

Idazlanen ebaluazio automatizatua Lehen Hezkuntzan: euskararen berezkotasunak

Evaluación automatizada de redacciones en Educación Primaria: particularidades del euskera

Évaluation automatisée des rédactions à l'école primaire : particularités du basque

Automated Assessment of Essays in Primary Education: Particularities of the Basque

IZETA BALENTZIAGA, Inar

Ikertzaile beregaina

iizeta001@ikasle.ehu.eus

ORCID: <https://orcid.org/0009-0007-1928-4655>

IRUSKIETA QUINTIAN, Mikel

Euskal Herriko Unibertsitatea UPV/EHU

Bilboko Hezkuntza Fakultatea

HiTZ Zentroa - Ixa Taldea

mikel.iruskieta@ehu.eus

ORCID: <https://orcid.org/0000-0002-6121-3902>

IBARRA LIZUNDIA, Irune

Euskal Herriko Unibertsitatea UPV/EHU

Hezkuntza, Filosofia eta Antropologia Fakultatea

irune.ibarra@ehu.eus

ORCID: <https://orcid.org/0000-0002-6655-4327>

DOI: <https://doi.org/10.59866/eia.v2i71.352>

Lan hau burutzeko, Gipuzkoako Foru Aldundiaren dirulaguntza jaso dugu, «Etorkizuna Eraikiz» ekimenaren bidez (2024-2025 deialdiko 17. proiektua). Halaber, egileek eskerrak eman nahi dizkiete laginak biltzeko aukera eman duen ikastetxeari; bereziki, zuzendaritzari, tutoreei, aholkulariari eta, nola ez, parte hartu duten haurrei.

Egileen zereginak:

Inar Izeta: Kontzeptualizazioa, metodologia, datuen tratamendua (digitalizazioa), softwarea (adimen artifizialaren entrenamendua), analisi formala, irudien ekoizpena eta jatorrizko zirriborroa idaztea.

Mikel Iruskietia: Kontzeptualizazioa, metodologia, softwarea (adimen artifizialaren entrenamendua), idazketa-berrikuspena eta edizioa.

Irune Ibarra: Kontzeptualizazioa, metodologia, datuen tratamendua (datuak eskuratzeko), idazketa-berrikuspena eta edizioa.

Jasotze-data: 2026-02-05

Behin-behineko jasotze-data: 2026-06-20

Behin betiko jasotze-data: 2026-08-06

Ikerketa honetan Lehen Hezkuntza (LH) idazmena nola garatzen den aztertu da, hainbat ikasmailatan. Euskarazko 181 testu digitalizatu eta CTAP (*Common Text Analysis Platform*) tresnaren bidez haien adierazle linguistikoak eratu dira, CAF (*Complexity, Accuracy, Fluency*) markoa erabiliz konplexutasun neurri egokienak hautatzeko. Emaitzek erakutsi dute zenbait adierazle alde esanguratsuak detektatzen dituztela ikasmailen artean. Irakasleek emandako kalifikazioen artean, diskriminatzen gaitasun handiena duen adierazlea aditz-sintagmen kopurua da.

Gako-hitzak: Euskara, idazmena, ebaluazioa, Lehen Hezkuntza, Hizkuntza Naturalaren Prozesamendua (HNP), CTAP (*Common Text Analysis Platform*).

En esta investigación se ha analizado el desarrollo de la expresión escrita en distintos niveles de Educación Primaria. Se han digitalizado 181 textos en euskera y, mediante la herramienta CTAP (*Common Text Analysis Platform*), se han extraído sus indicadores lingüísticos, utilizando el marco CAF (*Complexity, Accuracy, Fluency*) para seleccionar las medidas de complejidad más adecuadas. Los resultados han mostrado que los indicadores detectan diferencias significativas entre los niveles escolares. En 4.º de Educación Primaria, el número de sintagmas verbales es el indicador con mayor capacidad para discriminar entre las calificaciones otorgadas por el profesorado.

Palabras clave: Euskara, escritura, evaluación, Educación Primaria, Procesamiento del Lenguaje Natural (PLN), CTAP (*Common Text Analysis Platform*).

Cette recherche a analysé le développement de l'expression écrite à différents niveaux de l'enseignement primaire. 181 textes en basque ont été numérisés et leurs indicateurs linguistiques ont été extraits à l'aide de l'outil CTAP (*Common Text Analysis Platform*), en utilisant le cadre CAF (*Complexity, Accuracy, Fluency*) pour sélectionner les mesures de complexité les plus appropriées. Les résultats démontrent que certains indicateurs mettent en évidence des différences significatives entre les niveaux scolaires. En CM1 (4e année du primaire), le nombre de syntagmes verbaux est l'indicateur ayant la plus grande capacité à discriminer les notes attribuées par les enseignants.

Mots-clés : Basque, écriture, évaluation, éducation primaire, Traitement Automatique du Langage Naturel (TALN), CTAP (*Common Text Analysis Platform*).

This research analyzed the development of writing skills across different grades of Primary Education. One hundred eighty-one texts in Basque were digitized, and their linguistic features were extracted using the CTAP (*Common Text Analysis Platform*) tool, employing the CAF (*Complexity, Accuracy, Fluency*) framework to select the most appropriate complexity measures. The results demonstrate that certain indicators detect significant differences between school levels. In the 4th grade of Primary Education, the number of verb phrases is the indicator with the greatest capacity to discriminate between the scores assigned by teachers.

Keywords: Basque, writing, assessment, Primary Education, Natural Language Processing (NLP), CTAP (*Common Text Analysis Platform*).

IZETA BALENTZIAGA, Inar, IRUSKIETA QUINTIAN, Mikel and IBARRA LIZUNDIA, Irune (2026) «Automated assessment of essays in primary education: particularities of the Basque», *Euskera Ikerketa Aldizkaria* 71 (2): 227-266.

1 Sarrera

- 1.1 Ikerketaren testuingurua eta helburuak
- 1.2 CAF marko teorikoa: konplexutasuna, zehaztasuna eta jariakortasuna
- 1.3 Euskararen ezaugarri tipologikoak idazmenaren analisian
- 1.4 Idazmenaren azterketa automatizatua: CTAP tresna

2 Metodologia

- 2.1 Diseinua
- 2.2 Parte-hartzaileak
- 2.3 Tresnak
- 2.4 Prozedura
- 2.5 Datuen analisia

3 Emaitzak

- 3.1 Adierazle linguistikoen hautaketa eta egokitasuna
- 3.2 Adierazle linguistikoen bilakaera LHn
 - 3.2.1 Estatistika deskribatzaileak
 - 3.2.2 Ezberdintasun estatistikoen analisia
 - 3.2.3 Bilakaeraren deskribapen kualitatiboa
 - 3.2.3.1 Aditz trinkoen eta perifrastikoen bilakaera
 - 3.2.3.2 Diskurtsoaren analisia: modalitatea eta perpaus-elkarketa
 - 3.2.3.3 Kohesioa: izenen gainjartze globala
- 3.3 Adierazle linguistikoen eragina kalifikazioetan ikasmailaren arabera
 - 3.3.1 Ezberdintasun estatistikoen analisia
 - 3.3.2 Ezberdintasunen deskribapen kualitatiboa

4 Eztabaida

5 Ondorioak

Bibliografia

Eranskinak

- 1 LH4ko testuak ebaluatzeko errubrika
- 2 LH6ko testuak ebaluatzeko errubrika
- 3 CTAP tresna euskarazko testuetara egokitzeko prozedura

1 Sarrera

1.1 Ikerketaren testuingurua eta helburuak

Idazmena garatzea funtsezko mugarria da ikasleen gaitasun kognitibo eta linguistikoen bilakaeran. Lehen Hezkuntzan (LH), ikasleek alfabetatzearen hasierako fasea gaindituta, testu egituratuagoak eta koherenteagoak ekoizten dituzte, eta testu horietan hizkuntza hobeki menderatzen dutenaren zantzuak erakusten dituzte, argiro. Garapen hori modu automatizatuan neurtzea erronka handia da ikerlarientzat, haurren garapen linguistikoa ez baita beti lineala eta adierazle ezberdinen arteko oreka aldatu egiten da adinean aurrera egin ahala (Arriola, Goikoetxea eta Iruskieta 2020).

Ikerketa honen helburua da ikasleen idazmenaren garapen linguistikoa eta horrek eskolako kalifikazioetan duen eragina modu automatizatuan aztertzea, ikuspegi kuantitatibo, kualitatibo, zeharkako eta ez-esperimental batetik. Zehazki, lau dimentsio linguistikoren (jariakortasuna, lexikoa, sintaxia eta kohesioa) bilakaera eta horiek eskolako kalifikazioetan duten eragina aztertu nahi da.

1.2 CAF marko teorikoa: konplexutasuna, zehaztasuna eta jariakortasuna

Lan honen marko nagusia nazioartean idazmena ebaluatzeko erabiltzen den CAF (*Complexity, Accuracy, Fluency*) eredua da (Bulté eta Housen 2012). Ereduek hiru dimentsio independentetan banatzen du hizkuntzaren erabilera, eta ikerketa honetan, dimentsio horiek euskararen ezaugarrietara eta azterketa automatikora egokitu dira:

- Jariakortasuna (*fluency*): Testua ekoizteko erraztasunari dagokio eta prozesamendu kognitiboaren arintasuna islatzen du. Ekoizpen-ahalmen gisa ulertzen da, eta testu-bolumenaren bidez neurtzen da, esaterako, hitz edo unitate linguistikoen kopuruaren bidez.
- Zehaztasuna (*accuracy*): Testuak hizkuntzaren arauetara duen atxikimendua da, akatsen maiztasunaren bidez neurtua. Lan honetan,

ordea, dimentsio hori eredu teorikotik kanpo utzi da, euskararen prozesamendu automatikoan akats gramatikalak fidagarritasunez neurtzeko tresna estandarizatuak falta direlako oraindik.

- Konplexutasuna (*complexity*): Ikasleak erabiltzen dituen egitura linguistikoen aberastasunean oinarritzen da. Euskararen tipologia aintzat hartuta, lan honek konplexutasuna aditzaren zentralitatearekin lotzen du: konplexutasun lexikoa aditz-moten kopuruan oinarritzen da eta konplexutasun sintaktikoa aditz-sintagmen kopuruan.

1.3 Euskararen ezaugarri tipologikoak idazmenaren analisian

Euskara hizkuntza eranskaria (aglutinatzailea) eta hitz-ordena malgu samarra duen hizkuntza izateak erronka nabarmenak planteatzen dizkio idazmenaren azterketa automatikoari. Ezaugarri tipologikoek (aglutinazioa, kasu-sistema aberatsa eta aditzaren zentralitatea) zuzenean eragiten dute testuaren egitura morfosintaktikoan zein diskurtsiboan, eta, ondorioz, erabilitako neurketa-adierazleen interpretazioan ere bai (Hualde eta Ortiz de Urbina 2003):

- Flexio-morfologia aberatsa: Euskarak morfema flexiboen sistema zabal du (kasua, numeroa, determinazioa eta aditz-komunztadura), eta horrek hitzen aldakortasun handia eragiten du (Euskaltzaindiaren Gramatika Batzordea 2021). Ondorioz, hitz-kopuru soiletan (tokenetan¹) oinarritutako neurketek lexikoaren erabilera errealaaren irudi partziala edo distortsionatua eman dezakete, lema bakar batek forma ugari har ditzakeelako (Aduriz *et al.* 2000). Testuinguru horretan, lema-mailako adierazleak egokiagoak dira ikaslearen errepertorio lexikoa modu fidagarriagoan islatzeko (McCarthy eta Jarvis 2010).
- Aditzaren eginkizun egituratzailea sintaxian: Euskaraz, aditza da perpausaren egitura sintaktikoaren nukleo nagusia, bertan kodetzen baitira argumentu-egitura, kasu-harremanak eta komunztadura-informazioa

¹ Tokena testu bateko oinarritzko unitaterik txikiena da; sistemak hitz, zenbaki edo puntuazio-marka gisa identifikatzen duena.

(Hualde eta Ortiz de Urbina 2003). Aditzaren bidez markatzen dira perpauseko partaideen arteko harremanak, baita perpausaren izaera gramatikala ere. Hori dela eta, aditz-formadun egituren kopurua adierazle egokia da idazlearen konplexutasun sintaktikoa neurtzeko (Biber, Conrad eta Reppen 1998).

- Kohesio erreferentziala: Euskararen *pro-drop* izaeraren ondorioz, erreferentziazko argumentuen jarraipena sarritan aditzaren morfologian txertatzen da, eta izen-sintagmen elipsia ohikoa da (Hualde eta Ortiz de Urbina 2003). Horrek aukera ematen du argumentu berak testuan zehar modu esplizitu zein inplizituan berrerrabiltzeko, kohesio diskurtiboa bermatuz. Testu-mailako argumentu-gainjartze globala neurtzeak testuaren barne-koherentzia ebaluatzeko informazio baliotsua eskaintzen du (Givón *et al.* 1983).

1.4 Idazmenaren azterketa automatizatua: CTAP tresna

Hizkuntza Naturalaren Prozesamenduaren (NLP) garapenak testuen analisi masiboa eta objektiboa ahalbidetu du (Azurmendi, Arregi eta López de Lacalle 2025). CTAP (*Common Text Analysis Platform*) tresnak testuen ehunka ezaugarri linguistiko automatikoki erauzteko aukera ematen die ikerlariei (Chen eta Meurers 2016). Nolanahi ere, kontuan hartu behar da tresna hori ingelesezako diseinatu zela. Hori dela eta, euskarazko testuak aztertzeke, ezinbestekoa da aldeztu aurretik euskararen prozesamendurako berariazko NLP tresnak integratzea (3. eranskina), hala nola Lexikoaren Behatokia corpusaren (Euskaltzaindia 2014) 2014ko datuetan oinarrituta entrenatutakoak. Nazioarteko ikerketek (McNamara, Crossley eta McCarthy 2010) erakutsi dute tresna horiek erabiliz lortutako adierazleak (kohesio-indizeak, sintaxi-neurriak, etab.) korrelazio altua dutela giza ebaluatzaileek emandako kalifikazioekin. Hala ere, LHko euskarazko testuekin egindako azterketak oraindik urriak dira, eta horixe da ikerketa honen ekarpen nagusia: idazmenaren garapena ongien islatzen duten adierazle linguistikoak identifikatzea eta LHn duten bilakaera aztertzea.

2 Metodologia

2.1 Diseinua

Ikerketa hau diseinu kuantitatibo, kualitatibo, zeharkako eta ez-esperimental baten bidez egituratu da. Helburua ez da aldagaiak laborategi-ingurune batean manipulatzeari, baizik eta une jakin batean ikastaldeen ekoizpen linguistikoa alderatzea. Metodologiari dagokionez, korrelazioetan oinarritutako ikuspegi deskriptibo batetik, adierazle linguistikoen bilakaera eta haien arteko elkarreragina aztertu dira LHn zehar. Horrez gain, analisi inferentzialaren bidez, egiaztatu da adierazle horiek ikasmailan eta ikasleen kalifikazioetan duten isla estatistikoki esanguratsua dela. Azkenik, adierazle bilakaera eta kalifikazioetan duten eragina ikuspegi kualitatibotik ere aztertu dira, ikasleen testuetako pasarte adierazgarrienak aztertuz.

Beraz, lau dira ikerketa galderak: (1) Zein dira LHko ikasmailak ongien bereizten dituzten adierazle linguistikoak? (2) Zein lotura dago adierazle horien artean? (3) Zein garapen dute adierazle horiek LHn zehar? (4) Ikasmaila bakoitzean, nolakoa da adierazle linguistikoen eta testuek jasotako kalifikazioen arteko elkarrekikotasuna? Horrenbestez, aldagai independenteak ikasmaila (LH1etik LH6ra) eta irakasleak testu bakoitzari emandako kalifikazioa (oinarrizko maila, erdiko maila eta maila aurreratua) dira, eta menpeko aldagaiak, aldiz, hautatutako adierazle linguistikoak eta horien bilakaera.

2.2 Parte-hartzaileak

Ikerketa honetako lagina Gipuzkoako eskola publiko bateko LHko ikasle-
lez osatuta dago. Guztira 186 ikaslek parte hartu zuten testuen idazketan,
lau mailatan banatuta: lehenengo maila (LH1), bigarren maila (LH2), lau-
garren maila (LH4) eta seigarren maila (LH6). Maila horiek ziklo bakoitza-
ren hasiera eta LHren amaiera ordezkatzeko hautatu ziren.

Datuak garbitzeko prozesuan, bost testu baztertu dira honako irizpide
hauen arabera: (a) Pedagogia Terapeutikoko (PT) laguntzailearekin idatzi-

tako testu bat, (b) eskatutako hizkuntzan idatzi ez diren bi testu eta (c) kaligrafia kaxkarra izateagatik irakurgaitzak diren beste bi testu. Ondorioz, azken lagina 181 parte-hartzailek ($N = 181$) osatu dute. Testuak anonimizatuta jaso dira. Parte-hartzaileen banaketa 1. taulan xehatzen da.

1. taula. Parte-hartzaileen banaketa ikasmailaren eta kalifikazioaren arabera

	Oinarrizko maila	Erdiko maila	Maila aurreratua	Guztira
LH1	7	10	1	18
LH2	12	19	31	62
LH4	20	46	20	86
LH6	4	6	5	15
Guztira	43	81	57	181

Iturria: Egileak prestatua.

Oharra: ($N = 181$).

2.3 Tresnak

Ikerketan bildutako datuak prozesatzeko eta aztertzeko, software espezifikoak erabili dira hiru fasetan antolatuta:

1. Digitalizazioa: *Transkribus* (Read-Coop Sce 2015) tresna erabili da eskuz idatzitako testuak formatu digitalera bihurtzeko. Euskarazko HTR ereduak erabili dira aurrelanketa egiteko eta, ondoren, eskuz zuzendu dira testu guztiak, idazlearen akatsak errespetatuta.
2. Adierazleen erauzketa: CTAP (*Common Text Analysis Platform*) tresna baliatu da adierazle linguistikoen neurketa automatikorako (Chen eta Meurers 2016). Tresna horrek NLPko teknikak erabiliz hizkuntzaren konplexutasun-neurriak erauzten ditu hautatutako adierazleen (hitz-token kopurua, aditz-moten kopurua, aditz-sintagmen kopurua eta

izenen gainjartze globala) balio kuantitatiboak kalkulatu. Jatorrizko tresnak euskarazko testuak onartzen ez dituenaz, azpitresna espezi-fikoak integratu dira (ikus 3. eranskina).

3. Datuen analisia: Antolaketa *Microsoft Excel* bidez egin da eta analisi estatistikorako *Python* programazio-lengoaia erabili da, zehazki, *Pandas* (McKinney 2010), *SciPy* (Virtanen et al. 2020) eta *Seaborn* (Waskom 2021) liburutegiak.

2.4 Prozedura

Ikerketaren prozedura bost fasetan garatu da:

1. Lehenik, datu-bilketa egin da, Eusko Jaurlaritzaren ebaluazio diagnostikoaren esparruan. Ikasleek beren ohiko ikasgeletan idatzi dituzte testuak, ordubete inguruko saio batean eta tutoreak gainbegiratura. Testuak ikasmaila bakoitzari egokitutako gai baten gainekoak dira. Horrez gain, aurretik saio bat egin dute testu-mota bakoitzaren egitura lantzeko: irudi baten deskribapena (LH1), asteburuko bizipenen kontakizuna (LH2), norberaren autobiografia (LH4) eta albiste baten gaineko iritzi-artikulua (LH6).
2. Jarraian, tutoreek berek ebaluatu dituzte testuak, errubrika bateratuen arabera (ikus 1. eta 2. eranskinak). Zuzenketa horien kalifikazioak aldagai kategoriko gisa jaso dira, hiru mailatan sailkatuta: oinarrizko maila, erdiko maila eta maila aurreratua.
3. Ondoren, eskuz idatzitako testuak digitalizatu dira. Testuak eskaneatu ondoren, *Transkribus* plataformara igo dira irudiak, testu editagarri bihurtzeko.
4. Segidan, testuak CTAP plataformara (Chen eta Meurers 2016) igo dira eta bakoitzari zegozkion metadatuak (ikasmaila eta kalifikazioa) esleitu zitzaizkion.
5. Azkenik, CTAP tresnari (Chen eta Meurers 2016) euskarazko testuak aztertzeke azpitresnak integratu ondoren (ikus 3. eranskina), adierazle

linguistiko guztiak aukeratu eta testuen ebaluazio automatizatua burutu da NLP algoritmoak aplikatuz. Emaitzak .csv formatuan jaso ziren, eta testu bakoitzak adierazle linguistiko bakoitzean lortutako balio kuantitatiboa zehaztu zen.

2.5 Datuen analisisa

Datuen prozesamendua bost fasetan burutu da:

1. Lehenik, aurreprozesamendua eta garbiketa egin da *Microsoft Excel* bidez, baliagarriak ez diren testuak kentzeko. Informazioa taula batean antolatu eta .csv formatura esportatu da.
2. Bigarrenik, adierazle linguistikoen hautaketa edo *screening* estatistiko bat egin da. Hautaketa horretarako, CAF (*Complexity, Accuracy, Fluency*) ereduaren egokitzapen bat hartu da abiapuntutzat; konplexutasuna (*complexity*) hiru azpidimentsiotan banatu da (konplexutasun lexikoa, konplexutasun sintaktikoa eta kohesioa), jariakortasunaren (*fluency*) dimentsioa mantenduz. Zehaztasunaren (*accuracy*) dimentsioa ez da kontuan hartu, euskararen prozesamendu automatikoan akats gramatikalak fidagarritasunez neurtzeko tresna estandarizatuak falta direlako oraindik. Jarraian, CTAP tresnak (Chen eta Meurers 2016) eskainitako adierazle guztiak dimentsioetan banatu dira eta dimentsio bakoitzeko adierazle egokiena aukeratu da, irizpide nagusia ikasmailen arteko aldeak hautemateko ahalmena izanik. Horretarako, ikasmailarekin korrelazio altuena (Spearmanen *rs*) erakusten duten adierazleak hautatu dira. Bestalde, adierazleen arteko harremana neurtu da Spearmanen *rs* korrelazio-koefizientearen bidez.
3. Behin datu-basea prestatuta, analisi inferentzialari ekin zaio. Datuen banaketa ez-normala eta bariantza heterogeneokoa dela egiaztatu ondoren (Shapiro-Wilk eta Levene proben bidez), metodo ez-parametrikokoak lehenetsi dira. Hori horrela, taldeen arteko konparaketa globalerako (bai ikasmailen artean, bai kalifikazioen artean) Kruskal-Wallis *H* proba erabili da. Estatistikoki esanguratsuak diren kasuetan,

post-hoc analisiak burutu dira binakako konparaketen bidez (Mann-Whitney U), Bonferroni zuzenketa aplikatuz I. motako errorea (alfa akatsa) saihesteko eta analisi guztietan esangura-maila $\alpha = 0,05$ ean ezarri da.

4. Segidan, lerro-grafikoak, kaxa-diagramak eta korrelazio-mapak sortu dira, emaitza esanguratsuak irudikatzeko.
5. Amaitzeko, ikasleen testuak banan-banan aztertu dira. Horretarako, adierazle linguistikoek ikasmailetan zehar izandako garapena islatzen duten pasarte adierazgarrienak hautatu dira, baita kalifikazioen artean diskriminatzeko baliagarriak izan daitezkeenak ere.

3 Emaitzak

3.1 Adierazle linguistikoen hautaketa eta egokitasuna

Arestian aipatu bezala, dimentsioak egokien islatzen dituzten adierazle linguistikoak aukeratu dira (adierazleak kalkulatzeko prozedura zehatza 3. eranskinean ikus daiteke), ikasmailekiko duten korrelazioaren arabera (ikus 2. taula):

1. Hitz-token kopurua, puntuazio-markak eta zenbakiak baztertuz: Adierazle horrek testuan agertzen diren hitz-agerpen guztien kopurua neurtzen du, errepikapenak barne. Neurri horrek testuaren produkzio-kantitatea edo luzera efektiboa islatzen du ($r_s = 0,83, p < 0,001$).
2. Aditz-moten kopurua: Aditz-mota gehiago erabiltzeak gaitasun linguistiko handiagoa erakusten du ($r_s = 0,84, p < 0,001$).
3. Aditz-sintagmen kopurua: Aditz-sintagmen kopuru handiagoak per-pausa egituratzeko gaitasun handiagoa eta egitura sintaktiko konplexuagoak islatzen ditu ($r_s = 0,81, p < 0,001$).
4. Izenen gainjartze globalaren desbideratze estandarra, lemetan oinarritua: Balio altuek testuaren kohesio erreferentzial handia adierazten dute ($r_s = 0,58, p < 0,001$).

2. taula. Adierazleen hautaketa eta ikasmailekiko korrelazioa

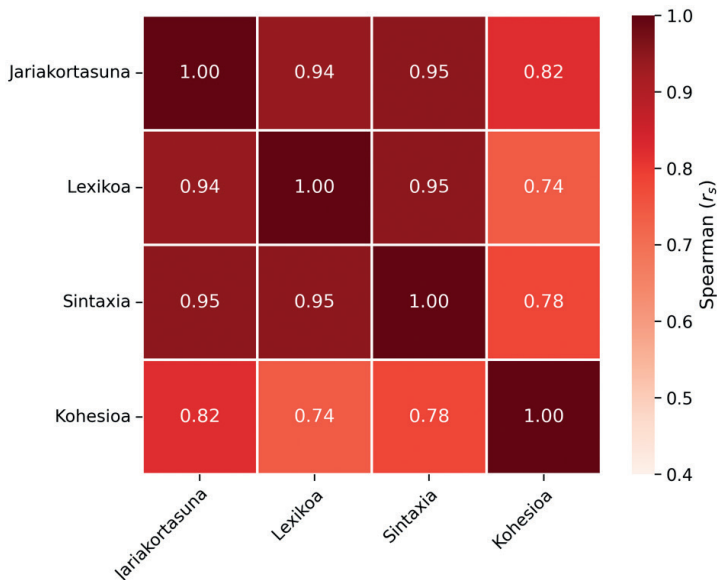
Dimentsioa	Adierazlea (CTAP)	Spearmanen r_s	p
Jariakortasuna	Hitz-token kopurua (puntuazioa eta zenbakiak baztertuta)	0,833	< 0,001
Lexikoa	Aditz-moten aniztasuna (modalak barne)	0,837	< 0,001
Sintaxia	Osagai sintaktikoen kopurua: aditz-sintagma	0,807	< 0,001
Kohesioa	Argumentuen gainjartze globala (lemetan oinarrituta)	0,580	< 0,001

Iturria: Egileak prestatua.

Oharra: ($N = 181$). Korrelazio guztiak esanguratsuak dira: $p < 0,001$.

Behin adierazleak hautatuta, haien arteko harremana aztertu da eta korrelazioen analisiak patroiz esanguratsua azaleratu du (ikus 1. irudia). Batetik, jariakortasunaren, lexikoaren eta sintaxiaren arteko korrelazioak oso altuak dira ($r_s > 0,94$). Bestetik, kohesioak portaera bereizia erakutsi du, korrelazio nabarmenki baxuagoekin gainontzeko adierazleekiko ($r_s \leq 0,82$).

1. irudia. Dimentsio linguistikoen arteko korrelazioen bero-mapa (Separmanen r_s)



Iturria: Egileak prestatua ($N = 181$).

Oharra: Bero-mapak dimentsio bakoitzeko adierazle linguistikoen arteko Spearmanen korrelazioak (r_s) irudikatzen ditu. Korrelazio guztiak estatistikoki esanguratsuak ($p < 0,001$) dira.

3.2 Adierazle linguistikoen bilakaera LHn

3.2.1 Estatistika deskribatzaileak

Atal honetan, hautatutako adierazle linguistikoek mailaz maila izan duten portaera aztertzen da. Datuen banaketak normaltasunaren printzipioa eta bariantzen homogeneotasuna urratzen dituenenez ($p < 0,05$ kasu guztietan), metodo ez-parametrikokoak erabili dira. Joera zentraleko neurri gisa mediana (Mdn) erabili da eta sakabanatze-neurri gisa kuartil arteko zabalera (IQR).

3. taulan ikus daitekeenez, adierazle gehienek garapen nabarmena erakusten dute LH1etik LH4ra. Jariakortasunean, nabaria da ekoizpenak LH2tik LH4ra izan duen gorakada ($Mdn = 35$ eta $Mdn = 159$), non baliorik altuena erregistratzen den. LH6 mailan ekoizpenak apur bat behera egiten du, baina maila altuan mantentzen da ($Mdn = 144$).

Konplexutasun lexikoak eta sintaktikoak antzeko bilakaera dute: konplexutasun lexikoaren kasuan, hazkundera ez da LH4an ($Mdn = 29$) gelditzen, LH6ko ikasleek lortu baitute medianarik altuena ($Mdn = 37$). Konplexutasun sintaktikoan aldiz, LH2tik LH4rako igoeraren ondoren ($Mdn = 5$ eta $Mdn = 25$), LH6an paretsu ageri da ($Mdn = 23$), egonkortasuna iradokiz.

Azkenik, kohesioaren bilakaera zeharo ezberdina da: LH1 eta LH2 mailetako testuetan ez dago izenen gainjartzerik ($Mdn = 0$); LH4an gailurra jotzen du ($Mdn = 0,68$) eta LH6an jaisten da ($Mdn = 0,23$).

3. taula. Adierazleen estatistika deskribatzaileak (Mdn eta IQR)
ikasmailaren arabera

	Jariakortasuna	Lexikoa	Sintaxia	Kohesioa
LH1	12 (6,5)	3 (2,25)	2 (3)	0 (0)
LH2	35 (22)	8 (3)	5 (3)	0 (0,1)
LH4	159 (64)	29 (18)	25 (12)	0,68 (0,81)
LH6	144 (67,5)	37 (12,5)	23 (10)	0,23 (0,15)

Iturria: Egileak prestatua, CTAP bidez lortutako datuetan oinarrituta ($N = 181$).
Oharra: Balioek mediana (Mdn) adierazten dute eta parentesi arteko datuek kuartil arteko zabalera (IQR).

3.2.2 Ezberdintasun estatistikoen analisia

Ikasmailen arteko ezberdintasunak aztertzeko, datuek normaltasunaren eta bariantza-homogeneotasunaren printzipioa betetzen ez dutenez,

Kruskal-Wallis H proba hautatu da. Emaitzek frogatu dute lau adierazleetan ikasmailen arteko ezberdintasunak estatistikoki esanguratsuak direla ($p < 0,001$). Gainera, kalkulaturako efektuaren tamainak oso handiak dira kasu guztietan ($\eta^2 \geq 0,68$). Horrek baieztatzen du dimentsio linguistiko bakoitzerako aukeratu ditugun adierazleek bilakaera nabarmena dutela ikasmailetan aurrera egin ahala (ikus 4. taula).

4. taula. Ikasmailen arteko aldeak adierazle linguistikoetan: Kruskal-Wallis H probaren emaitzak

Dimentsioa	KW (H)	p	η^2
Jariakortasuna	128,62	$< 0,001$	0,743
Lexikoa	122,56	$< 0,001$	0,707
Sintaxia	119,16	$< 0,001$	0,687
Kohesioa	118,13	$< 0,001$	0,681

Iturria: Egileak prestatua ($N = 181$).

Oharra: Konparaketak Kruskal-Wallis H probaren bidez egin dira eta alde guztiak estatistikoki esanguratsuak dira ($p < 0,001$). η^2 balioek efektu-tamaina handia adierazten dute kasu guztietan ($\eta^2 \geq 0,681$).

Ezberdintasun horien jatorria zehazteko, binakako konparaketak (*post-hoc*) burutu dira Mann-Whitney U proben bidez, Bonferroni zuzenketarekin (p -balio doitua: $p \times 6$) (ikus 5. taula). Jariakortasuna eta sintaxia lehen urteetan garatu dira nagusiki: LH1etik LH2ra eta LH2tik LH4ra aldeak oso esanguratsuak dira ($p < 0,001$), efektu-tamaina ertain-handiekin ($r \geq 0,36$). Hala ere, LH4 eta LH6 mailen artean garapena egonkortu egin da; ez da alde esanguratsurik aurkitu ($p = 1$) eta efektu-tamaina baxuek ($r < 0,05$) gaitasun horiek LH4rako sendotu egiten direla iradokitzen dute.

Konplexutasun lexikoan, garapena motelagoa izan da. LH1 eta LH2 artean alde esanguratsuak dauden arren, efektuaren tamaina ertaina da

($r = 0,31$) eta LH2tik aurrera ez da hazkunde esanguratsurik hautematen maila hurbilen artean. Aipagarria da LH1 eta LH6 artean esangura galtzen dela Bonferroni zuzenketaren ondorioz ($p_{adj} = 0,307$).

Kohesioak, berriz, garapen berantiarragoa eta iraunkorragoa erakutsi du. Lehen zikloko mailen artean (LH1-LH2) ez dago alderik ($p = 0,399$), baina LH2tik LH4ra jauzi kualitatibo handia gertatzen da ($r = 0,80$). Beste dimentsioetan ez bezala, kohesioak beherakada handia jasaten du LH4tik LH6ra ($p = 0,001$, $r = 0,37$).

5. taula. Ikasmaien arteko bikote-konparaketak (*post-hoc*) adierazle linguistikoetan: Mann-Whitney *U* probaren emaitzak

Dimentsioa	Konparaketa	p-adj	Esang.	r (Efektutamaina)	Puntuazio altuena
Jariakortasuna	LH1 eta LH2	< 0,001	Bai	0,617	LH2
	LH1 eta LH4	< 0,001	Bai	0,629	LH4
	LH1 eta LH6	< 0,001	Bai	0,836	LH6
	LH2 eta LH4	< 0,001	Bai	0,827	LH4
	LH2 eta LH6	< 0,001	Bai	0,617	LH6
	LH4 eta LH6	1	Ez	0,046	-
Lexikoa	LH1 eta LH2	0,042	Bai	0,306	LH2
	LH1 eta LH4	0,039	Bai	0,271	LH4
	LH1 eta LH6	0,307	Ez	0,370	-
	LH2 eta LH4	1	Ez	0,102	-
	LH2 eta LH6	1	Ez	0,048	-
	LH4 eta LH6	1	Ez	0,049	-

Sintaxia	LH1 eta LH2	0,009	Bai	0,359	LH2
	LH1 eta LH4	< 0,001	Bai	0,614	LH4
	LH1 eta LH6	< 0,001	Bai	0,836	LH6
	LH2 eta LH4	< 0,001	Bai	0,804	LH4
	LH2 eta LH6	< 0,001	Bai	0,619	LH6
	LH4 eta LH6	1	Ez	0,028	-
Kohesioa	LH1 eta LH2	0,399	Ez	0,172	-
	LH1 eta LH4	< 0,001	Bai	0,615	LH4
	LH1 eta LH6	< 0,001	Bai	0,779	LH6
	LH2 eta LH4	< 0,001	Bai	0,790	LH4
	LH2 eta LH6	< 0,001	Bai	0,457	LH6
	LH4 eta LH6	0,001	Bai	0,375	LH4

Iturria: Egileak prestatua ($N = 181$).

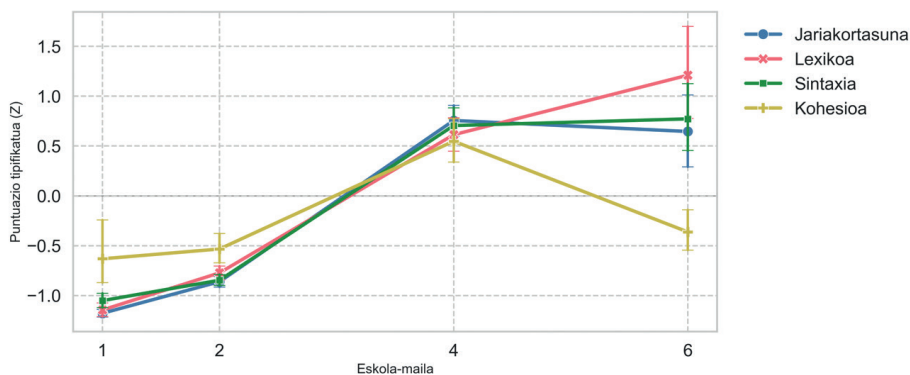
Oharra: Konparaketak Mann-Whitney U probaren bidez egin dira. p -adj: Bonferroni zuzenketaren bidez lortutako p -balio doitua ($p \times 6$). *Esang.:* Bai (alde esanguratsua, $p < 0,05$); Ez (alderik ez, $p \geq 0,05$). r (Efektu-tamaina): 0,1 (txikia), 0,3 (ertaina) eta 0,5 (handia). Puntuazio altuena lortu duen eskola-maila ere adierazita dago.

2. irudian adierazle linguistikoen bilakaera agertzen da ikasmailetan zehar. Adierazle bakoitzak eskala ezberdina duenez, datuak estandarizatu egin dira puntuazio tipifikatuen bidez (Z puntuazioak), guztien garapen-erritmoa grafiko bakarrean ikusteko. Puntu bakoitzak taldearen batezbesteko aritmetikoa adierazten du eta errore-barrek % 95eko konfiantza-tartea (KT % 95) irudikatzen dute. Ardatz bertikalak puntuazio tipifikatuak adierazten ditu; horrenbestez, 0 lerroak lagin osoaren batezbestekoa markatzen du.

2. irudian ikus daitekeen grafikoak argi islatzen ditu arestian atzemandako joera nagusiak: goranzko joera eta egonkortzea dago jariakortasunean; lexikoan eta sintaxian LH1etik LH4ra, malda handiekin. Hala ere, aurreko

Mann-Whitney U proban ez bezala, LH2tik LH4ra gorakada bat ageri da lexikoa. LH4tik LH6ra, ordea, lerro guztiak (lexikoa izan ezik) horizontalak dira edo errore-barrak gainjartzen dira. Gainjartze horrek bisualki berresten du bi maila horien artean ez dagoela alde estatistiko esanguratsurik, eta garapenaren egonkortzea irudikatzen du. Kohesioaren portaera ezberdina da, ordea: hasierako egonkortasunaren ondoren, LH4n gailurra jotzen du eta LH6n nabarmen egiten du behera, lagin osoaren batezbestekotik behera kokatuz berriro.

2. irudia. Adierazle linguistikoen bilakaera ikasmailatan zehar



Iturria: Egileak prestatua ($N = 181$).

Oharra: Balioak puntuazio tipikoen bidez (Z puntuazioak) adierazi dira dimentsioen arteko konparaketa errazteko. Puntu bakoitzak taldearen batezbesteko aritmetikoa adierazten du eta errore-barrek % 95eko konfiantza-tartea (KT % 95) irudikatzen dute.

3.2.3 Bilakaeraren deskribapen kualitatiboa

Adierazleek ikasmailaz ikasmila izan duten bilakaera estatistikoki aztertu ondoren, funtsezkoa da emaitzak ikuspegi linguistikotik interpretatzea, ikasleen testuetako pasarte adierazgarriak aipatuz. Analisi honetan alde batera utzi da jariakortasunaren adierazlea (hitz-tokenen kopurua, puntuazio-markak eta zenbakiak baztertuz, ikus 3. eranskina), hizkuntzaren garapen kualitatiboari buruzko informazio nahikorik ematen ez duelako.

3.2.3.1 ADITZ TRINKOEN ETA PERIFRASTIKOEN BILAKAERA

Aditz trinkoen errepertorioa zabaldu egiten da LHn. LH1 mailan, *izan* («olagarroa epailea da») eta *egon* («marrazoa igerilekuan dago») aditz trinkoak erabili dituzte irudiko elementuen izaera eta kokapena deskribatzeko. LH2 mailan, lehengoez gain, asteburuko bizipenak kontatzeko mugimendua adierazten duten *joan* («etxera nindoan») eta *etorri* («[lagunaren izena] etxera zetorren») aditzak erabili dituzte. LH4 mailan, ikasleek beren autobiografia idaztean, *eduki* aditza sarri erabili dute («bederatzi urte dauzkat», «begi marroiak dauzkat»), adina eta jabetza adierazteko. LH6 mailan, aldiz, ez da aditz trinko berririk erabili; aurrekoak erabili dituzte egoera eta argudioak azaltzeko.

Aditz perifrastikoen eta aspektuaren bilakaera ere nabarmena da LHn. LH1 mailan, irudien deskribapenak egiteko, aditz perifrastiko burutugabeak erabili dituzte gehienbat, pertsonaiak une horretan egiten ari diren ekintza jarraituak zehazteko («katua arrantzan ari da», «txakurra zukua edaten ari da»). LH2 mailan, ordea, asteburuan egindakoaren berri emateko aditz perifrastiko burutuak darabiltzate nagusiki («pelikula ikusi nuen», «dantza egin nuen»). LH4 mailako autobiografietan, iraganeko ekintzak gogoratze-ko aditz perifrastiko burutuak erabiltzeaz gain, orainaldikoak ere erabili dituzte, haien zaletasunak eta ohiturak adierazteko («igeri egiten dut»). LH6 mailako iritzi-artikuluetan, orainaldiko aditz perifrastikoez gain, etorkizuneko aditz-formak ere erabiltzen dituzte («denbora gehiegi pasatuko dute»).

3.2.3.2 DISKURTSOAREN ANALISIA: MODALITATEA ETA PERPAUS-ELKARKETA

LH1 eta LH2 mailetan ez da modalitatezko egiturarik ageri; lehenengo mailan ikasleak irudi baten deskribapen faktikora mugatzen dira eta bigarren mailan iraganeko gertaerak kontatzen dituzte, besterik gabe. LH4 mailan, ordea, garapen kualitatibo garrantzitsu bat gertatzen da, ikasleek beren autobiografietan desirak eta asmoak adierazten baitituzte. Etorkizuneko nahiak eta lanbideak proiektatzen dituzte: «handitan albaitaria izan nahi

nuke»². Ahalera («kirola egin dezaket») eta betebeharra adierazten duten egiturak ere agertzen dira («larrialdietara joan behar izan nuen»). LH6 mailan modalitatearen garapena are nabarmenagoa da, aukerak eta arriskuak aurkezteko ahalera-egiturak erabiltzen dituztelako («mugikorretik min asko egin daiteke»), baita ezintasuna islatzen dutenak ere («ezingo genien abisatu gurasoei»). Betebeharra adierazten duten egiturak ugaritzen dira («neurri bat hartu behar da») eta kausa-ondorio harremanak zein ondorio hipotetikoak ezartzen dituzte baldintzazko egituren bidez («mugikorra debekatzen badute, ezingo gara bata-bestearekin komunikatu»).

Perpaus-elkarketak bilakaera nabarmena du, eta idazten duten testu-motaren konplexutasunarekin batera garatzen da. LH1 mailan, irudiaren deskribapena egitean, alborakuntza erabiltzen dute («Marrazoa dago. Mutilak betaurrekoak ditu.»). LH2 mailan, aldiz, asteburuko gertaerak kronologikoki kontatzen dituzte, *eta* edo *gero* juntagailuak («atzo parkera joan nintzen eta gero izozkia jan nuen») eta helburuzko mendeko perpausak («baserrira joan nintzen han bazkaltzeko») baliatuz. LH4 mailako autobiografietan egitura konplexuagoak eratzen dira; denborazko menderakuntza-egiturak («txikia nintzenean burua ireki nuen») eta kausazkoak («triste nago datorren urtean bosgarren mailara joango garelako») txertatzen dituzte. Azkenik, LH6 mailako iritzi-artikuluetan perpaus-elkarketa landuagoa ageri da; argudiaketari eusteko kausa-ondorio harremanak eta ondorio hipotetikoak ezartzen dituzten baldintzazko menderakuntza-egiturak («ume bat bakarrik badago eta min handia hartzen badu, mugikorretik 122ra deitu dezake») agertzen dira.

3.2.3.3 KOHESIOA: IZENEN GAINJARTZE GLOBALA

LH1 eta LH2 mailetan, izenen gainjartze globalaren desbideratze estandarra nulua izan da, baina arrazoi ezberdinengatik. LH1 mailan, ikasleek irudiko elementuak zerrendatu dituzte («Marrazoa dago. Mutilak betaurrekoak ditu.»), eta esaldi bakoitzeko elementu berri bat aurkeztu dutenez, ia

² Adibideetan ikasleek idatzitako pasarteak hitzez hitz jaso dira.

ez da lemen errepikapenik hauteman eta gainjartze-balioek (ikus 3. eranskina) ez dute aldakortasunik erakutsi.

LH2 mailan, berriz, asteburuko bizipenen kontakizun kronologikoa egin dute; perpausak bata bestearen atzetik kateatu dituzten arren, gai eta tesuinguruak etengabe aldatu dira («larunbatean Azpeitian karrerako kotoxeak ikusi nituen eta gero baserrira joan nintzen bazkaltzera eta arratsaldean gimnasia erritmikoa ikustera joan nintzen»). Ondorioz, esaldien arteko lema-gainjartze balioak konstanteak izan direnez, izenen gainjartze globalaren desbideratze estandarra nulua izan da. Horrez gain, badira sistema automatikoak detekta ezin ditzakeen kohesio-elementuak: aditzaren elipsia («bonboiak eman zizkiguten eta ogia ere bai») eta subjektuaren elipsia («Larunbatean Leitzara joan nintzen [...] aspertu egin nintzen») egiten da.

LH4 mailan, autobiografiak izenen gainjartze globala baldintzatu du, ikasleek testua bloke tematikoetan antolatu baitute (aurkezpena, familia, iragana eta zaletasunak). Egitura horrek kohesio-matrizean balio altuagoak eta desbideratze estandar handiagoa eragin ditu; bereziki nabarmena da familia-kideen errepikapena, pertsona horiek kontakizunaren hainbat ataletan agertzen baitira. Hala ere, sistema automatikoak gainjartzea literalki neurtzen duenez, ez du testuaren kohesioa bere osotasunean neurtu: pertsonen dagokienez, hasieran «ama» aipatu badute eta aurrerago amaren izena erabili badute, sistema automatikoak ez du gainjartze hori neurtu. Sistemak detekta ezin ditzakeen beste kohesio-elementu batzuk bi hauek dira: aditzaren elipsia («nire amamak [amonaren izena] izena du eta nire aitak [aitaren izena]») eta *pro-dropa* («[herriaren izena] bizi naiz eta 4A gelan nago»).

LH6 mailan, iritzi-testuek ardatz tematiko nagusi eta egonkor bat dute-nez, gai horiei dagozkien lemak maiz errepikatu dira. Hala ere, argudiaketa blokeka antolatu denez, lemak bloke barruan errepikatu ondoren, hurrengo blokerak pasatzean, desagertu eta berriak agertu dira, eta puntuazioak jaitسي dira. Beste muga bat antzeman da maila honetan: sinonimoen erabilerak («mobila», «mugikorra», «telefonoa») elementu zentralen gainjartzea ezkutatatu du neurketa automatikoan. *Pro-dropa* («mugikorrak ez dira jokoetan jolasteko bakarrik [...] debekua jarrita ere askok erabiltzen jarraitzen dute

[...] baina hobea izango zen ez bazuten edukiko») ere agertzen da, eta sistema automatikoak neurtu ezin duenez, adierazleak ez du testuaren kohesioa bere osotasunean neurtu.

3.3 Adierazle linguistikoen eragina kalifikazioetan ikasmailaren arabera

3.3.1 Ezberdintasun estatistikoen analisia

Adierazle linguistikoek irakasleek emandako kalifikazioetan (oinarrizko maila, erdiko maila eta maila aurreratua) duten eragina esanguratsua den aztertzeko Kruskal-Wallis H proba aukeratu da, datuek ez dituztelako baldintza parametrikoak betetzen.

6. taulako emaitzek erakusten dutenez, hainbat kasutan efektu-tamaina handiak ($\eta^2 \geq 0,22$) detektatu dira, nahiz eta esangura-maila $p < 0,05$ mugatik gora egon. Hori bereziki nabarmena da LH1eko jariakortasunean ($\eta^2 = 0,27$) eta lexikoan ($\eta^2 = 0,22$), baita LH6ko jariakortasunean ($\eta^2 = 0,32$) eta syntaxian ($\eta^2 = 0,43$) ere. Esangura estatistikorik ezak, efektu-magnitude altu horiekin batera, iradokitzen du taldeen artean alde handiak egon arren, maila horietako lagin-tamaina txikiek ($N = 18$ eta $N = 15$) mugatu egin dutela alde horiek ziurtasun estatistikoz berresteko gaitasuna.

Hala ere, LH1 mailan ezberdintasun esanguratsuak aurkitu dira syntaxian ($H = 6,26$, $p = 0,044$) eta kohesioan ($H = 8,33$, $p = 0,015$), non efektuaren tamainak oso handiak diren ($\eta^2 > 0,3$). Alabaina, LH2 mailan lotura hori desagertu egin da dimentsio guztietan, esangura-mailek eta efektu-tamainek erakutsi duten bezala ($p > 0,05$, $\eta^2 \leq 0,01$). LH4 mailan, lexikoan ($p = 0,019$) eta syntaxian ($p = 0,004$) alde esanguratsuak atzeman dira, azken horren kasuan, efektuaren tamaina ertain-handia ($\eta^2 = 0,11$) izanik. Azkenik, LH6 mailan esangura galdu da dimentsio guztietan ($p > 0,05$), nahiz eta efektuaren tamaina gehienak handiak izan, arestian aipatu bezala.

6. taula. Kalifikazioen arteko aldeak adierazle linguistikoetan ikasmailaren arabera: Kruskal-Wallis H probaren emaitzak

Maila	Dimentsioa	KW (H)	p	Esang.	η^2
LH1	Jariakortasuna	5,519	0,063	Ez	0,271
	Lexikoa	4,892	0,087	Ez	0,222
	Sintaxia	6,26	0,044	Bai	0,328
	Kohesioa	8,333	0,015	Bai	0,487
LH2	Jariakortasuna	1,068	0,586	Ez	– 0,016
	Lexikoa	2,731	0,255	Ez	0,013
	Sintaxia	2,247	0,325	Ez	0,004
	Kohesioa	1,397	0,497	Ez	– 0,01
LH4	Jariakortasuna	5,081	0,079	Ez	0,038
	Lexikoa	7,877	0,019	Bai	0,072
	Sintaxia	11,02	0,004	Bai	0,11
	Kohesioa	2,751	0,253	Ez	0,009
LH6	Jariakortasuna	4,545	0,103	Ez	0,318
	Lexikoa	3,045	0,218	Ez	0,131
	Sintaxia	5,419	0,066	Ez	0,427
	Kohesioa	0,833	0,659	Ez	– 0,146

Iturria: Egileak prestatua ($N = 181$).

Oharra: KW (H): Kruskal-Wallis H estatistikoa. p : esangura-maila (esanguratsua $p < 0,05$ denean). *Esang.*: Bai (alde esanguratsua); Ez (alderik ez). η^2 : efektu-tamaina, non 0,01 (txikia), 0,06 (ertaina) eta 0,14 (handia) diren erreferentzia-balioak. Kasu batzuetan efektu-tamaina handia izan arren ($\eta^2 > 0,14$), esangura-maila ez da lortu ($p > 0,05$), batez ere lagin-tamaina murrizteneko mailetan.

Behin Kruskal-Wallis H probaren bidez alde orokorrak identifikatu ondoren, Mann-Whitney U *post-hoc* analisisia burutu da, ezberdintasun horiek zein taldetan gertatu diren zehazteko (ikus 7. taula). Emaitez erakutsi dutenez, Kruskal-Wallis H proban agertutako esangura gehienak lausotu egin dira bikote-analisan, Bonferroni zuzenketaren zorrotasunagatik.

LH1 mailan, syntaxian eta kohesioan atzemandako aldeak ez dira esanguratsuak izan bikoteen artean ($p_{adj} > 0,05$). Kruskal-Wallis H proba orokorrak alde esanguratsuak detektatu dituen arren ($p = 0,044$ eta $p = 0,015$), binakako Mann-Whitney U *post-hoc* analisisian esangura hori desagertu egin da ($p_{adj} > 0,05$). Litekeena da gertakari hori bi faktore nagusiren ondorio izatea: Bonferroni zuzenketaren zorrotasunagatik, aldi berean hainbat konparaketa egitean, zuzenketak p -balioaren muga asko estutzen duelako I. motako errorea (alfa errorea) saihesteko eta horrek esangura lortzea zailtzen duelako. Bestalde, potentzia estatistiko falta ere izan liteke gertakari horren kausa: LH1eko laginaren tamaina eta banaketa (oinarrizko mailako zazpi testu, erdiko mailako hamar eta maila aurreratuko bakarra) oso txikia eta desorekatua izanik, Kruskal-Wallis H proba gai izan arren taldeen arteko alde orokorrak antzemateko, Mann-Whitney U probak ez du nahikoa potentzia estatistikorik alde horiek bikote zehatzen artean atzemateko.

Era berean, LH4ko lexikoan, hasiera batean aldea esanguratsua bazen ere ($p = 0,019$), *post-hoc* analisisian ez da taldearteko alderik aurkitu, talde horretako lagina handiagoa izanik ere kalifikazio guztietan. Esangura LH4ko syntaxian mantendu da soilik eta kasu zehatz batzuetan bakarrik: oinarritzko mailaren (1) eta erdiko mailaren (2) artean ($p_{adj} = 0,005$) eta hasierako mailaren (1) eta maila aurreratuaren (3) artean ($p_{adj} = 0,042$). Aldiz, erdiko mailaren eta aurreratuaren artean (2 eta 3) ez dago alderik ($p_{adj} = 1$), beraz, sintaxiari dagokionez, LH4ko kalifikazioen arteko arrakala nagusia oinarritzko mailan dauden ikasleen eta gainerakoen artean dago (ikus 3. irudia).

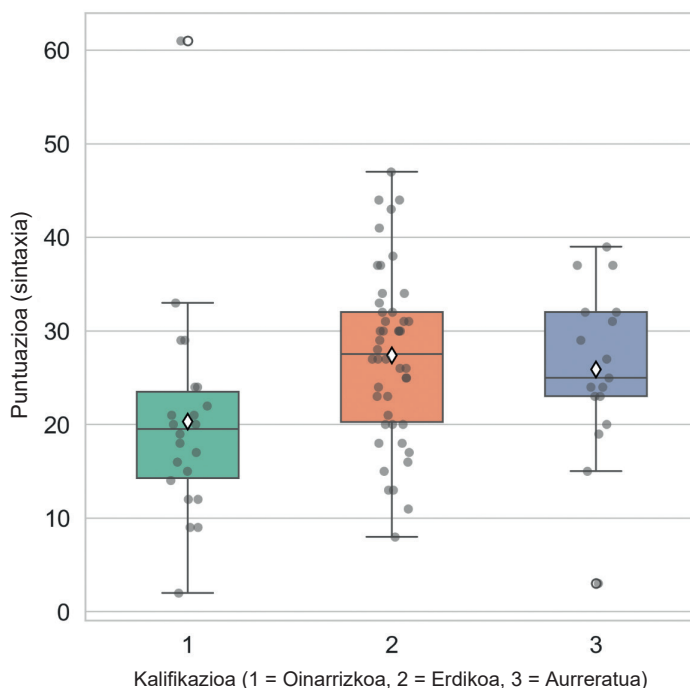
7. taula. Kalifikazioen arteko bikote-konparaketak (*post-hoc*) adierazle linguistikoetan eta ikasmaila bakoitzean: Mann-Whitney *U* probaren emaitzak

Maila	Dimentsioa	Konparaketa	<i>p</i> -adj	Esang.	<i>r</i> (Efektu-tamaina)	Kalifikazio altuena
LH1	Sintaxia	1 eta 2	0,12	Ez	0,517	-
		2 eta 3	0,82	Ez	0,385	-
		1 eta 3	0,584	Ez	0,567	-
	Kohesioa	1 eta 2	1	Ez	0,091	-
		2 eta 3	0,140	Ez	0,495	-
		1 eta 3	0,124	Ez	0,567	-
LH4	Lexikoa	1 eta 2	0,563	Ez	0,161	-
		2 eta 3	1	Ez	0	-
		1 eta 3	0,771	Ez	0,184	-
	Sintaxia	1 eta 2	0,005	Bai	0,385	2
		2 eta 3	1	Ez	0,052	-
		1 eta 3	0,042	Bai	0,395	3

Iturria: Egileak prestatua ($N = 181$).

Oharra: Kalifikazioak hiru mailatan banatu dira: 1 (oinarrizko maila), 2 (erdiko maila) eta 3 (maila aurreratua). Konparaketak Mann-Whitney *U* probaren bidez egin dira. *p*-adj: Bonferroni zuzenketaren bidez lortutako *p*-balio doitua ($p \times 3$). Esang.: Bai (alde esanguratsua, $p\text{-adj} < 0,05$); Ez (alderik ez, $p\text{-adj} \geq 0,05$). *r* (Efektu-tamaina): 0,1 (txikia); 0,3 (ertaina) eta 0,5 (handia). Kalifikazio altuena (puntuazio altuena lortu duen kalifikazioa) ere adierazita dago.

3. irudia. Syntaxiko puntuazioen banaketa kalifikazioen arabera LH4 mailan



Iturria: Egileak prestatua ($N = 181$).

Oharra: Kaxa-diagramek LH4ko ikasleek syntaxian lortutako puntuazioen banaketa erakusten dute kalifikazio-kategorien arabera: 1 (oinarrizko maila), 2 (erdiko maila) eta 3 (maila aurreratua).

3.3.2 Ezberdintasunen deskribapen kualitatiboa

Analisi estatistikoak erakutsi du LH4 mailan aditz-sintagmen kopurua neurtzen duen adierazlea dela testuak kalifikazioen arabera diskriminatzeko gaitasuna duen bakarra. Zehazki, aldea oinarrizko mailako ikasleen eta gainerako (maila ertaina eta aurreratua) artean hauteman da, baina ez maila ertainaren eta aurreratuaren artean. Emaita estatistiko horiek testuen azterketa kualitatiboaren bidez argitu eta osatu daitezke.

Aditz-mota kopuruaren adierazleak ez du kalifikazioen arteko bereizketa egiteko gaitasunik erakutsi; hala ere, modalitate-formek kalifikazioen arteko ezberdintasun kualitatiboak argiago erakuts ditzakete. Oinarrizko mailako testuetan modalitatearen erabilera oso mugatua da, deskribapen estatiko eta ekintza burutueta mugatu direlako ikasleak («ni bizi naiz [ikaslearen herrian]») eta nahia adierazteko saiakerak segmentazio-arazoezin (hitzak elkartuta idaztea) gertatu direlako («nire aitak nahizuen militarra izan»). Erdi mailako testuetan modalitatearen oinarrizko erabilera agertzen dira nahia edo asmoa adierazteko, nahiz eta akats morfosintaktikoak eta segmentazio-arazoak ohikoak izan oraindik («nahi konuke izan andereñoa»). Maila aurreratuko testuetan segmentazio-arazoak nabarmen murrizten diren arren, modalitate-adierazpenak egitura perifrastiko sinpleetara mugatzen dira («nik ez dut nahi joan»). Aitzitik, modalitate hipotetikoa adierazten duten *nuke*, *nezake* edo *banintz* bezalako formak oso urriak dira.

Aditz-sintagmen kopuruari dagokionez, analisi estatistikoak erakutsi du adierazle horrek kalifikazioen arteko alde esanguratsuak hauteman ditzakeela. Oinarrizko mailako testuetan aditz-sintagmen segidak gailendu dira; ikasleek egitura predikatibo sinpleak pilatu dituzte, euren arteko lotura sintaktiko gutxirekin («oraintxe bertan [ikaslearen herrian] bizi naiz [ikaslearen eskolan] ikasten dut 4D gelan nago»). Erdiko mailako testuetan, aldiz, aditz-sintagmen arteko koordinazioa justaposizioaren edo *eta* juntagailuaren bidez lortzen da («ni txikia nintzenean haur eskolara joan nintzen, han [lagunaren izena] ezagutu nuen eta nire lagunik onena izan zen»). Maila aurreratuko testuetan, berriz, aditz-sintagmen interakzio hierarkikoa nabarmendu da, menderakuntza-egituretan txertatutako egitura predikatiboen bidez, denborazko egituretan bereziki («eskolan hasi nintzenean arbelaren kontra jo nintzen»).

Azkenik, izenen gainjartze globalaren desbideratze estandarrek ez du erakutsi kalifikazioen artean diskriminatzeko gaitasun esanguratsurik. Bi fenomenok lagundu dezakete hori azaltzen: batetik, oinarrizko eta erdiko mailako testuetan, ortografia-akatsak direla eta, lematizazioa ez da egokia izan beti eta horrek kohesioaren neurketa zigortu du. Bestetik, maila aurreratuko testuetan desbideratze estandarra altuagoa izan arren, aldea ez da

estatistikoki esanguratsua izan: ikasleek bloke tematikoak egituratu dituzte (aurkezpena, familia, iragana eta zaletasunak), blokearen barruan lemak errepikatuz, baina hurrengo blokerak pasatzean aurreko lemak errepikatu ez direnez, gorabeherak sortu dira gainjartze-matrizean (ikus 3. eranskina). Aurretik aipatutako elipsiek ere (aditzaren elipsia: «nire amamak [amona-ren izena] izena du eta nire aitak [aitaren izena]» eta *pro-drop*a: «[herriaren izena] bizi naiz eta 4. mailan nago») agerian utzi dute adierazle honen muga: sistema automatikoak ez dituen ezagurri horiek neurtzen, kohesioaren zenbait elementu analisi kuantitatibotik kanpo geratu dira.

4 Eztabaida

Ikerketaren emaitzek LHko idazmenaren garapenari buruzko interpreta-zio zehatza ahalbidetu dute, eta, orokorrean, bat datoz sarreran aipatutako euskararen ezaugarri tipologikoekin. Jariakortasunaren, lexikoaren eta sin-taxiaren arteko korrelazio altuek adierazten dute hiru dimentsio horiek garapen-bloke integratu bat osatzen dutela; hau da, idatzizko ekoizpena handitzearekin batera, hizkuntza-aberastasuna ere areagotzeko joera dago. Hala ere, goranzko joera hori bereziki LH1 eta LH4 artean nabaritu da di-mentsio horietan, adierazleen emaitzak egonkortu egin baitira azken zikloan.

Kohesioak, aldiz, portaera independentea erakutsi du. Lehen zikloan ba-lio minimoak erregistratu dira, LH4 mailan gailurra jo du eta LH6 mailan beherakada nabarmena antzeman da. Beherakada horrek, ordea, ez du zer-tan kohesioaren ahultzerik adierazi; aitzitik, litekeena da ikasleek sinoni-moak eta elipsirako baliabideak erabiltzean, sistema automatikoak fenome-no horiek neurtu ez izana. Horrek Givónek (1983) eta beste autore batzuek azaldutakoarekin bat egiten du: euskararen *pro-drop* izaeraren ondorioz, argumentuen jarraipena sarritan aditzaren morfologian kodetzen da eta izen-sintagmen elipsia ohikoa da diskurtoan.

Emaitzek agerian utzi dute euskararen flexio-morfologiaren konplexuta-sunak (Euskaltzaindiaren Gramatika Batzordea 2021) analisi kuantitati-boaren interpretazioa baldintzatzen duela. Hitz-tokenetan oinarritutako

neurketek lexikoaren aberastasuna modu partzialean islatzeko arriskua dute, forma bakar batek aldaera morfologiko ugari izan baititzake (Aduriz *et al.* 2000). Ondorioz, lema-mailako adierazleak egokiagoak dira errepertorio lexikoa modu fidagarriagoan neurtzeko (McCarthy eta Jarvis 2010). Halaber, nabarmena da aditzak perpausaren egitura sintaktikoan duen funtzio zentral eta egituratzailea (Hualde eta Ortiz de Urbina 2003). Horrenbestez, aditz-formadun egituren analisia egokia izan da garapen sintaktikoa ebaluatzeko (Biber, Conrad eta Reppen 1998).

Bilakaeraren azterketa kualitatiboak erakutsi du aditz-errepertorioak eta perpaus-elkarketak garapen progresiboa izan dutela Lehen Hezkuntzan, testu-mota bakoitzaren konplexutasunarekin batera. LH1 mailan deskribapen-egiturak nagusitzen dira (*izan* eta *egon* aditzak, eta alborakuntza); LH2 mailan, asteburuko gertaeren kontakizun kronologikoak ageri dira, juntagailuen (*eta*, *gero*) eta aditz perifrastiko burutuen bidez; LH4 mailako autobiografieta egitura konplexuagoak garatzen dira: denborazko eta kausazko menderakuntza-egiturak, desioak adierazteko modalitate-egiturak eta aditzaren zein subjektuaren elipsia. Azkenik, LH6 mailako iritzi-artikulueta argudiaketa-egitura konplexuak nagusitzen dira, kausa-ondorio harremanak eta ondorio hipotetikoak adierazten dituzten baldintzazko menderakuntza-egituren bidez. Arestian aipatu bezala, kohesioaren neurketa automatikoak aldakortasun nabarmena erakutsi du, sinonimoen erabilera eta sistema automatikoak neurtu ezin ditzakeen elipsi zein *pro-drop* mekanismoen ondorioz.

Irakasleek emandako kalifikazioen artean diskriminatzeko gaitasuna erakutsi dute adierazleek, eta ondorioztatu da LH4 mailako sintaxiak soilik erakutsi duela gaitasun hori, zehazki, oinarritzko mailako ikasleen eta gainerakoen (maila ertaina eta aurreratua) artean. Litekeena da lagin-tamaina txikiek eta Bonferroni zuzenketaren zorrotasunak potentzia estatistikoa mugatu izana. Ildo berean, Kruskal-Wallis probaren emaitzek esangura orokorra adierazi badute ere, *post-hoc* analisietan ez da beti esangura hori mantendu, eta horrek, bikoteen arteko konparaketa zehatzetarako datu gehiago behar direla iradokitzen du. Ikuspegi kualitatiboak perpaus-elkarketaren bilakaera argi eta koherentea erakutsi du: oinarritzko mailetan lotura sintaktikorik gabeko egitura predikatibo sinpleen pilaketa eta albo-

rakuntza nagusitzen dira, erdiko mailan juntadurak ageri dira eta maila aurreratuko testuetan menderakuntza-egiturak garatzen dira.

Azkenik, ikerketa honen zenbait muga metodologiko eta estatistiko azpimarratu behar dira. Batetik, baliteke LH1 eta LH6 mailetakoin tamaina txikiek analisi estatistikoen potentzia murriztu izana; ondorioz, Kruskal-Wallis proban islatutako zenbait alde orokor *post-hoc* analisisetan lausotu egin dira Bonferroni zuzenketaren zorroztasunagatik. Bestetik, maila bakoitzean testu-mota desberdinak erabiltzeari lotutako heterogeneotasunak emaitzen konparagarritasuna baldintza dezake. Hori ikerketadiseinuaren berezko muga da, adin desberdinetako haurren garapen kognitiboak eta linguistikoak testu-eskaera ezberdinak justifikatzen baitituzte. Kohesioaren kasuan, frogatu da erabilitako adierazle automatikoak (izenen gainjartze globalaren desbideratze estandarra) ezin dituenenez zenbait kohesio-elementu detektatu (hala nola sinonimoen eta elipsirako baliabideen erabilera), neurketa kuantitatiboak kohesioa bere osotasunean neurtu ez izana puntuazio baxua eragin du. Aipatu beharra dago ezin izan dela CAF (*Complexity, Accuracy, Fluency*) markoa bere osotasunean aztertu, ez baita oraindik zehaztasuna (*accuracy*) neurtzeko tresna automatiko fidagarriarik garatu euskararako.

5 Ondorioak

Ikerketa honek Lehen Hezkuntzako idazmenaren garapena aztertu du, CTAP tresna eta CAF (*Complexity, Accuracy, Fluency*) markoa euskararen berezitasun tipologikoetara (flexio-morfologia, aditzaren zentralitatea eta *pro-drop* izaera) egokituz. Lortutako emaitzek erakutsi dute ebaluazio automatizatuaren ikuspegitik, euskara bezalako hizkuntza eranskari eta morfologikoki konplexuetan, hitz-tokenetan oinarritutako metrikek lexikoaren erabileraren irudi partziala edo distortsionatua eman dezaketela. Testuinguru horretan, lema-mailako adierazleak erabiltzea eta aditz-formadun egituren analisi kuantitatibo nahiz kualitatiboa uztartzea estrategia egokiak eta fidagarriak dira ikasleen hizkuntza-aberastasuna eta garapen sintaktikoa ebaluatzeko.

Gaitasun sintaktikoaren garapenari dagokionez, emaitza kuantitatiboek eta kualitatiboek erakutsi dute kalifikazio-mailak bereizteko gaitasun estatistiko esanguratsua LH4ko sintaxi-adierazlean soilik identifikatu dela. Azterketa kualitatiboak, gainera, bilakaera progresibo argia erakutsi du: oinarrizko mailako lotura sintaktikorik gabeko egitura predikatibo sinpleen pilaketatik Erdiko mailako koordinazio-egituretara eta maila aurreratuko menderakuntza-egituretara igaro dira ikasleak.

Kohesioaren neurketari dagokionez, emaitzek adierazle automatikoaren zenbait muga metodologiko azaleratu dituzte. Berezik, izenen gainjartze globalean oinarritutako metrikek ez dituzte behar bezala jasotzen kohesioa eraikitzeke erabiltzen diren hainbat baliabide diskurtsibo, hala nola elipsia, *pro-drop* mekanismoak eta sinonimoen erabilera. Ondorioz, neurketa kuantitatiboek ez dute kohesio-maila bere osotasunean islatu, eta hori puntuazio baxuetan antzematen da.

Inplikazio pedagogikoei dagokienez, emaitzek iradokitzen dute aditz-egituren, perpaus-elkarketen, modalitatearen eta menderakuntza-egituren garapenak lotura estua duela ikasleen idazmenarekin. Horrenbestez, alderdi horiek era sistematikoan lantzea eraginkorra izan daiteke idazmena garatzeko.

Azkenik, etorkizuneko ikerketek lagin handiagoak hartu beharko lituzkete kontuan, Lehen Hezkuntzako ikasmaila guztiak barne hartuz, eta, ahal den neurrian, testu-mota homogeenagoak erabiliz. Era berean, beharrezkoa da kohesioa modu integralagoan neurtzeko adierazle berriak garatzea, baita euskaraz zehaztasuna (*accuracy*) modu fidagarrian ebaluatzeke tresna automatiko espezifikoak diseinatzea ere. Horrela, CAF (*Complexity, Accuracy, Fluency*) markoaren hiru dimentsioak osotasunean azter litezke, euskarazko idazmenaren ebaluazio automatizatuan aurrerapauso esanguratsuak emateko.

Bibliografia

ADURIZ, Itziar, AGIRRE, Eneko, ALDEZABAL, Izaskun *et al.* (2000). «A word-grammar based morphological analyzer for agglutinative languages», in *Proceedings of the 18th Conference on Computational Linguistics (COLING 2000)*. Association for Computational Linguistics: 3-9. <https://doi.org/10.3115/990820.990821>

AGERRI, Rodrigo, BERMUDEZ, Josu eta RIGAU, German (2014). «IXA pipeline: Efficient and ready to use multilingual NLP tools», in *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*. European Language Resources Association (ELRA): 3823-3828. <https://doi.org/10.6317/2b6q4azr2a9a>

ARRIOLA, Jose Mari, GOIKOETXEA, Josu eta IRUSKIETA, Mikel (2020). «Hizkuntza-teknologiak hizkuntzen ikas-irakaskuntzan: zenbat aukera, hainbat erronka», in *Hizpide* 95: 35-50. <https://doi.org/10.54512/LOUF1436>

AZURMENDI, Ekhi, ARREGI, Xabier eta LÓPEZ DE LACALLE, Oier (2025). «Automatic essay scoring and feedback generation in Basque language learning», *arXiv*. <https://doi.org/10.48550/arxiv.2512.08713>

BIBER, Douglas, CONRAD, Susan eta REPPEN, Randi (1998). *Corpus linguistics. Investigating language structure and use*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511804489>

BULTÉ, Bram eta HOUSEN, Alex (2012). «Defining and operationalising L2 complexity», in *Dimensions of L2 performance and proficiency dimensions of L2 performance and proficiency*. John Benjamins Publishing Company. <https://doi.org/10.1075/llt.32.02bul>

CHEN, Xiaobin eta MEURERS, Detmar (2016). «CTAP: A web-based tool supporting automatic complexity analysis», in *Proceedings of the 26th International Conference on Computational Linguistics: System demonstrations (COLING 2016)*. The International Committee on Computational Linguistics. <https://doi.org/10.17863/CAM.39630>

EUSKALTZAINDIA. (2014). *Lexikoaren Behatokia*. <https://lexikoarenbehatokia.euskaltzaindia.eus/>

EUSKALTZAINDIAREN GRAMATIKA BATZORDEA. (2021). *Euskararen Gramatika*. Euskaltzaindia.

GIVÓN, Talmy, HINDS, John, GASSER, Michael, *et al.* (1983). *Topic continuity in discourse: A quantitative cross-language study*. John Benjamins. <https://doi.org/10.1075/tsl.3>

HUALDE, José Ignacio eta ORTIZ DE URBINA, Jon (2003). *A Grammar of Basque*. Mouton de Gruyter.

IXA TALDEA. (2022). *IXA taldea A motako ikertalde finkatua. Eusko Jaurlaritz*a. <https://www.ixaeus/node/13571>

MCCARTHY, Philip eta JARVIS, Scott (2010). «MTLD, Vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment», *Behavior Research Methods* 42 (2): 381-392. <https://doi.org/10.3758/BRM.42.2.381>

MCKINNEY, Wes (2010). «Data structures for statistical computing in Python», in *Proceedings of the 9th Python in Science Conference*. SciPy: 56-61. <https://doi.org/10.25080/majora-92bf1922-00a>

MCMAMARA, Danielle, CROSSLEY, Scott eta MCCARTHY, Philip (2010). «Linguistic features of writing quality», *Written Communication* 27: 57-86. <https://doi.org/10.1177/0741088309351547>

READ-COOP SCE. (2015). *Transkribus*. <https://www.transkribus.org/> (2024/07/10ean kontsultatua).

VIRTANEN, Pauli, GOMMERS, Ralf, OLIPHANT, Travis, *et al.* (2020). «SciPy 1.0: fundamental algorithms for scientific computing in Python», *Nature Methods* 17 (3): 261-272. <https://doi.org/10.1038/s41592-019-0686-2>

WASKOM, Michael (2021). «Seaborn: Statistical data visualization», *Journal of Open Source Software* 6 (60): 3021. <https://doi.org/10.21105/joss.03021>

Eranskinak

1 LH4ko testuak ebaluatzeko errubrika

ERRUBRIKA - LH4			
Irizpideak	Oinarrizko maila	Erdiko maila	Maila aurreratua
1. Egokitasuna			
Helburua (% 10)	Ez du eskatutako testu-mota ekoiztu.	Eskatutako testu-mota ekoiztu du (nahiko sinplea).	Eskatutako testu-mota egoki ekoiztu du.
Lerro-kopurua			
Grafia (% 10)	Grafia irakurtezina da. Hizki larri eta xeheak nahasten ditu. Testuak zirrimarrak ditu edota egiten dituzten berridazketak ezin irakur daitezke.	Grafia ez da oso argia, irakurketa zailtzen du. Testuak zirrimarrak edota berridazketak ditu, irakurketa zailtzen dutenak.	Grafia argia da. Testuak ez du zirrimarrarik edota berridazketarik.
2. Koherentzia eta kohesioa			
Egitura eta lokailuak (% 10)	Ez ditu ideiak paragrafoetan antolatzen. Testuaren egitura ez da argia. Ez dago oinarrizko antolatzaile edota lokailurik edo gaizki erabiltzaile daude.	Ideiak paragrafoetan antolatu ditu. Testuaren egitura nahiko argia da. Badira oinarrizko testu-antolatzaileak.	Ideia ezberdinak paragrafoetan antolatzen ditu. Testuaren egitura nabaria da. Lokailuak eta testu-antolatzaile desberdinak erabiltzen ditu.
Informazioa/ edukia (% 15)	Edukiak ez du beti eskatutakoarekiko loturarik edota pobrea da.	Informazioa eskatutakoari egokitzen zaio.	Informazioak eskatu denarekin lotura dauka eta aparteko osagaiak gehitzen ditu.

3. Zuzentasuna			
Hiztegia/ lexikoa (% 10)	Hiztegi mugatua eta hitz-errepikapen ugari.	Hiztegi nahiko mugatua, hitzak errepikatzen dira batzuetan.	Hiztegi zabala erabiltzen du eta ez ditu hitzak errepikatzen.
Puntuazioa: puntuazio- marken erabilera egokia. Testuan puntuak erabili ditu (% 10)	Puntuazio-akats dezente daude.	Puntuazio-markak ondo erabiltzen ditu orokorrean.	Puntuazio-markak (puntu, koma...) ondo erabiltzen ditu.
Ortografia (% 10)	Akats ortografiko nabariak (rr/r...). Hainbat hizki norabidez aldatzen ditu. Hainbat hizki ispilu gisa idatzita daude. Ez ditu hainbat hizki idazten.	Akats ortografikoak ageri dira, baina ez hitz arruntetan. Hitzak elkarren arteak ongi bereizten ditu, baina aditzak ez ditu banatzen. Batzuetan bakarrik errespetatzen ditu hizki larriak eta xeheak.	Ortografia-arauak ezagutu eta errespetatzen ditu. Hitzak ongi bereizten ditu. Esaldi-hasieran eta puntu ondoren, larriz. Izen propioetan, lehen hizkia larriz. Erakusleetan «h». Aditzetan «z».
Morfosintaxia (% 15)	Oinarrizko akats ugari. Esaldiak gaizki osatuta. Lotu gabeko perpauak.	Akats batzuk ageri arren, testua ulergarria da. Esaldiak nahiko ondo osatuta daude.	Ez dago ia akatsik. Ondo osatutako esaldiak daude.
Hizkuntza- erabilera/ erregistroa (% 10)	Hizkuntza ez da egokia, lagunartekoa eta ahozko hizkuntzarekin adierazita. Erregistroak nahasita.	Hizkuntza egokia erabiltzeko saiakera egin du. Hala ere, ahozko hizkuntzari lotutako ezaugarriren bat edo beste ageri da batzuetan.	Hizkuntza egokia erabili da (lagunarteko edo ahozko hizkuntza eta euskara batua desberdindu ditu).

2 LH6ko testuak ebaluatzeko errubrika

ERRUBRIKA - LH6			
Irizpideak	Oinarrizko maila	Erdiko maila	Maila aurreratua
1. Egokitasuna			
Helburua	Ez du eskatutako testu-mota idatzi. Emandako informazioa oso eskasa da (hiztegia/lerro-kopurua). Hizkuntza ez da egokia, lagunartekoa edo ahozko hizkuntzarekin adierazia.	Eskatutako testu-mota ekoiztu du, baina nahiko sinplea da.	Narrazioa edo deskribapena da.
Lerro-kopurua			
Grafia	Grafia irakurtezina da. Letra larri eta xeheak nahasten ditu. Testuak zirrimarrak ditu edota egiten dituzten berridazketak ezinezkoak dira irakurtzen.	Grafia ez da oso argia, irakurketa zailtzen du. Testuak zirrimarrak edota berridazketak ditu, irakurketa zailtzen dutenak.	Grafia argia da. Testuak ez du zirrimarra edo berridazketarik.
2. Koherentzia eta kohesioa			
Egitura eta lokailuak	Ideiak ez daude era egokian antolatuta. Ez du lokailurik eta testu-antolatzailearik.	Ideiak paragrafoetan antolatzen ditu. Ez du antolatzaile eta lokailu egokirik erabiltzen edo behin eta berriz berdinak erabiltzen ditu.	Ideia ezberdinak paragrafoetan egoki antolatzen ditu. Lokailuak eta testu antolatzaile desberdinak erabiltzen ditu.
Informazioa/ edukia (% 15)	Edukia ez dago beti eskatutakoarekin lotuta eta pobrea da.	Informazioa eskatutakoari egokitzen zaio.	Informazioak eskatu denarekin lotura dauka eta aparteko osagaiak gehitzen ditu.

3. Zuzentasuna			
Hiztegia	Hiztegi mugatua eta hitz-errepikapen ugari.	Hiztegi mugatua.	Hiztegi zabala erabiltzen du eta ez ditu hitzak errepikatzen.
Puntuazioa: puntuazio-marken erabilera egokia. Testuan puntuak erabili ditu (% 15)	Puntuazio-marka gutxi erabiltzen dira eta akatsak nabari dira.	Puntuazio-markak egoki erabiltzen dira, nahiz eta akats batzuk ageri.	Puntuazio-markak (puntu, koma...) ondo erabiltzen ditu.
Ortografia: grafiaren erabilera egokia (% 15)	Akats ortografiko ugari.	Akats ortografikoak ageri dira, baina ez hitz arruntetan.	Ortografia-arauak ezagutu eta errespetatzen ditu.
Hizkuntzaren erabilera	Hizkuntza erregistroan lagunarteko diskurtsoa ageri da.	Ahozko zein idatzizko erregistroa nahasirik ageri da.	Erregistro formala, egokia.

3 CTAP tresna euskarazko testuetara egokitze prozedura

Jatorrizko CTAP tresnak (Chen eta Meurers 2016) sei hizkuntza onartzen ditu, eta, euskara ez dagoenez integratuta, beste ikerketa-proiektu batean garatutako euskarazko bertsioa erabili da lan honetarako (IXA Taldea 2022). Adierazle guztiak kalkulatzeko, aurreprozesamendu estandar bat jarraitu da. Lehenik, NLPAnnotator.xml arduratu da testu garbiak jaso eta sistemak uler ditzan egokitzeaz. Segidan, testuak esaldietan banatu dira SentenceAnnotator.xml erabiliz. Jarraian, testuak tokenizatu egin dira TokenAnnotator.xml bidez. Beharrezkoa denean, POSAnnotator.xml arduratu da etiketatze morfosintaktikoaz, hitz bakoitzari bere kategoria gramatikala esleitzuz. CTAP tresnak prozesu-kate (*pipeline*) hori kudeatzeko AAE (*Aggregate Analysis Engine* edo Analisi-Motor Bateratua) erabili du (Chen eta Meurers 2016), testuaren inguruko datu linguistikoak sortuz. AAE ez dagoenez diseinatuta euskarazko elementu linguistikoak bereizteko, IXA Taldeak garatutako IXAPIPES (Agerri, Bermudez eta Rigau 2014) integratu da. Datu linguistiko horiekin, CTAP tresnak (Chen eta Meurers 2016) gainontzeko hizkuntzetan neurtzen dituen adierazle linguistikoak euskaraz kalkulatu ditu, 201 guztira. Jarraian, lan honetarako aukeratu diren adierazle linguistikoak zehazki nola kalkulatzaren diren azaltzen da:

1. Hitz-token kopurua, puntuazio-markak eta zenbakiak baztertuz (*Number of Word Tokens, excluding punctuation and numbers*): adierazle horrek arestian aipatu den aurreprozesamendu estandarren ondoren, POSAnnotator.xml erabili du testuko puntuazio-markak eta sinboloak identifikatu eta baztertzeko. Jarraian, hitz-kategoria duten elementuak soilik zenbatu dira.
2. Aditz-moten kopurua, modalak barne (*Number of POS Feature, Verb Types, including Modals*): adierazle horrek ere aurreprozesamendu estandarra jarraitzen du. Jatorrizko tresnak beste hizkuntzetako berariazko aditz-etiketak baliatzen dituenaz, IXAPIPES (Agerri, Bermudez eta Rigau 2014) analizatzaileak *Universal Dependencies* (UD) deituriko nazioarteko estandarra erabiltzen du, hizkuntzen egitura gramatikala eta sintaktikoa modu uniformeaz etiketatzen dituenaz. Euskaraz aditz modal isolaturik ez dagoenez eta funtzio hori aditz laguntzaileen

flexioen bidez bideratzen denez, UD estandarreko bi etiketa nagusi erabili dira: VERB (aditz trinkoak, mendekoak nahiz konposatuak eta partizipioak) eta AUX (aditz laguntzaileak eta haien flexioaren bidezko modalak). Sistemak POSAnnotator.xml erabili du etiketa horiek iragazi eta zenbatzeko, non VERB etiketak aditzaren forma lexiko nagusiak biltzen dituen eta AUX etiketak euskarazko aditz-sistemako elementu funtzionalak, horien artean, modalitatea adierazten duten formak (ahalera eta subjuntiboa).

3. Aditz-sintagmen kopurua (*Number of Syntactic Constituents, Verb Phrase*): adierazle hori kalkulatzeko, aurreprozesamendu estandarren ondoren, ParseTreeAnnotator.xml erabili da zuhaitz sintaktikoak sortzeko eta horietatik aditz-sintagmei dagozkien adarrak isolatu eta zenbatzeko.
4. Izenen gainjartze globalaren desbideratze estandarra, lemetan oinarritua (*Referential Cohesion, Standard Deviation of Global Noun Overlap, lemma-based*): testuko kontzeptuak zein neurritan errepikatzen diren kalkulatzeko, lema bera esaldi ezberdinetan zenbat aldiz agertzen den neurtu da. Lema-mailako neurketa bereziki garrantzitsua da euskararen kasuan, bere izaera eranskaria eta deklinabide-sistema direla eta (McCarthy eta Jarvis 2010). Prozesu horretan, aurreprozesamendu estandarren ondoren, POSAnnotator.xml arduratu da izenak bakarrik iragazteaz eta LemmaAnnotator.xml erabili da izen horien lema edo erroa ateratzeko. Jarraian, CTAP tresnak (Chen eta Meurers 2016) gainjartze-matrize bat osatu du; adibidez, lehen esaldiak bigarren eta hirugarren esaldiarekin lema bat partekatzen badu, esaldi horren puntuazioa 2 izan da. Segidan, puntuazio guztien arteko batezbestekoa kalkulatu da, eta amaitzeko, testu osoaren desbideratze estandarra. Emaizta baxua bada, testuko esaldi guztiek antzeko puntuazioa dutela esan nahi du eta testuaren kohesioa orekatua izan dela adierazten du horrek; are gehiago, errepikapenik ezean ere (balio konstantea mantentzen denean) emaitza zero da. Emaizta altua bada, berriz, esan nahi du esaldi batzuk oso konektatuta daudela eta beste batzuk ez daudela batere konektatuta. Konektibitate-gorabehera horiek kohesioaren ezegonkortasuna islatzen dute, eta horrek irakurleari testua ulertzea zailtzen dio, gaia etengabe zentratu eta aldatzen delako.