edo latindarra duten morfemez osatuta dauden terminoak dira, adibidez hipogluzemia edo fotodermatitis. Termino neoklasikoak oso errotuta daude medikuntzaren alorrean.

McCray et al. (1988) lanean konposatu neoklasikoak bi morfema motetan sailkatzen dituzte: erroak eta amaierakoak. Horiekin batera, aurrizki (*non-*, adibidez *nononcogenic* terminoan) zein atzizki (*-al*, adibidez *nocomial* terminoan) orokorrak ere erabili dituzte azterketan. Literaturan ez dago koherentziarik izendatze horien gainean, eta erroei aurrizki ere deitzen diete. Gure lanean, aurrerago ikusiko dugun bezala, erroak eta aurrizkiak elkarrekin landuko ditugu, aurrizki terminoa erabiliz beraiek

**NeoTerm**  
**ingelesezko**  
**termino**  
**neoklasiko bat**  
**jaso eta euskal**  
**ordainak sortzen**  
**dituen programa**  
**da eta ezagutza**  
**adierazten duten**  
**erregelatan du**  
**oinarria.**

izendatzeko, zabalduagoa dagoelakoan. Amaierako morfemen kasuan, atzizki izendatuko ditugu.

Euskarazko ordainak sortzeko, hiru urrats inplementatu ditugu NeoTermen:

1. Afixuen identifikazioa: jatorri-terminoaren (ingelesezkoak kasu honetan) afixu neoklasikoaren identifikazioa, eta horien araberrako zatitzea.
2. Afixuen baliokidetzak: jatorri-terminoan identifikatutako afixuen baliokidetzak lortzea helburu hizkuntzan (euskara, gure kasuan).
3. Afixuen konposaketa: euskararen arau morfofonologikoak errespetatuz afixu baliokideen konposaketa sortzea.

Hurrengo irudian (5. irudia) *hypophosphatemia* terminoa NeoTermi esker nola euskaratzen dugun erakusten dugu urratsez urrats. Kontuan izan, afixu guztiak ez direla lexikoietan agertzen, eta transliterazioa behar dela termino hau euskaratzeko.

**5. irudia: NeoTermen oinarri-sistemaren sorkuntza prozesua.**

**Jatorri-terminoa:** *hypophosphatemia*  
**Identifikatutako afixuak:** *hypo+phos+ph#+at+emia*  
**Afixu baliokideak:** hipo+fos+f+at+emia  
**Euskal ordaina:** hipofosfatemia

Adibidean, lehenik, afixuak identifikatzen ditugu, identifikatu gabeko karaktere kopurua minimizatuz eta afixu kopuru minimoa aukeratuz (*hypo+phos+ph#+at+emia*). Ostean, afixuen baliokideak jasotzen ditugu lexikoietatik eta identifikatu gabeko zatiak transliteratzen ditugu (hipo+fos+f+at+emia). Bukatzeko, afixuak konposatzen ditugu (hipofosfatemia). Afixuen identifikaziorako, ezinbestekoa da lehenik hauek zerrendatzea eta euskarazko ordainak lortzea.

*Afixuen hiztegien sorkuntza*

Jatorri-terminoaren afixuen identifikaziorako, ingelesezko medikuntzako afixuen hainbat zerrenda eskuratu ditugu 1.703 aurrizki eta 630 atzizki lortuz: Stedman-en medikuntza hiztegiko “*Medical Prefixes, Suffixes, and Combining Forms*” (Stedman, 2005), Wikipediaren “*List of medical roots, suffixes and prefixes*” (Wikipedia, 2013), McCarthy-ren “*Suffix Prefix Dictionary*” (McCarthy, 2016) hiztegia eta, garapenean sortutako zalantzazko kasuetan, “*Mosby’s Medical Dictionary*” hiztegia erreferentzia hartuta hainbat aurrizki gehitu ditugu.

Ordainen sorkuntzarako zerrenda elebarkar horietako afixuak banan-banan aztertu ditugu euskarazko baliokidetzak aurkitzeko. Horretarako, afixuen itzulpenak hiztegi espezializatueta termino-ordain pareetatik ondorioztatu ditugu. Adibide gisa, 1. taulan *encephal-* aurrizkiaren euskal baliokidea ondorioztatzeke prozesua erakusten dugu.

1. taula: *encephal-* aurrizkiaren euskal ordainak.

Ingeleseke terminoak	Euskarazko ordainak
anencephalia	anentzefalia
echoencephalogram	ekoentzefalograma
electroencephalograph	elektroentzefalografo
encephalitis	entzefalitis
encephalomyelitis	entzefalomielitis
encephalopathy	entzefalopatia
leukoencephalitis	leukoentzefalitis
...	...

Ikus daitekeenez, hiztegi sarrera guztietan ingeleseke *encephal-* aurrizkia, euskarazko ordainetan entzefal- aurrizki moduan agertzen da. Prozesu hori bera egin dugu zerrendetako 2.333 afixuekin (lortu ditugun afixuen baliokidetzak <http://ixa2.si.ehu.eus/neoterm/> web-orrian daude eskuragarri).

Sortutako hiztegi elebidunetik, bi lexikoi bereizi ditugu, bata aurrizkiarako eta bestea atzizkiarako.

### Identifikazioa

Afixuen identifikazioa bi urratsetan egiten da. Lehenengo urratsetan identifikatzea lortzen ez bada, bigarren urratsetan identifikatzen saiatuko da. Bigarren urratsetan ere ezin bada identifikatu, terminoa euskaratu gabe gelditzen da.

Lehenengo urratsetan, nahi adina aurrizki eta derrigorrezko atzizki bakar bat bilatzen dugu terminoan. Adibidez, ingeleseke *radionecrosis* terminoarentzat, *radio+necr+osis* edo *radio+nekro+si* zatiak identifikatuak genituzke (aurrizki eta atzizki guztiak).

Atzizkia derrigorrezkoa egitea erabaki dugu, atzizkien baliokidetzak aurrizkienak baino konplexuagoak edo ezberdinak direla ikusi dugu-

**Hiztegi sarrera guztietan ingeleseke *encephal-* aurrizkia, euskarazko ordainetan entzefal- aurrizki moduan agertzen da. Prozesu hori bera egin dugu zerrendetako 2.333 afixuekin.**



**KabiTerm,  
termino  
konplexuen  
barruan agertzen  
diren beste  
terminoetan  
oinarritzen da  
termino  
konplexuak  
euskaratzeko.**

lako. Hau da, aurrizki gehienek transliterazio arauak jarraitzen dituzten bitartean, atzizkiek aldaera ezberdinak dituzte (ingelesezko *-itis*, euskaraz *-itis* mantentzen den bitartean, *-sis* atzizkia *-si* gisa euskaratzen da).

Ikusi dugunez, afixuen identifikazioan zenbaitetan anbiguotasuna egoten da, eta hainbat aukera proposatzen ditu NeoTermek. Anbiguotasuna gutxitzeko, afixu gutxien (edo luzeenak) dituzten aukerak erabili ditugu. Adibidez, ikus dezagun *photodermatitis* terminorako transduktoreak proposatzen dituen afixuen konbinazio ezberdinak:

- *photo+dermat+itis*: 3
- *photo+derm+at+itis*: 4
- *phot+o+dermat+itis*: 4
- *phot+o+derm+at+itis*: 5

Terminoak, 3, 4 edo 5 afixutan bana daitekeela ikus dezakegu. Zehaztutako irizpideari jarraiki, lehenengo aukera izango da aukeratua afixuen baliokidetzak lortzeko eta euskal ordaina sortzeko.

Bigarren urratsean, identifikazio malguagoa definitu dugu, aurrizki eta atzizki zerrendetan aurkitzen ez diren zatiak ere identifikatuz. Horretako definituko dugun gutxieneko baldintza, terminoaren atzizkia lexikoian agertzea da, nahiz eta aurrizki guztiak lexikoian ez egon. Malgutasun honekin anbiguotasuna sortzen da eta hori ebazteko, jarraian azalduko dugu irtenbide bat.

*Diverticulitis* terminoaren kasuan, *verticul* zatia ez da afixuen lexikoian agertzen (zati hauek # ikurra dute amaieran). Hala ere, transduktoreak proposatu egiten ditu zatiketak lexikoetan agertzen diren afixuak identifikatu gabeko zatiekin konbinatu egiten baititu. Identifikatutako zatien ondoren, zati kopurua eta identifikatu gabeko zatien karaktere kopurua gehitu ditugu.

- *diverticul#+itis*:  $2 + 10 = 12$
- *divertic#+ul+itis*:  $3 + 8 = 11$
- *di+verticul#+itis*:  $3 + 8 = 11$
- *di+vertic#+ul+itis*:  $4 + 6 = 10$

Aukeraketa egiteko orduan, bi baldintza erabili ditugu: luzeran motzena den identifikatu gabeko zatia bilatzen dugu, eta aldi berean, lexikoietan agertzen diren afixu gutxien egotea bilatzen dugu (*di+vertic#+ul+itis*, kasu honetan). Identifikatu gabeko zatia transliterazioa erabiliz euskaratzen da.

### *Baliokidetza*

Baliokidetza, identifikatzeko moduaren arabera egiten da: lexikoen bidez ala transliterazioaren bidez.

Afixu bezala identifikatu gabeko zatiak transliteratzeko 40 erregela definitu ditugu. Erregela horiek sortzeko, euskararen morfofonologia arauak errespetatzeko erregelak transliteraziorako egokitu ditugu (ikus konposaketa atala) eta berri batzuk definitu ditugu ere. Berriak definitzeko, erregela batzuk enpirikoki ondorioztatu ditugu, euskararen ortografia arauak betetzen dituztela bermatuz. Adibidez, *v* letra, *b* letra-rekin ordezkatzeko dugun beti arau ortografikoari jarraituz, edo *c* letra *z*-rekin *e*, *i* edo *y* letra jarraian izanez gero, eta *k*-rekin gainontzekoetan, arau morfofonologikoen ezartzen duten moduan.

Erregela horietako batzuk adibideekin erakusten ditugu 2. taulan.

**2. taula: Transliteraziorako hainbat erregela.**

	en	eu	Adibidea	Erregela
1	ÿ	i	yttrium -> itrio	y -> i    [#.l ʔ] _ Kon
2	ph	f	pharyngostome -> faringostoma	ph -> f
3	rh	rr	rheostosis -> erreostosis	rh -> rr
4	ck	k	amsinckine -> antsiskina	ck -> K
5	rs	rts	arsenic -> artseniko	r -> r t l l _ Txis Bok
6	xs	s	isoxsuprine -> isosuprina	x -> 0 l l _ s
7	cc	kz	occipital -> okzipital	cc -> k z l l _ [ e l i l y ]
...	...	...	...	...

### *Konposaketa*

Afixuen konposaketarako, sortutako ordainek euskararen arau ortografikoak eta morfofonologikoak betetzeaz arduratzen diren 28 erregela definitu ditugu. Arau ortografikoen artean *v* bezalako letren ordezkapenak aurkitzen ditugun bitartean, arau morfofonologikoetan hobikarien eraginez txistukarien sabaikaritzea aurkitzen dugu (euskaraz, entzefalitis esango genuke, eta ez \*enzefalitis).

Erregela horiek definitzeko, alde batetik Euskaltzaindiaren ortografia arauak izan ditugu kontuan (Euskaltzaindia. Luis Mitxelena, 1968), eta bestetik enpirikoki aztertutako aldaerak.

**Afixuen konposaketarako, sortutako ordainek euskararen arau ortografikoak eta morfofonologikoak betetzeaz arduratzen diren 28 erregela definitu ditugu.**

**KabiTermek ingeleseko termino konplexu bat jasotzen du sarreran, eta, baliabideak izanez gero, euskarazko ordainak proposatzen ditu. Baliabide horiek, habiaratutako terminoen ordainak eta euskaratze-patroiak dira.**

Aurreko adibideari jarraiki, afixuen konposaketak sortutako aldaeren artean daude gaurko erdaretan *r* letraz hasten diren hitzak (3. taulan 1. erregela). Hitz horiek euskaraz mailegatzean, bokalez hasten dira. Kasu gehienetan *e* letra erabiltzen da horretarako (erradio, errezeta...) nahiz eta kasu bakan batzuetan *a* letra ere erabili izan den (arropa, arrosa...).

Hurrengo taulan (3. taula) erregela gehiago ikus ditzakegu.

**3. taula: Erregela morfofonologiko batzuk.**

	en	eu	Adibidea	Erregela
1	ˈr	e r r	radionecrosis -> erradionekrosi	r -> err ll [.#. l ] _
2	ˈs	e s	spermatocoele -> espermatozele	s -> es ll [.#. l ] _ (+) Kon
3	s s	s	dysomnia -> disomnia	s -> 0 ll _ + s
4	n z	n t z	encephalitis -> entzefalitis	n -> n t ll _ Txis + Bok
5	r s	r e s	hypersplenism -> hiperesplenismo	r + -> r e ll _ Txis Kon
6	r r	r	hyperreflexia -> hiperreflexia	r r -> r ll Kon + _
7	m p	n p	symphysis -> sinfisis	m -> n ll _ + [ b l p l t l f ]
8	k	z	thoracic -> toraziko	
9	c c	k	buccolabial -> bukolabial	K Z -> k ll _ + [Kon l a l o l u]
...	...	...	...	...

### 2.3 Termino habiaratuen euskaratzea: KabiTerm

KabiTerm termino konplexuen barruan agertzen diren beste terminoen oinarritzen da termino konplexuak euskaratzeko.

KabiTermek ingeleseko termino konplexu bat jasotzen du sarreran, eta baliabideak izanez gero, euskarazko ordainak proposatzen ditu. Baliabide horiek, habiaratutako terminoen ordainak eta euskaratze-patroiak dira. Aurrerago azalduko dugun moduan, KabiTermen lana errazteko, AnaMed deituriko analizatzailea prestatu dugu. Analizatzaile hori, KabiTermek beharrezkoa duen informazioa biltzeaz arduratzen da, eta termino habiaratuak identifikatzen eta prestatzen ditu, KabiTerm euskaratzeaz soilik ardura dadin.

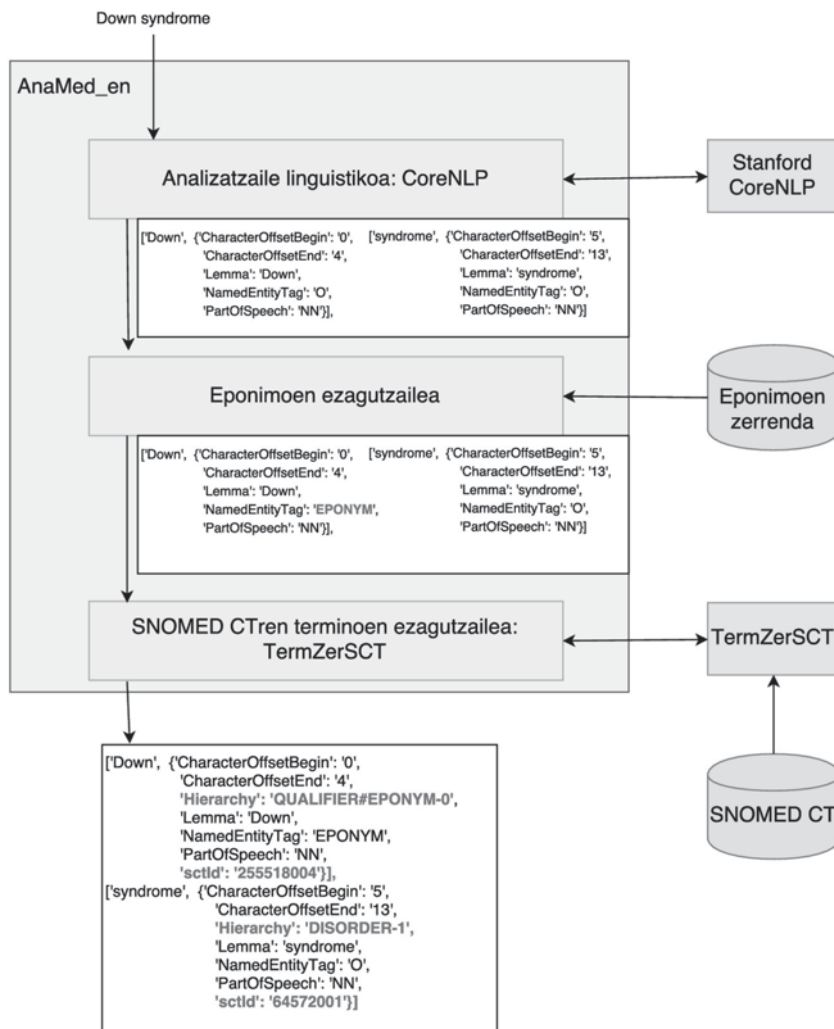
Orokorrean, AnaMedek analizatzaile linguistiko automatiko bat hartzen du oinarritzat, eta eponimoen eta SNOMED CTren terminoen identifikazioa integratzen ditu. Bere arkitektura 6. irudian ikus dezakegu. Irudian ikusten den “Down syndrome” adibidearekin ikusiko dugu AnaMeden funtzionamendua.

1. **Analizatzaile linguistikoak** hitz bakoitzaren lema (*lemma*), kategoria gramatikala (*PartOfSpeech*), entitate mota (pertsonea, tokia ala erakundea den, *NamedEntityTag*) eta hitzaren posizioa (*CharacterOffsetBegin* eta *CharacterOffsetEnd*) ematen dizkigu.
2. Bigarrenik, **eponimoen ezagutzaileak** kontzeptuen izendatzetan agertzen diren pertsona-izenak identifikatzen ditu. Izan ere, medikuntzan eponimoak oso ohikoak dira, bereziki gaixotasunak eta sindromeak adierazteko. Irudiko “Down-en sindromea” edo “Alzheimer-en gaixotasuna” terminoetakoak adibidez. Irudian ikusten dugun bezala, “Down” hitzaren *NamedEntityTag* etiketa eguneratua izan da *EPONYM* balioarekin.
3. Azkenik, **SNOMED CTren terminoen ezagutzaileak**, pasatako terminoaren baitan agertzen diren beste SNOMED CTren terminoak identifikatzen ditu, eta euri dagozkien hierarkia (gaixotasuna, gorputz-egitura, prozedura etab. ote den) eta SNOMED CTko kontzeptu identifikadorea (zenbaki bat) etiketatzen ditu (*Hierarchy* eta *sctId* etiketak hurrenez hurren). Irudian ikus daitekeenez, Down hitzari, kalifikatzaile hierarkia (*qualifier*) eta 255518004 kontzeptu identifikadorea esleitu dizkio, eta *syndrome* hitzari aldiz, nahasmenduen hierarkia (*disorder*) eta 64572001 identifikadorea.

---

**SNOMED CTren terminoen ezagutzaileak, pasatako terminoaren baitan agertzen diren beste SNOMED CTren terminoak identifikatzen ditu, eta euri dagozkien hierarkia (gaixotasuna, gorputz-egitura, prozedura etab. ote den) eta SNOMED CTko kontzeptu identifikadorea (zenbaki bat) etiketatzen ditu.**

6. irudia: AnaMed analizatzailearen arkitektura.



AnaMedek sortutako informazioa KabiTerm sistemak erabiltzen du termino nagusian termino habiaratuak identifikatzeko. Termino habiaratu horiek oinarri hartuta, ingelesezko terminoen egituren analisisa egin dugu AnaMeden bidez. Hurrengo taulan (4. taula), egitura horien adibide batzuk ikus ditzakegu. Adibidez, *malignant neoplasm of renal calyx* termino konplexuan bi termino nagusi aurkitu ditugu *malignant neoplasm* nahasmenduen (*disorder*) hierarkiakoa bata eta *renal calyx* gorputz-egituren (*body structure*) hierarkiakoa bestea. Aipatu beharra dugu, habiaratutako termino gehiago ere aurkitzen dituela AnaMedek, hala nola *calyx* gorputz-egitura, *malignant* kalifikatzailea (*qualifier*) edo *neoplasm* nahasmendua.

**4. taula: AnaMeden bidez lortutako egiturak eta multzokatzeak.**

Terminoak	Multzokatzea	Egitura
structure of radial tuberosity	structure of radial_tuberosity	structure+of+BODYSTR
Baelz's disease	Baelz's disease	EPONYM+'s'+DISORDER
malignant neoplasm of renal calyx	malignant_neoplasm of renal_calix	DISORDER+of+BODYSTR

Habiaratutako termino guztiakin sortu ditugu egiturak, eta agerpen kopuruaren eta dependentsia kopuruaren arabera sailkatu ditugu. Hau da, egitura zenbat terminotan azaltzen den zenbatu dugu, eta termino hori zenbat terminotan agertzen den habiaratuta kontuan hartu dugu. Egitura horietako gutxi batzuk ikus daitezke 5. taulan (“Agerp.” izenburuak agerpen kopuruari egiten dio erreferentzia eta “Depen.” izenburuak egitura horrek beste terminoekiko duen dependentsia kopuruari). Adibidez, QUALIFIER+DISORDER egiturarako, 4.469 agerpen aurkitu ditugu, eta gainera 74.208 terminoetan agertzen da egitura hau habiaratuta. Lehen-tasun altuko egitura dela erabaki dugu, agerpen asko izateaz gain, gainerako termino konplexuak euskaratu ahal izateko oso garrantzitsua izango delako. QUALIFIER+neoplasm egitura aldiz, bere horretan 4 aldiz agertzen bada ere, beste termino konplexuek lau termino horiekiko duten dependentsia oso altua da (28.642 terminoen barruan agertzen da).

**5. taula: SNOMED CTren terminoen egituren agerpenak eta dependentsiak beste terminoekiko.**

Egitura	Adibidea	Agerp.	Depen.
QUALIFIER+DISORDER	unstable diabetes mellitus	4.469	74.208
QUALIFIER+neoplasm	malignant neoplasm	4	28.642
PROCEDURE+of+BODYSTRUCTURE	amputation of finger	5.082	33.181
...	...	...	...

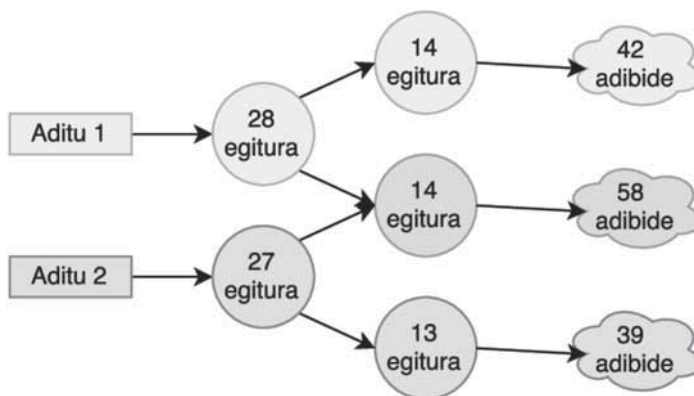
Bi medikuri egitura horietako adibideak eman dizkiegu, horiei euskarazko ordainak emateko. Izan ere, Zabala et al. (2012) lanean erakusten duten bezala, hiztegi zein itzultzaileek jatorrizko terminoaren egituraren kalkoetara jotzeko joera duten bitartean, adituek euskararen berezko egituretara jo ohi dute. Hori horrela izanik, medikuek sortutako lagina hobetsi dugu hiztegietatik erauzitako adibideak baino. Adituek sortutako adibide horiek euskaratze-patroiak definitzeko oinarria izan dira. Adibideak lortzeko prozesua bi fasetan banatu dugu.

Lehenengo faserako 41 egitura aukeratu ditugu eta bakoitzeko gutxienez zoriz erauzitako 3 adibide. Aditu bakoitzari 28 eta 27 egitura eman diz-

**Hiztegi zein itzultzaileek jatorrizko terminoaren egituraren kalkoetara jotzeko joera duten bitartean, adituek euskararen berezko egituretara jo ohi dute. Hori horrela izanik, medikuek sortutako lagina hobetsi dugu hiztegietatik erauzitako adibideak baino.**

kiogu hurrenez hurren, zeinetatik 14 egitura komunak diren, 7. irudian ikus dezakegun moduan. Denera 100 eta 97 adibideren euskarazko ordainak jaso ditugu, euretako 58 termino komunean dituztelarik.

**7. irudia: Adituei banatutako lagina.**



Komunean dituzten adibideen kasuan, adostasun altua neurtu dugu, eta desadostasun kasuekin, beraien arteko adostasuna bilatu dugu, irizpide berdinak erabil ditzaten. Hurrengo taulan (6. taula) adibide horietako batzuk ikus ditzakegu, non lehenengo adibidean ordain ezberdinak proposatu dituzten (nahiz eta gerora adostasuna lortu den), bigarreanean adostasuna izan duten, eta hirugarren eta laugarrenetan adibide ezberdinak dituzten. Adibideetan ikus daitekeenez adibide-terminoak ez dira sinpleak eta osasunari buruzko ezagutza behar da euskaratze egokia lortzeko.

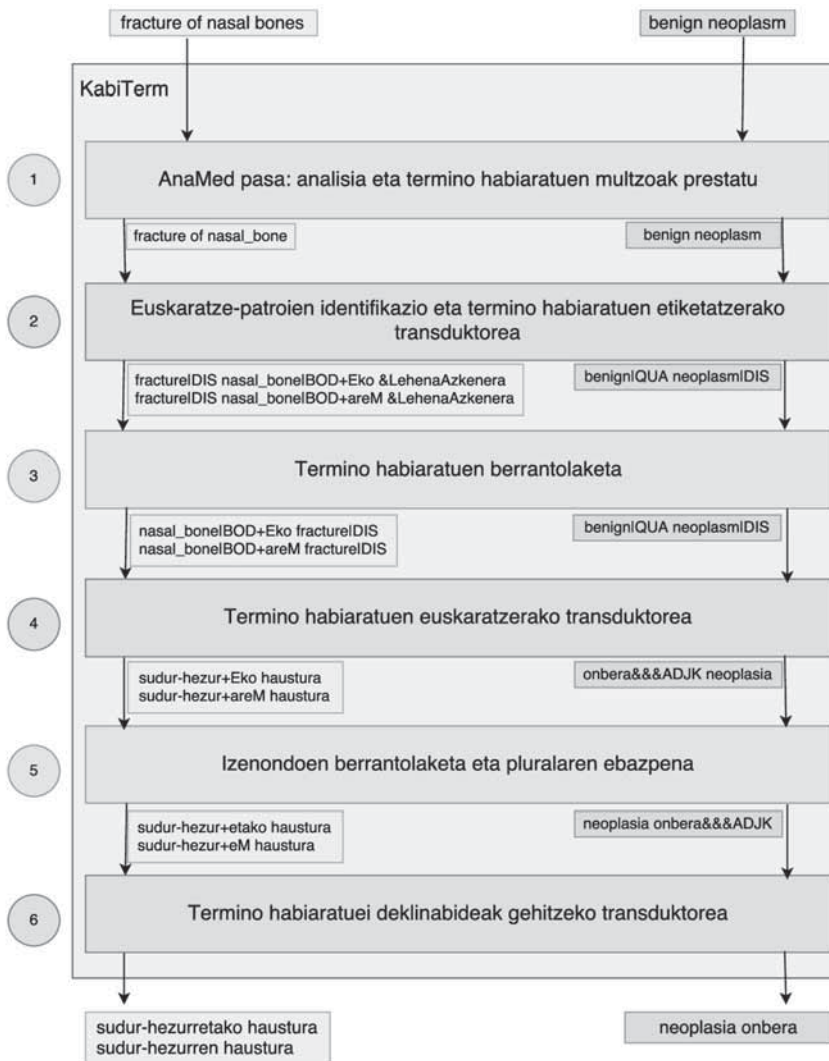
**6. taula: Adituen euskaratzeen adibide batzuk.**

Ingelesa	Aditu 1	Aditu 2
cryotherapy to cranial nerve	nerbio kranialaren krioterapia	garezurreko nerbioen krioterapia
calcium regulating agent overdose	kaltzioaren agente erregulatzailiek eragindako gaindosia	kaltzioaren agente erregulatzailiek eragindako gaindosia
open fracture of scaphoid bone of wrist	eskumuturreko eskafoide hezuraren haustura irekia	
adrenergic neurone blocking drug adverse reaction		neurona adrenergikoen blokeatzaileek eragindako kontrako efektua

Oinarrizko irizpideak argi izanik, bigarren faserako, 25na egitura prestatu ditugu, aditu bakoitzak denera 100 adibide inguru dituztelarik. Bi faseak elkarturik, 340 adibide inguru izan ditugu euskaratze-patroiak definitzeko eta horietatik abiatuta 53 euskaratze-patroi definitu ditugu.

Jarraian, besteak beste KabiTermek egiten duen patroi horien erabilera aztertuko dugu. Kabiterm tresnak urratsez urratseko aldaketak egiteko, transduktoreak erabiltzen ditu eta bere funtzionamendua 8. irudian irudikatu dugu:

**8. irudia: KabiTermen arkitektua eta funtzionamenduaren adibideak.**



1. Aurrena, sarrera terminoa AnaMedek analizatzen du, termino habiaratuak identifikatuz eta multzokatuz: *fracture of nasal bones* terminoaren kasuan, *fracture* nahasmendua da eta *nasal\_bone* gorputz-egitura. Termino habiaratuak multzokatzeaz gain, lematizazio ere beharrezkoa izan da adibide honetan, *nasal bone* SNOMED CTren terminoa delako, baina ez *nasal bones* forma pluralean.



**Euskaratze-  
patroien  
identifikazio eta  
termino  
habiaratuen  
etiketatzerako  
transduktoreari  
deitzen diogu, eta  
honek, dagokion  
euskaratze-  
patroia  
aplikatuz,  
termino  
habiaratuei  
euskaratzeko  
beharrezko  
etiketak gehitzen  
dizkie.**

2. Bigarrenik, euskaratze-patroien identifikazio eta termino habiaratuen etiketatzerako transduktoreari deitzen diogu, eta honek, dagokion euskaratze-patroia aplikatuz, termino habiaratuei euskaratzeko beharrezko etiketak gehitzen dizkie. Kasu honetan, DISORDER+of+BODY-STRUCTURE egitura identifikatu du, eta horri dagokion euskaratze-patroia aplikatu dio: *fracture* terminoari “|DIS” etiketa gehitu dio, nahasmendu bat delako, *nasal\_bone* terminoari aldiz, “|BOD+Eko” etiketa eta “|BOD+areM” etiketak gehitu dizkio, gorputz-egitura izateaz gain, deklinabidea ere gehitu behar izan diolako (“+Eko” eta “+areM” adibidean). Horri guztiari, ordena aldatzeko etiketa ere gehitu dio, lehenengo terminoa amaieran jartzeko (“&LehenaAzkenera”).

3. Jarraian, termino habiaratuen berrantolaketa egiten dugu, aurreko urratsak sortutako etiketari jarraituz (&LehenaAzkenera). Horrela, *fracture* terminoa lehenetik azkenera pasatzen da.

4. Laugarrenik, termino habiaratuen euskaratzeko transduktoreari deitzen diogu. Horrela, euskarazko ordainak jasotzen ditugu: “sudur-hezur+Eko haustura” eta “sudur-hezur+areM haustura” lortuz (hierarkien etiketak desagertu dira eta ingelesezko termino bakoitzaren euskarazko ordaina dugu irteeran).

5. Ostein, termino habiaratuetako bat jatorrian pluralean zegoenez, horren deklinabideak eguneratzen ditugu pluralaren forma jaso dezan: “+Eko” “+etako” bihurtzen da, eta “+areM” “+eM”. Kasu honetan, “sudur-hezur+etako haustura” eta “sudur-hezur+eM haustura” lortu ditugu. Gogoratu, sarrerako terminoa pluralean dagoela (*fracture of nasal bones*). Adibide honetan gertatzen ez bada ere, izenondoren bat egonez gero, horren berrantolaketa ere gauzatzen da urrats honetan. Izan ere, euskaraz, izenondoak izenaren ondoren kokatzen dira, eta izenlagunak izenaren aurretik. Adibidez, “onbera” izenondoa, izenaren ostean kokatzen da, adibidez, “neoplasia onbera” terminoan ikusten dugun moduan. Izenlagunak, aldiz, izenaren aurrean gelditzen dira ingelesezko adjektiboen antzera, adibidez “jaiotzetiko kiste” terminoan “jaiotzetiko” izenlaguna izenaren aurretik ikus dezakegu.

6. Azkenik, termino habiaratuei deklinabideak gehitzeko transduktoreari deitzen diogu, euskarazko termino konposatuak lortuz: “sudur-hezurretako haustura” eta “sudur-hezurren haustura”.

Azaldu berri dugun euskaratze-prozesu honetan hainbat faktore hartu ditugu kontuan. Genitiboaren kasuan, zenbaitetan ez da erraza izaten lekuzko genitiboa ala edutezko genitiboa erabili behar den erabakitzea. Adibidez, “*abdominal aorta*” terminorako, Euskaltermek lekuzko geniti-

boa, “abdomeneko barrunbe” erabiltzen du eta Anatomiako Atlasak genitiboa erabiltzen du, “abdomenaren barrunbe”. Adituekin irizpide bat ezartzen saiatu gara: kokapena adierazi nahi denean, lekuzko genitiboa; osoa-zatia erlazioa adierazi nahi denean, edutezko genitiboa (Zabala et al., 2012). Hala ere, irizpide hori automatizatzea ez da argia, eta adituen iritzi jantzia ezinbestekoa da, testuingurua eta horren ulertzea ezinbestekoa delako. Hori horrela izanik, eta lan honen helburua ez denez erreferentziazko ordainak sortzea, baizik eta ordain hautagaiak sortzea, gainsorkuntzaren bidea aukeratu dugu kasu horietan, eta bi modutara osatu ditugu terminoak.

Terminoaren euskaratze-patroiak ahalik eta zabalenak izan daitezzen saiatu gara, orokortzea posible izan den heinean. Hona hemen kontuan izan ditugun fenomeno batzuk:

- Preposizioen euskaratze anitza: ingelesezko preposizio asko SNO-MED CTko kalifikatzaileen hierarkian daude, baina preposizioak ez dira denak berdin euskaratzen. Adibidez, *with* normalean soziatibo deklinabidearekin ordezkatzen da, *on* inesiboarekin, *to* helburuzkoarekin, etab. Gure erregelatan, deklinabide-markak ez dira lexikoi bidez gehitzen, eta erregela zehatzek gehitzen dituzte beharrezkoak direnetan.
- Adjektiboen ordena: Kalifikatzaileen hierarkiaren kasuan, hasiera batean orokortzea ezinezkoa zela iruditu zitzaigun, adjektiboek euskaraz ordenan aldaerak dituztelako. Hala ere, izenondo eta izenlagun ezberdintasunari esker orokortzea lortu dugu. Izenondoa bada, izenaren ostera mugitzen dugu kalifikatzailea den adjektiboa, eta bestelakoetan dagoen bezala uzten dugu, ingelesezko terminoaren ordena berdinean. [QUALIFIER]+[DISORDER] egitura duten *benign neoplasm* eta *congenital cyst* adibideen kasuan, euskaraz “neoplasia onbera” eta “jaiotzetiko kiste” dira ordainak; “onbera” izenondoa izanik izenaren ostera mugitu dugu, eta “jaiotzetiko” izenlaguna izatean izenaren aurrean utzi dugu. Hori kontuan izanik, adjektiboen kasuan, izenondo ala izenlagun informazioa lexikoietan gorde dugu.
- Genitiboa izenetan: euskaratze-patroi batzuek termino habiaratu bati genitiboaren deklinabidea gehitzea eskatzen dute. Kasu horietan, termino hori izena dela ziurtatu behar izan dugu, gainerako kasuetan ez baitu genitiboaren marka onartzen. Izena ez den kasuetan, terminoa bere horretan utzi dugu, deklinabide markarik gabe. Adibidez, ingelesezko *hypertrophic rhinitis* terminoan, *hypertrophic* gorputz-egituren hierarkiakoa da, eta *rhinitis* nahasmendu bat. BODYSTRUCTURE+DISORDER egiturarako euskarazko NA-

---

**Adjektiboen ordena:**  
**Kalifikatzaileen hierarkiaren kasuan, hasiera batean orokortzea ezinezkoa zela iruditu zitzaigun, adjektiboek euskaraz ordenan aldaerak dituztelako. Hala ere, izenondo eta izenlagun ezberdintasunari esker orokortzea lortu dugu.**

**Estandar  
klinikoen  
itzulpena lan  
sakona eta  
garestia da, eta  
aditu askoren  
denbora luzea  
eskatzen du.  
Itzulpen  
automatikoak  
estandar  
klinikoaren  
itzulpenean  
laguntza handia  
eskain dezake,  
itzulpenaz  
arduratzen diren  
adituen lana  
biziki arinduz.**

HASMENDU+aren edo +ko GORPUTZ-EGITURA egitura balioki-dea sortzen dugu. Adibideari jarraituz, “errinitis hipertrofikoko” edo “errinitis hipertrofikoaren” sortuko genuke, ordain akastunak izanik. Izan ere, patroi mota horiek (genitiboa eskatzen dutenak), izenetan betetzen dira, eta ez adjektiboetan.

- Termino habiaratuak pluralean: sistemaren estaldura hobetzeko, terminoen formak lexikoietan ez aurkitzean, horien lema bilatu ditugu. Izan ere, AnaMedek estrategia berdina erabiltzen du terminoen identifikazioan. Forma pluralean badago, transduktoreei terminoaren forma singularra pasatzen diegu eta pluralaren marka gehitzen diogu, deklinabidea ondo gehitu ahal izateko (gogoratu 8. irudiko *nasal bones* adibidea).

### 3. ONDORIOAK ETA ETORKIZUNERAKO LANA

Artikulu honetan, estandar klinikoen euskaratze automatikorako garatu dugun algoritmoa eta tresnak aurkeztu ditugu. Termino neoklasikoak euskaratzeko NeoTerm deituriko sistema sortu dugu, zeinak ingelesezko termino neoklasikoak abiapuntu izanik, horien morfologia baliatzen duen euskarazko ordainak sortzeko. Aurrizki eta atzizkien konposaketan oinarritzeaz gain, transliteraziorako modulu bat ere garatu dugu, Neo-Termen eragina areagotuz.

Termino konplexuetarako aldiz, KabiTerm sistema garatu dugu. Honek termino habiaratuen egitura baliatzen du ingelesezko termino konplexuak euskaratzeko. Ideia nagusia zera da: termino konplexuen barruan maiz beste termino batzuk agertu ohi dira habiaratuta. Habiaratutako terminoa, barrukoa, euskaratuta badago, euskaratze-patroiak definitu ahal dira termino konplexu osoaren ordainak lortzeko. Helburu horrekin termino habiaratu horien SNOMED CTren hierarkiak erabili ditugu terminoen egiturak aztertzeko eta euskaratze-patroiak definitzeko.

Estandar klinikoen itzulpena lan sakona eta garestia da, eta aditu askoren denbora luzea eskatzen du. Itzulpen automatikoak estandar klinikoaren itzulpenean laguntza handia eskain dezake, itzulpenaz arduratzen diren adituen lana biziki arinduz. Horretaz gain, itzulpen prozesuaren sendotasuna eta koherentzia mantentzeko tresna oso lagungarria da, termino zehatzen ordain berdinak behin eta berriz erabiltzea ahalbidetzen baitu inongo kosturik gabe.

Hemen aurkeztutako algoritmoaren inplementazio osatuari esker, SNOMED CT estandar klinikoaren euskarazko lehen bertsio automatikoa

lortu dugu. Hemendik aurrera, bertsio horren eskuzko balidazioa izango da eginkizun nagusia. Gure aldetik, balidazio hori bi urratsetan egitea proposatzen dugu. Aurrenik, mediku zein erizainek komunitatean kontzeptu bakoitzerako termino egokiak aukeratzeko kanpaina prestatuko dugu. Horrela, kontzeptu bakoitzeko semantikoki zuzenak izan daitezkeen termino multzo bat lortuko dugu, erabiltzaileek aukeratuak. Bigarrenik, terminologia klinikoan adituak diren hizkuntzalari, itzultzaile zein mediku eta erizainek azken balidazioa egingo dute, terminoen egokitasuna bermatu eta erreferentziazkoa izan daitezkeen kontzeptu bakoitzerako termino hobetsia aukeratzeko. SNOMED CT-ren euskarazko bertsioa sortzeko definitutako algoritmoa une honetan Gaixotasunen Nazioarteko Sailkapenaren bertsio bat itzultzeko (10.ena modifikazio klinikoarekin) erabiltzen ari gara, Osakidetzarekin elkarlanean. SNOMED CTren euskaratzearekin batera, XuxenMed, medikuntzara egokitutako zuzentzaile ortografikoa sortu dugu.

Estandar kliniko eleaniztunek, hizkuntzen arteko baliokidetzak eskaintzeaz gain, ikuspegi informatikotik begiratuta ate asko irekitzen dizkigute. Adibidez, osasun-txostenen kalitatezko itzulpen automatikorako oinarria izan daitezke, edota hizkuntza ezberdinetan idatzitako osasun-txostenetan informazioa modu kohesionatu batean erauzteko oinarria ere. SNOMED CT euskaraz lortzearekin, euskara gaztelania zein ingelesa bezain hizkuntza indartsuen parean jartzen dugu eta horrela, oztopo soziolinguistikoei aurre egiteko baliabide berri bat lortzen dugu. ●

## **OHARRA**

1. <http://ixa.si.ehu.es/sites/default/files/dokumentuak/8874/NAGUSIA.pdf>

## **BIBLIOGRAFIA**

- Desjardins L. Le santé des francophones du Nouveau-Brunswick. Petit-Rocher, Société des Acadiens et des Acadiennes du Nouveau-Brunswick, 2003.
- European Observatory on Health Care Systems. Luxembourg: Health system review. Health Systems in Transition, 1999.
- Euskaltzaindia. Luis Mitxelena. 0. araua. ortografia, 1968.
- Gerkens S. eta Merkur S. Belgium: Health system review. Health Systems in Transition, 12(5):1-266, 2010.
- McCarthy E.M. <http://macroevolution.net>, 2016.
- McCray A.T., Browne A.C., eta Moore D.L. The semantic structure of neoclassical compounds. Proceedings of the Annual Symposium on Computer Application in Medical Care, page 165. American Medical Informatics Association, 1988.

---

***Estandar kliniko eleaniztunek, hizkuntzen arteko baliokidetzak eskaintzeaz gain, ikuspegi informatikotik begiratuta ate asko irekitzen dizkigute. Adibidez, osasun-txostenen kalitatezko itzulpen automatikorako oinarria izan daitezke, edota hizkuntza ezberdinetan idatzitako osasun-txostenetan informazioa modu kohesionatu batean erauzteko oinarria ere.***

Stedman T.L. Stedman's Medical Dictionary, chapter Medical Prefixes, Suffixes, and Combining Forms. Lippincott Williams & Wilkins, twenty-eighth edition edition, 2005.

Wikipedia. List of medical roots, suffixes and prefixes – Wikipedia, The Free Encyclopedia. [http://en.wikipedia.org/w/index.php?title=List\\_of\\_medical\\_roots,\\_suffixes\\_and\\_prefixes](http://en.wikipedia.org/w/index.php?title=List_of_medical_roots,_suffixes_and_prefixes), 2013.

Zabala I., San Martin I., Lersundi M., Azkue J.J., eta Mendizabal J.L. The Elaboration of Human Anatomy Terminology for the Basque Language: the Contribution of Translators, Linguists and Experts. *Terminàlia*, 15– 25, 2012.