

# IGARRITZ: euskarazko testu-iragarpenerako web-ingurune egokituak

(IGARRITZ: adapted web environment for text prediction in Basque)

Mikel Iruskieta\*, Iker de la Iglesia, Unai Atutxa  
HiTZ-Ixa taldea (UPV/EHU, Donostia, Gipuzkoa)

Lierni Ortiz

Lekeitio BHI

**LABURPENA:** Motrizitate mugatua duten ikasleek —garun-paralisi baten ondorioz mugaren bat dutenek, adibidez— tresna egokituak izaten dituzte testuak idazteko; esaterako, begiradaren jarraipenerako hardware bat, zeinarekin ordenagailuan letrak aukeratu eta sistemak iragartzen dituen hitzak aukeratu baitaitezke. Halako sistemek euskaraz idazteko baliabideak izaten dituzte, edo euskarazko hitz zerrendak sartuta iragarpenak aukeratzeko abagunea ematen dute. Edozein testu-iragarpenen xedea zera da, testua idazteko esfortzua murriztea, baita testu luzeagoak azkarrago idatzi ahal izatea ere. Irisarekin idaztea hamar atzamarrekin idaztea baino geldoa eta nekezagoa da; horrez gain, testu-iragarpenerako tresnak beste hizkuntzetan baino okerrago ibiltzen dira euskaraz, eta, sarritan, nekez laguntzen dute. Lan honen helburua honako hau da: euskarazko testu-iragarpenerako web-ingurune egokitu bat egitea adimen artifizialeko teknikak erabiliz. Horretarako, euskarazko hizkuntza-ereduetan oinarrituriko web-interfaze bat sortu dugu, IGARRITZ izenekoa. Sortutako testu-iragarleak *Transformer* arkitektura erabiltzen du, eta HiTZ/roberta-eus-euscrawl-base-cased hizkuntza-eredua berrentrenatu dugu hezkuntzarako corpusarekin (ikasleen testuak, Gizapedia entziklopedia eta hezkuntzari buruzko testuak: Wikipediakoak eta *Berriakoak*). Bukatzeko, tresna ebaluatu eta egun eskura dagoen beste sistema batekin konparatu dugu, garun-paralisia duen Bigarren Hezkuntzako ikasle batek ekoizitako testuekin. Emaitzen arabera, IGARRITZek euskarazko testu-iragarpena hobetu egiten du, eta begiradarekin idazten duen pertsonak adierazi du berak askoz errazago eta gehiago idazten duela. Horrez gain, ebaluazio automatikoa egin dugu, eta hor ere emaitzak hobetzea lortu dugu.

**HITZ GAKOAK:** Testu-iragarpena, begiradaren jarraipena, euskara, hezkuntza, adimen artifiziala, hizkuntza-ereduak

**ABSTRACT:** Students with limited mobility, for example, those caused by brain paralysis, have adapted tools for writing texts, such as eye-tracking hardware, to select letters and predict words. For instance, they can use eye-tracking hardware to select letters and choose words predicted by the system. These systems offer resources for writing in Basque, and predictions can be customized by inputting Basque word lists. The primary aim of text prediction is to alleviate the effort involved in typing and to facilitate faster or increased text production. However, writing with Iris is slower and more challenging compared to conventional typing with ten fingers. Furthermore, predictive text functionality in Basque is comparatively less effective than in other languages, offering minimal quality output. Thus, the objective of this study is to develop an adapted web environment for Basque text prediction employing artificial intelligence techniques. To achieve this goal, we have developed a web interface named IGARRITZ based on the HiTZ/roberta-eus-euscrawl-base-cased language model, utilizing a Transformer architecture. It was re-trained with an educational Basque corpus sourced from student texts, educational texts from Gizapedia, Wikipedia, and Berria. Finally, we evaluated the tool and compared it with another currently available system using texts produced by a secondary school student with cerebral palsy. The results indicate that IGARRITZ enhances text prediction in Basque. The student, who writes using eye-tracking technology, reported that the writing process has become significantly easier and more efficient. Additionally, our automatic evaluation demonstrated improved results compared to the existing system.

**KEYWORDS:** text prediction, eye tracking, Basque, education.

\* **Harremanetan jartzeko / Corresponding author:** Mikel Iruskieta, HiTZ Hizkuntza Teknologiko Euskal Zentroa-Ixa, Euskal Herriko Unibertsitatea (UPV/EHU), Sarriena auzoa, z/g (48091-Leioa, Euskal Herria). – mikel.iruskieta@ehu.eus – https://orcid.org/0000-0002-6121-3902

**Nola aipatu / How to cite:** Iruskieta, Mikel; De la Iglesia, Iker; Atutxa, Unai; Ortiz, Lierni (2025). «IGARRITZ: euskarazko testu-iragarpenerako web-ingurune egokituak». *Ekaia*, ale berezia 2025, 29-46. (https://doi.org/10.1387/ekaia.26214).

Jasotze-data: 2024, maiatzak 13; Onartze-data: 2024, ekainak 28.

ISSN 0214-9001 - eISSN 2444-3255 / © UPV/EHU Press 2025



Lan hau Creative Commons Aitortu-EzKomertziala-PartekatuBerdin 4.0 Nazioartekoa lizentzia baten mende dago

## 1. SARRERA

Mugikortasun mugatua duten eta goiko gorputz-adarrak mugitzeko zailtasunak dituzten ikasleek zenbait baliabide behar izaten dituzte eskolako lanak egiteko eta curriculumak garatzeko. Mugimendurako zailtasunez gain ahozko komunikaziorako arazoak ere baldin badituzte, komunikatzen eta irakurketa-idazketan laguntzen dien gailu teknologikoren bat behar izaten dute, dela testuak irakurtzeko, dela pentsamenduak eta ezagutzak idazketaren edo beste komunikazio-sistema alternatibo baten bidez adierazteko. Ahozko komunikazioa erabili ezin denean, halaberrez erabili behar dira komunikazio-sistema handigarriak edota ordezkioak (KSHO) [1]. KSHO sistema motak bitan sailka daitezke [2, 11]: i) laguntza gabeko komunikazio-sistemak (esku-alfabetoa eta zeinu-hizkuntza, besteak beste), eta ii) komunikazio-sistema lagunduak (komunikazio-softwareak).

Esandakoarekin lotuta, psikomotrizitateari eragiten dion garun-paralisia duten DBHko ikasleentzat oso zaila da euskaraz testu luzeak sortzea, ohiko testu-editoreen iragarpenek edo hiztegieta oinarritutako testu-iragarleek produkzio-aukera txikia eskaintzen dutelako. Besteak beste, euskaraz hizkuntza eranskari izatea zailtasun gehitua da, eta, ondorioz, hitz-erro bakararekin hainbat hitz osa daitezkeenez, zaildu egiten da testu-iragarpeneko tresnen lana. Desgaitasun hori duten ikasleek tresna egokituak izaten dituzte testuak idazteko; esaterako, begiradaren jarraipenerako programekin, letrak aukeratu eta sistemak iragartzen dituen hitzak aukeratu ditzakete. Hala ere, ordenagailua begiradaren jarraipenarekin erabiltzeak badu zailtasun nagusi bat: prozesu motela da begiekin letraz letra idaztea, eta erabili behar duen pertsonak ahalegin handia egin behar du testu labur bat idazteko. Beraz, idazketa azkartzeko eta, aldi berean, ahalegina txikiagotzeko, aterabide aproposa izan daiteke hitzak auresateko sistema bat erabiltzea. Izan ere, horrelako sistemen bidez, teklatura gutxiagotan sakatzu idatz daitezke hitzak. Idazteko laguntza horiek mugikorrek, testu-prozesadore batzuek eta abarrek dituztenen antzekoak dira.

Lan honetan, DBHko 4. mailako ikasle batekin lan egin dugu. Garun-paralisia izanik, begiradaren jarraipenarekin ordenagailua erabiltzeko teknologia erabiltzen du (*eye-tracking*),<sup>1</sup> eta honako baliabide hauek baliatu ditzake: i) Grid3, Easyclick, edota ii) AsTeRICS-Grid [3] (Espainiako estatuko beste unibertsitate batzuek sortutako sistema, esaterako). Aukera horien artean, ikasle horrek Grid3 darabil ikastetxean, eta begiradaren jarraipenera egokitutako programa horren laguntzaz maneia dezake ordenagailua, pantailako teklatura birtualaren bidez: fitxategiak ireki edo ixteko, eta testuak idazteko. Horrelako sistemek izaten dituzte euskaraz idazteko baliabideak, edo, euskarazko hitz-zerrendak sartuta, iragarpenak aukeratzeko

---

<sup>1</sup> Irisbond edo Tobii sistemak ordenagailua irisarekin erabiltzeko baliabideak dira.

modua ematen dute. Alabaina, testu-iragarleak beste hizkuntzetan baino okerrago ibiltzen dira euskaraz, eta nekez laguntzen dute edozein testu-iragarlek duen xedea erdiesten; alegia, testua idazteko esfortzua murriztea eta testu gehiago edo azkarrago idaztea.

Egun, ikasle horrek ez du euskarazko idazkera arin eta eroso errazten dion sistema eraginkorrik; izan ere, idazteko dituen zailtasunek eta idazteko abiadurak testuen kalitateari ere eragiten diote, ahalegin handia egin behar duenez joera baitu esaldi laburragoak edota gutxi landuak erabiltzeko. Berbarako, asmatutako gertaera bat idazteko eskatu diogunean, 11 minutu eta 20 segundoan egitura sinpleko bi esaldi idatzi ditu; 19 hitz (103 letra)<sup>2</sup> Esfortzua txikiagotuz gehiago eta azkarrago idazteko aukera eskaintzeaz gain, iragarpen-sistema egoki batek tekleatze- edota ortografia-akatsak saihesten ditu, eta horrek ere ikaslearen komunikazioari mesede egingo dio. Bestalde, uste dugu euskarazko kalitatezko testuak aurreratzeko sistema bat izateak aurrerapena ekarriko lukeela, aipatutako ikasle horientzat ez ezik baita antzeko egoeran leudekeen beste ikasle askorentzat ere, teklaturaren erabiltzeko zailtasunak dituzten ikasleek eskura izan baitezakete idazketa erraztu eta bizkortzeko sistema bat.

Lan honen helburua honako hau da: euskarazko testu-iragarpenerako eta irisarekin idazteko programekin bateragarria den web-ingurune egokitu bat egitea, adimen artifizialeko teknikak erabiliz. Horretarako, euskarazko hizkuntza-ereduetan oinarrituriko web-interfaze bat sortu dugu, IGARRITZ izeneko. Sortutako testu-iragarleak *Transformer* arkitektura erabiltzen du [4]; zehazki, RoBERTa kodifikatzaile-arkitekturan [5] oinarritutako HiTZ/roberta-eus-euscrawl-base-cased hizkuntza-eredua [6] erabili dugu. Hizkuntza-eredu hori erabili dugu euskarazko hitzen distribuzioa eta hitzen arteko erlazioak ikasteko, euskaraz osatutako corpus handi batekin entrenatu baitugu aurretik. Aurretiko entrenamendu (*pre-trained*) horretan ikasitako erregistroak askotarikoak dira, eta DBHn erabiltzen direnez bestelakoak. Beraz, hizkuntza-eredua berregokitzeko eta iragarpen-hitzak ikaslearen beharrekin ahalik eta hobekien bat etortzeko, eredua berentrenatu dugu hezkuntzari buruzko corpusarekin (Gizapedia eta hezkuntzari buruzko testuak: Wikipediakoak eta *Berriakoak*), zeina berariaz bildua eta egokiagoa baita, etengabeko aurretiko entrenamendua (*continual-pretraining*) deritzon teknika erabiliz.

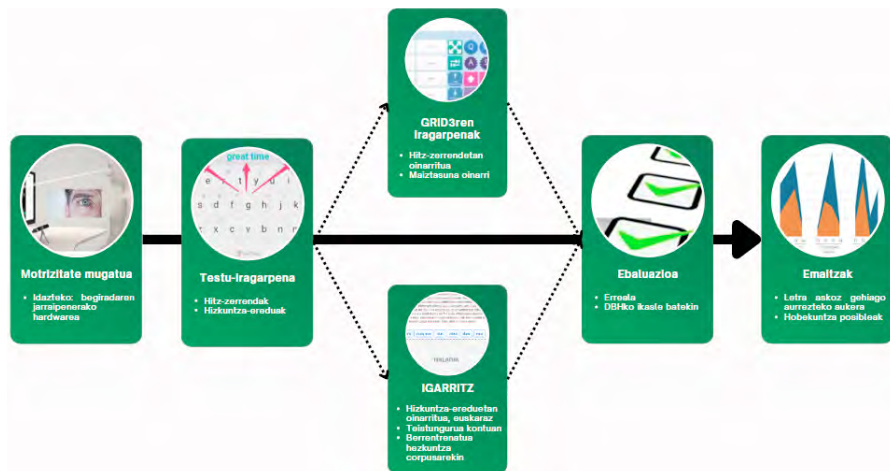
Duen funtzioa kontuan izanik, IGARRITZ honela sailkatu dugu lan honen [7] irizpideen arabera: komunikazioa erraztea eta areagotzea; zehazki, euskarazko testu idatzia iragartzea irisarekin. Bestalde, prezioa edo erabi-

---

<sup>2</sup> Batzuetan, darabilen gailua kalibratu behar izateak ere eragiten dio. Adibide honetan, behar izan dituen 11 minutu pasatxo horietatik, 1 minutu eta 5 segundo behar izan ditu idazketa-prozesuaren erdian gailua kalibratzeko.

lera doakoa da, eta online atzitzen da. Materialari dagokionez, baliabide interaktiboa da: erabiltzaileari hurrengo hitza iragartzen dio IGARRITZek, eta, letrak sartu ahala, idatzitako testuingurura eta sartutako letrara egokitzen du iragarpena. Softwarea hizkuntza-teknologian oinarrituta dago, hots, euskarazko hizkuntza-ereduak eta hezkuntzari lotutako corpusak erabili ditugu hurrengo hitza zein den iragartzeko. Erabilitako baliabideen izaerari dagokionez, desgaitasun motorra dutenentzat euskarazko testuak idazteko testu-iragarpena egiten duen tresna da IGARRITZ.

Esan bezala, testua iragartzen duten tresnek giza esfortzua eta mezua- ren ekoizpen-denbora murriztu behar dituzte; beraz, irisaren bidez teklatura sakatzeko maiztasuna gutxitu behar dugu. Emaitzen arabera, IGARRITZek euskarazko testu-iragarpena hobetzen du, eta irisarekin idazten duen pertsonak adierazi du berak askoz errazago eta gehiago idazten duela. Sortu- tako IGARRITZ testu-iragarlea ebaluatu, eta egundaino ikasleak erabili- tako testu-iragarlearekin konparatu dugu, aipatutako ikasleak ekoiztiko testuekin. Horrez gain, ebaluazio automatikoa egin dugu, eta hor ere emai- tza onak lortu ditugu; esanguratsuak, gainera (ikus 1. irudian prozesu ho- rien laburpena). Bukatzeko, IGARRITZ testu-iragarlea CLARIAH-EUS webgunean [12] probatu daiteke, hemen: <http://igarritz.clariah.eu/>.



## 1. irudia. IGARRITZ aplikazioaren sorkuntza.

Hurrengo ataletan, lan honetan ezarri dugun helburua lortzeko jarraitu dugun metodologia azalduko dugu. Horretarako, lehenbizi, begiradarekin idazteko sistemaz arituko gara, eta sistema horietan euskarak dituen mugak aztertuko ditugu (2. atala). Ondoren, testuak iragartzeko hizkuntza-ereduak izango ditugu mintzagai, eta, jarraian, IGARRITZ sortzeko erabili

dugun hizkuntza-eredua zein den zehaztuko dugu (3. atala), baita hizkuntza-eredua elikatzeke erabili ditugun corpusen berri emango ere (4. atala). Gero, hurrengo atalean, IGARRITZ web-zerbitzuaz arituko gara (5. atala); hasteko, web-zerbitzua zer den eta haren erabilera azalduko ditugu, eta, horren ostean, IGARRITZ sisteman egin ditugun egokitzapenak eta horiek egin izanaren arrazoiak emango ditugu. Azkenik, ebaluazioaz arituko gara (6. atala). Ebaluazio bikoitza egin dugu. Alde batetik, IGARRITZ automatikoki nola ebaluatu dugun ikusiko dugu; eta, beste alde batetik, kasuan oinarrituta egin dugun eskuzko ebaluazioa aztertuko dugu. Azken horretan, Grid3 tresnaren iragarpenak (ikasleak orain arte erabiltzen izan dituenak) eta IGARRITZenak alderatuko ditugu.

## 2. BEGIRADAREKIN IDAZTEKO SISTEMAK ETA EUSKARA

Orain arte, testu-iragarpenerako, Grid3 programak barneratua duen sistema erabili izan du ikasle horrek.<sup>3</sup> Grid3 KSHOrako programa bat da, eta idazketa nahiz ordenagailua erabiltzen laguntzen duen *eye-tracker*-arekin (Irisbond edo Tobii) bateragarria da. Programa horrek memorian gordetzen ditu erabilitako hitzak, azkarrago idazteko, eta norbere mezu laburrak gorde daitezke, behar denean erabiltzeko. Horrez gain, idatzitakoa entzun daiteke, TTS sistemak integratuta dauzkalako, eta, halaber, bestelako programa zehatz batzuetara sartzeko lasterbideak erabilerrazak dira.

Euskara, ordea, ez dago Grid3 programaren hizkuntza-aukeren artean. Edonola ere, aukera bat da 10.000 hitzeko kanpo-hiztegi batekin elikatzea, iragarpenak egin ditzan. Hasiera batean, Euskal Hiztegiaren Maiztasun Egitura (EHME) [8] corpusean oinarrituz, euskarazko 10.000 hitz erabilienekin elikatuta erabili izan du ikasleak Grid3. Erabiltzailea teklatuan idazten joan ahala, sistemak zerrenda horretan oinarritutako 6 iragarpen erakusten ditu, eta ordena alfabetikoaren arabera proposatzen. Hitzaren igartze-tasari begira, ikusi dugu iragarpenen ordena alfabetikoa izateak ez duela laguntzen; adibidez, erabiltzaileak «a» letra idatzi ondoren, proposatzen zaion lehen hitza «abendua» da beti, horrek zentzurik eduki ez arren. Arazo hori gainditzeko, hainbat ahalegin egin ditugu Grid3n; esaterako, hitz-zerrenda bat sartzea, hitzaren maiztasuna kontuan hartuz. Sortutako zerrenda berrian, hitz-errepikapenak sartu ditugu: EHMEko 100 hitz erabilienak 13 aldiz idatzi ditugu zerrendan, ondorengo 100 hitz erabilienak

---

<sup>3</sup> Esana dugun legez, ikasleak Grid3 sistema erabiltzen du begiradaren bidez ordenagailua maneiatu ahal izateko; eta, zerbait idaztera doanean, Grid3k berak proposatutako iragarpenak erabiltzen ditu. IGARRITZ erabiltzean, ez gara ariko erabiltzen Grid3ren iragarpenak, baina bai Grid3 sistema, funtsezkoa baita ikasleak ordenagailua maneiatzeko.

12 aldiz... 10.000 hitzeko testu-mugara iritsi arte. Modu horretan, Grid3k lehenago proposatzen ditu hitz errepikatuenak, ordena alfabetikoaren arazoari aurre eginez. Errepikapen horiek sartu behar direnez, ordea, Grid3k 2.800 hitz inguru soilik dauzka erregistratuta bere sisteman. Gainera, aurretik idatzitakoaren testuingurua kontuan hartu gabe egiten du iragarpena, corpusaren maiztasunari bakarrik begiratzen baitio. Beraz, Grid3 tresnaren iragarpenak ikaslearen beharrianetara apur bat ekartzea lortu arren, mugak oso nabarmenak dira, eta ikasleak dituen beharrianetatik oso urruti daude.

### **3. TESTUAK IRAGARTZEKO HIZKUNTZA-EREDUAK**

Hiztegiekin testuak iragartzeko metodologia baino sistema aurreratuagoa proposa daiteke hizkuntza-ereduak erabiliz, halakoek testuingurua kontuan hartzen dutelako, hau da, saioan bertan aurretik idatzitakoa kontuan hartzen dute iragarpeneko hautagaiak egokitzeko eta proposatzeko. Hizkuntza-ereduak hizkuntzaren ezaugarriak gordetzen dituzten eredu konputazionalak dira, testuak irakurtzeko, analizatzeko edota sortzeko. Hizkuntza-ereduekin aipatutako hitzen arteko erlazioak ikas-teko eta eredu estatistikoak lortzeko, testu multzo handiak behar ditugu. Eredu jakin bat sortzeko prozesu horri entrenamendua esaten diogu, eta, zenbat eta datu gehiagorekin entrenatu, orduan eta hizkuntzaren irudikapen estatistiko aberatsagoa lortuko dugu. Hortaz, esan genezake ereduak honako hauek atzeman ditzaketela: i) corpuseko hizkuntzaren erregular-tasunak, eta ii) hitzen arteko harreman estatistikoak (harreman estatistiko horiek esanahi- eta estilo-ñabardurenak ere izan daitezke), distantziak distantzia.

Hizkuntza-ereduaren zereginei dagokienez, zeregin nagusia da testuinguru zehatza duen testu baten balizko hurrengo hitzei probabilitateak esleitzea. Ataza horri esker, hizkuntza-ereduak, probabilitate handiko hitzak erabiliz, koherentetzat jo daitezkeen testu bat iragar dezake. Haietako bat edo beste bat aukeratzeak hurrengo hitzen probabilitateak ere aldatuko dituzenez, hitz bakoitzaren ondorengo probabilitate berri horiek kalkulatzeko ahalmenari esker, esanguratsuak izan daitezkeen hitzak iragar ditzake, testuingurua kontuan izanik eta testuingurura etengabe egokituz.

Horren ondorioz, aurrez definitutako hiztegi-tan edo hitz-zerrendetan oinarritutako sistemak ez bezala, hizkuntza-ereduek aukera ematen dute iragarpen adierazgarri eta koherenteagoak sortzeko. Horrez gain, hizkuntza-ereduak corpus espezializatuarekin entrenatzen edo berentrenatzen badi-tugu, erregistro horretara egokitzen da testu-iragarpena; besteak beste, hiztegia, hitzen ordena eta testuaren estiloa imitatuz.

Azken urteetako aurrerapenei esker, ikasketa automatikoko metodo zaharren zenbait muga gainditu dira *Transformer* arkitekturarekin sor-

tzen diren hizkuntza-ereduetan [4]. Izan ere, *Transformer* arkitekturaren arreta-mekanismoari esker [10], iragarpenak egiteko, hitz-sekuentzia jakin bateko zenbait hitzek duten garrantzia haztatzen du ereduak. Zehazki, RoBERTa arkitektura noranzko bikoitzeko arreta-mekanismoaz baliatzen denez, ohiko hizkuntza-ereduak ez bezala, aurreko eta osteko testuinguru kontuan izan dezake testuaren erdiko hitzak iragartzeko [9]; aurreko esaldi bat zuzentzerakoan edo aldatzerakoan, esaterako, testu berrira hobeto egokitzeko egin ohi dira horrelakoak.

#### 4. ATAZARI EGOKITZEKO ERABILITAKO CORPUSAK

IGARRITZen iragarpenak sortzeko erabilitako oinarrizko hizkuntza-eredua, HiTZ/roberta-eus-crawl-base-cased, askotariko testu motekin entrenatu da, eta, horren ondorioz, sarritan, DBHko ikasle batek erabiltzen duen hizkerarekin bat ez datozen hitz eta esaldiak iragartzen ditu. Horregatik, oinarrizko hizkuntza-eredu horrek sortutako testuak atazara doitzeko, hezkuntzari buruzko testuak bildu ditugu: Gizapedia, ikasleen idazlanak eta hezkuntzari buruzko testuak (Wikipediakoak, Elkarrekoak eta *Berria*-koak) (ikus 1. taula).<sup>4</sup>

**1. taula.** Hizkuntza-ereduko corpusaren eta berrentrenamendurako corpusen estatistikak

Rola	Corpusa	Dokumentuak	Tokenak	URLa
Hizkuntza-eredua	Euscrawl	12.528k	423M	<a href="https://www.ixea.us/euscrawl/">https://www.ixea.us/euscrawl/</a>
Doiketarako berrentrenamendua	Gizapedia	5.688	714k	<a href="https://gizapedia.org/">https://gizapedia.org/</a>
	Wikipedia	1.962	3,3M	<a href="https://eu.wikipedia.org/wiki/Kategoria:Hezkuntza_Programako_artikuluak">https://eu.wikipedia.org/wiki/Kategoria:Hezkuntza_Programako_artikuluak</a>
	Berria	354	161k	<a href="https://www.berria.eus/ikasgela/gaiak/hezkuntza">https://www.berria.eus/ikasgela/gaiak/hezkuntza</a>
	Elkar	43	1,1M	—
	DBHko ikasleen idazlanak	1.469	194k	—

<sup>4</sup> Online ez dauden eta erabili ditugun corpusak ataza zehatzak egiteko baimena izan dugu soilik.

Gure hasierako hizkuntza-eredua bildutako corpusa erabilia ikasle baten hizkerara doitzeko, eredia entrenatzen jarraitu dugu, haren aurretiko entrenamenduan erabili dugun ataza berberari jarraituz, hots, etengabeko aurretiko entrenamenduaren teknika erabiliz. Ataza horretan, testuaren hitzak maskaratu edo ezkutatu egiten dira, eta maskaratutako leku horietan dauden hitzak iragarri behar ditu ereduak. Entrenamenduan, Huggingface liburutegiak lehenetsitako balioez gain, honako hiperparametro hauek erabili ditugu:  $2e-5$  ikasketa-tasa ( $\gamma$ ),  $0,01$  pisuen gainbehera ( $\lambda$ ) eta AdamW optimizazio-algoritmoa ( $\beta_1 = 0,9$ ,  $\beta_2 = 0,9$  eta  $\epsilon = 1e-6$ ).

## 5. IGARRITZ WEB-ZERBITZUA

Hizkuntza-ereduekin eta corpusarekin azaldutako metodologia erabilia, IGARRITZ web-zerbitzua sortu dugu, eta nahi duenak erabiltzeko moduan jarri dugu (ikus 1. irudia); Firefox esploratzailea erabiliz probatu daiteke, hemen: <http://igarritz.clariah.eus>

IGARRITZek bi atal nagusi ditu: hitzen iragarpenaren eremua eta, haren gainean, hitza aukeratzeko hautagaien testu-kaxa. Iragarpenari dagokionez, sistemak hainbat hitzen aurreikuspena egiten du. Batetik, ordura arteko testua edota idatzitako letrak kontuan izanik, balizko hurrengo 6 hitz erakusten ditu. Proposaturiko hitzen artean, erabiltzaileak behar duen hitzik ez badago, letra bat idatz dezake teklatuarekin edo teklatu birtualarekin. Halaber, IGARRITZek badu teklatu birtuala jartzeko eremu bat, kendu edo jarri egin daitekeena, erabiltzaileak hala nahi badu. Sagua edo begirada aukeratu nahi den hitzaren edo letraren gainean  $0,6$  segundo mantenduz edo klik eginez aukeratzen da, eta, denbora pasatu ahala, kaxa koloreztatu egiten da, erabiltzaileak zer aukeratzen ari den jakin dezan.

Bestalde, iragarpenak irakurtzean begirada mantenduz gero nahi ez den hitzen bat aukeratzeko arriskua dagoenez, eragiketak desegiteko botoiak jarri ditugu IGARRITZ-en. Horrela, erabiltzaileak idatzitako letrak ez dira galduko, eta aurreko egoerara bueltatzeko aukera izango du. Botoi horiek aukeratzeko, begirada denbora gehiagoz mantendu behar da, zailagoa izan dadin botoiok nahi gabe aukeratzea.

Bukatzeko, idatzitakoa kopiatzeko eta beste programa batean itsasteko edota idatzitakoa ezabatzeko aukerak eskaintzen ditu IGARRITZek. Testua nahi gabe ezabatuz gero, atzera egiteko eragiketeta desegiteko aukera dago (ikus 1. irudia).





2. irudia. IGARRITZ aplikazioa eta interakzioen azalpena.

### 5.1. Web-zerbitzuko idazketa-sisteman egindako egokitzapenak

Begiradaren jarraipenarekin ordenagailua maneiatzeko, begirada detektatzeko hardware bat erabiltzen da (*eye-tracker*-a: Irisbond edo Tobi, besteak beste), eta, horrez gain, prozesu hori errazten duten programak (Grid3, Easyclick...) ere kontuan izan behar ditugu. IGARRITZ tresna begiradaren jarraipenarekin idazteko diseinatu badugu ere, Diseinu Unibertsalaren kontzeptua gogoan izan dugu, ahalik eta testuinguru gehienetara egokitzeko. Hori argi ikusten da egin dizkiogun aldaketetan eta tresnak eskaintzen dizkigun aukeretan. Aldaketa eta aukera horiek bi fasetan egin ditugu batez ere: i) erabiltzailearekin lanean ari den irakaslearen gomendioei jarraituz egin diren aldaketak, hots, aditu baten aholkuz, eta ii) erabiltzailearekin zenbait proba egin ondoren egin diren aldaketak.

### 5.2. Erabiltzaileari begira egindako aldaketak (proba aurrekoak)

Honako aldaketa hauek egin ditugu:

- Aukeratzea. Baliabide honi esker, klik egin gabe eta begirada edo sagua botoiaren gainean mantenduz aukeratu ahal izango ditugu iragarritako hitzak. Aukeratzeko, ikasleak berak irisarekin klik egiteko konfigurazio-denbora errespetatu dugu, hots, 0,6 segundoan mantendu behar dugu begirada edo sagua iragarpenean, aukeratu ahal izateko.
- Aukeraketa nabarmentzea. Begirada non dagoen jakiteko, IGARRITZeko botoiak kolorez betetzen hasten dira, ezkerretik es-

- kuinera, begirada bertan dagoen bitartean. Botoiak guztiz betetzen direnean egiten dute klik.
- «Kopiatu» botoia. Testu-kaxan idatzitako testu guztia kopiatzen du, eta beste nonbaitera eramatea errazten.
  - «Ezabatu» botoia. Testu-kaxan dagoen testu guztia ezabatzen du. Botoi hori nahi gabe ez aukeratzeko, denbora gehiago egon behar da begirada edo sagua haren gainean ezarrita.
  - «Berregin» eta «desegin» botoiak. Historian aurrera edo atzera joateko botoiak dira. «Ezabatu» botoiarekin, egindako eragiketak zein iragarpen-aukeraketak desegin daitezke. Beraz, iragarpen bat oker aukeratuz gero, «desegin» botoia sakatuta, erabiltzaileak idatzi dituen letrak ez dira galduko, eta berriro agertuko zaizkio iragarpen guztiak.
  - Botoien tamaina. Iragarritako hitza aukeratzeko, markoa handitu dugu, irisarekin errazago aukeratu ahal izateko.
  - Kanpo-teklatura jartzeko eta menua ezkutatze botoia. IGARRITZ beste teklatu edo sistema batzuetan integratzeko eta pantailan espaziorik ez galtzeko, erabiltzaileak teklatuaren hutsunea ken lezake.
  - Puntuazioa zuzendu. Iragarritako hitza aukeratzean, espazioa ere jartzen du. Puntuazioa idaztean, aurretik jarritako espazioa ezabatu egiten da, eta aurreko hitzari lotzen zaio puntuazio-ikurra. Horrek guztiak erabiltzaileak espazioa gehitzen edo ezabatzen ibili beharra saihesten du.

### **5.3. Erabiltzailearekin egon ondoren egindako aldaketak**

Honako aldaketa hauek egin dizkiogu IGARRITZi:

- Iragarritako hitzen kokagunea. Probabilitate handiena duten hitzak erdian agertuko dira. Modu horretan, begirada gutxiago mugitzeaz gain, iragarritako hitzak gaineratik irakurtzeko begirada nora eraman behar duen jakingo du erabiltzaileak, eta hitz esanguratsuen aukeratzeko ahalegin txikiagoa egingo du.
- Botoien tamaina. Botoien horizontaltasuna mugatu, eta erdian jarri ditugu, bertikalago, begirada bazterretara asko mugitu gabe hitzak irakurri eta aukeratzeko.
- Historiala gorde. Begiradaren kontrola ez denez erabatekoa, gertatu izan da «ezabatu» botoia nahi gabe sakatu eta idatzitako guztia galdu izana. Hori saihesteko, ezabatzeko botoia sakatzeko denbora gehiago ezarri diogu, eta, horrez gain, baita historialean atzera eta aurrera egiteko aukera ere. Horrela, zerbait nahi gabe eginez gero, egindakoa erraz berreskura daiteke.

## 6. EBALUAZIOA ETA EMAITZAK

### 6.1. Ebaluazio automatikoa

Hitzak iragartzeko, ezkutatuta edo maskaratuta dauden hitzak igartzeko entrenatu dugu IGARRITZek darabilen hizkuntza-eredua, eta, beraz, aipatu behar da tresnaren helburua ez dela testu osoak sortzea, baizik eta testuaren une jakin batean hitz probableen zerrenda proposatzea. Ondorioz, testu-sorkuntzan erabiltzen diren ebaluazio-metrika eta -prozesu ohikoenak ez dira aproposak gure tresna ebaluatzeko. Metrika horien zenbait adibide dira BLEU, Rouge eta Perplexity neurriak, besteak beste. Hori horrela, IGARRITZek darabilen hizkuntza-ereduaren iragarpen-gaitasuna neurtzeko, gure atazari lotutako ebaluazio-prozesu egokitua diseinatu dugu.

Horretarako, hizkuntza-ereduaren ezaugarriez gain, atazan, tresnaren erabileran eta zenbait kasuistikatan oinarritu gara. Ebaluaziorako testu multzoan dauden hitz guztien asmatze-tasa neurtzea nekeza denez, testu bakoitzean gutxienez hiru letraz osatutako hitzen % 15 maskaratu dugu. Gutxieneko letra kopurua ezarri dugu hitz luzeak iragartzen zailagoak direlako. Bestalde, testuak moztu egin ditugu, adibide berriak sortzeko eta testua erabat burutu gabe dagoela irudikatzen. Azken kasu horietan, azken hitza beti maskaratu dugu, baita bestelako hitzen % 10 ere. Bestelako hitzak ere maskaratzean, erregularitasun sendoagoak eta orokortuagoak ikasi eta ebaluatu behar ditu hizkuntza-ereduak, eta ez da soilik mugatzen asmatu behar duen uneko hitzera.

Ebaluaziorako instantziak prestatuta, maskaratutako hitzen iragarpena eta ebaluazioa egin dugu. Erabiltzailearen erabilpena erreplikatzeko, hitz guztien artean probabilitate handiena duten sei hitz hartu ditugu. Hitz multzo horretako zerrendan jatorrizko hitza ez badago, hitzaren hurrengo letraren arabera bahetu ditugu hitzak, beste sei hautagai lortzeko. Jatorrizko hitza hautagaien zerrendan agertu arte edota hitza guztiz idatzi arte errepikatu dugu prozesu hori.

Horrekin, alde batetik, ereduaren zehaztasuna neurtu dugu. Bestalde, ereduak zenbat hitz aurreztuko lukeen ere neurtu dugu. Azken neurri hori oso garrantzitsua eta esanguratsua da ataza honetan, helburu nagusia ikaslearen testuen idatzizko produkzioa handitzea delako.

## 6.2. Kasuan oinarritutako eskuzko ebaluazioa

Ikasleak, Grid3 tresnaren iragarpenak erabiliz (euskarazko hitz zerrendak sartuta), 4 testu idatzi ditu, eta bideoan grabatu dugu prozesu osoa.<sup>5</sup> Ondoren, bideoen laguntzaz, testu horietan erabilitako hitzak bildu ditugu, eta ikasleak egindako modu berean idatzi ditugu, IGARRITZ erabilia, hots, hitzen testuingurua erabat mantenduz. Kasu batzuetan, ikasleak hitz bat osorik idatzi du, baina, gero, ezabatzea erabaki; horrelakoetan, hitza kontatu egin dugu. Ostera, letra solteak edo hitza osatzera iristen ez diren letra-errenkadak ezabatu baditu, horiek ez ditugu bildu, ez baitugu horrelakoak aurreikusterik. Bi arau horiek kontuan izanda, ebaluazioa egiteko, bi motatako hitzak aztertu ditugu:

- A. Testu iragarleak osorik aurreikusten dituen hitzak: ikasleak idatzitako hitza iragarleak zehatz-mehatz aurreikusi duenean, zenbatgarraren letran proposatzen duen jaso dugu, baita zenbat letra aurrezteko parada ematen duen ere. Adibidez, «irakasleak» hitza idatzi nahi izan dugunean, IGARRITZek «ir» idatzita ematen digu nahi izan dugun emaitza; beraz, kasu honetan, 2 letra idatzi behar izan ditugula etiketatu dugu, eta 8 letra aurreztu ahal izan ditugula.
- B. Hitza osorik idazteko moldaketaren bat behar dutenak: iragarleak ez ditu hitz guztiak asmatzen, baina, asmatzen ez dituen horien artean, batzuetan iragarpenek antz handia dute idatzi nahi den hitzarekin, eta, iragarpenari moldaketaren bat eginda, nahi dugun hitza idaztera hel gitezke. Horrelakoetan, moldaketak eginda zenbat letra aurreztu ahal izan ditugun jaso dugu. Posible da iragarleak egindako iragarpen batek baino gehiagok moldaketak egiteko aukera ematea; kasu horretan, letra gehien aurrezteko aukera ematen duena erabili dugu. Adibidez, «oinarritzeko» hitza ez du iragartzen IGARRITZek, baina, «o» idatzi dugunean, «oinarrituta», «oinarriturik» eta «oinarrituz» agertu dira 6 aurreikuspenen artean. Kasu honetan, «oinarrituz» hitza hartu dugu, «uz» ezabatu ondoren «zeko» gehitzeko. Horrela, 12 letra dituen hitzetik, «o» eta zeko» idatzi ditu (5 letra), eta, «uz», ezabatzeko birritan egin behar denez klik, orotara 5 letra aurreztu dituela jaso dugu. Beste bi aurreikuspenak ez ditugu hautatu, letra bat gehiago ezabatu beharko genukeelako.

Eskuzko ebaluazioari dagokion hitzen azterketa bi eratara egin dugu: orokorrean eta testuka. Horrez gain, ondoren, hitzen letra kopurua hartu

---

<sup>5</sup> Testuetako bat, 2. testua, ikasleak aurreko egunean hasitako gutun bat da; bideoan gutunaren bigarren zatia soilik grabatu denez, zati hori hartuko dugu testutzat. Bestalde, 3. eta 4. testuak testu beraren parte dira, baina, ikasleak bi egunetan idatziak direnez, bakoitza testu bat balitz bezala hartu dugu.

dugu aldagaizat, eta lau testuetan IGARRITZek egindako aurreikuspenak zer-nolakoak izan diren aztertu dugu. Horretarako, hitzak lau multzotan sailkatu ditugu: 0-3 letra, 4-6 letra, 7-9 letra eta 10 letra edo gehiago dituztenak.

Modu horretako ebaluazioa egin dugu, idazmenaren ekoizpenean lagundu beharreko ikasleari idazteko lan gehiago ez ematearren; gainera, ikaslearen erabilerak ez du iragarpen-ahalmena neurtzen. Beraz, benetan neurtu duguna tresnen iragarpen-ahalmena da, eta ez ikaslearen erabilera. Hori bai, ikasleari zuzenean galdetu diogu, eta erantzun du askoz azkarrago idazten duela IGARRITZekin, nahiz eta oraindik funtzio batzuk egokitu behar ditugun erosoago erabil dezan.

### 6.3. **Emaitzak**

Lan honetan lorturiko emaitza nabarmenena IGARRITZ sistema sortu izana da (<http://igarriz.clariah.eus>).

IGARRITZ-en erabilpena antzeratzen duen ebaluazio automatikoarekin, bi eredu testatu ditugu; alde batetik, oinarriko eredu, HiTZ/roberta-eus-euscrawl-base-cased, eta, bestetik, hezkuntzari buruzko testuekin doitutako eredu. Oinarriko hizkuntza-eredua erabilia, 70,9 mila letra aurrezten dira (% 19,57ko zehaztasunarekin (*accuracy*)); hezkuntzako corpusarekin doitutako hizkuntza-ereduarekin, ordea, 74,5 mila letra aurrezten dira (% 19,89ko zehaztasuna), 3,6 mila letra gehiago, hots, % 5 gehiago. Doitutako ereduak lehenago iragartzen du hitz aproposa, eta, horren ondorioz, erabiltzaileak letra gehiago aurrezten ditu, bi ereduek antzeko zehaztasuna izan arren.

IGARRITZ-en eskuzko ebaluazioari dagokionez, 1. taularen arabera, emaitzak honako hauek dira: aztertu ditugun 4 testuetan, IGARRITZekin emaitza askoz hobea lortu ditugu ikasleak orain arte erabilitako Grid3 tresnaren iragarpenekin baino. Testuen datu orokorreari erreparatzen badiogu, asmatutako hitzen bidez aurreztu genitzakeen letrak % 50,07 dira IGARRITZekin; Grid3ren iragarpenekin, aldiz, % 18,09. Azken horrekin alderatuta (% 13,9), moldaketekin aurreztu genetzakeen letra kopurua jaitsi egiten da IGARRITZekin (% 7,77), baina joera hori positiboa da, asmatutako hitzekin hitzak aurreztu ahala normala baita moldaketa gutxiago egiteko beharra izatea. Orokorrean ez ezik, testu guztietan hobetu ditu emaitzak IGARRITZek. Hala ere, kontuan hartu beharrekoa da testuen artean dagoen aldea. Esaterako, 2. testuan, asmatutako hitzei dagokienez, 238 letratatik 149 aurreztu ditugu, hots, % 62,61, eta, 4. testuan, ordea, ehuneko dezente jaisten da; % 26,67ra, hain zuzen. 2. testua aurrez idatzitako gutun baten jarraipena denez eta IGARRITZi aurreko zati hori erantsi diogunez hasieran, ondoriozta genetzake aurreko testuingurua lagungarria izan daitekeela oso iragarpenak hobetzeko. Halaber, esana dugunez, 4. tes-

tua ere beste testu zati baten jarraipena da, baina testu horrekin ez ditugu emaitza hain onak lortu. Beraz, testuingurua lagungarria dela suma badezakegu ere, badirudi edukiak eta gaiak eragin zuzena dutela.

**2. taula.** Eskuzko ebaluazioaren emaitzak; Grid3ren iragarpenekin eta IGARRITZekin aurreztu daitezkeen letrak

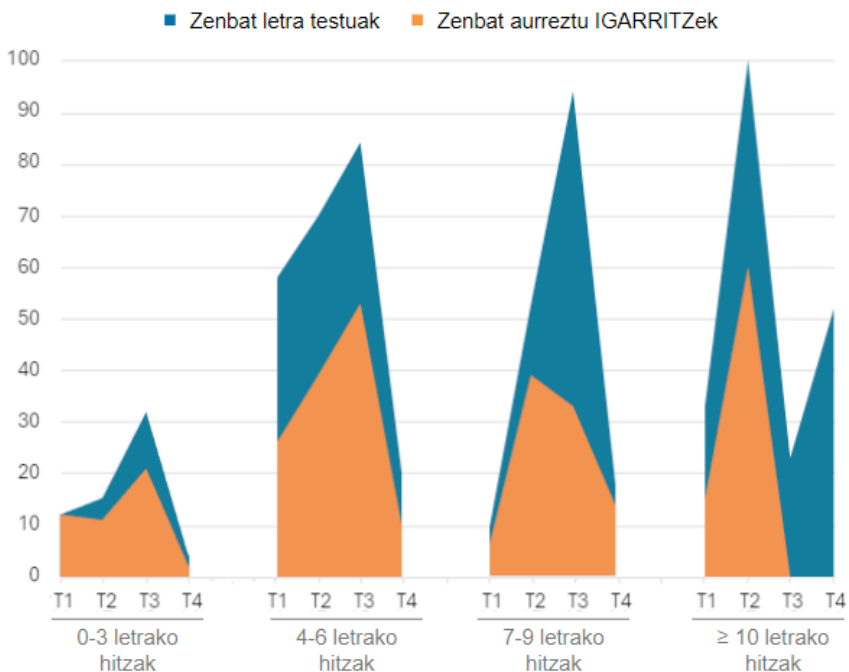
	Testuko letrak	Aurreztutako letrak		Moldaketetan aurreztutakoak	
		Grid3	Igarritz	Grid3 (bir.)	Igarritz (bir.)
1. testua	108	20	55	16	16
2. testua	238	45	149	54	9
3. testua	233	56	107	22	20
4. testua	90	0	24	4	7
<b>Orotara</b>	<b>669</b>	<b>121</b>	<b>335</b>	<b>93</b>	<b>52</b>
<b>Aurreztutakoa</b>		<b>% 18,09</b>	<b>% 50,07</b>	<b>% 13,90</b>	<b>% 7,77</b>

Zenbat letra aurreztu daitezkeen aztertu ondoren, IGARRITZek zer hitz mota iragarri eta aurrezteko aukera ematen duen aztertu dugu; ikus 2. taula. Letra kopurua kontuan izanda, ikus dezakegu 0-3 letra dituzten hitzak direla maiztasun txikiena dutenak (% 9,42). Alabaina, hitz mota horietan du IGARRITZek iragarpen-tasa eraginkorrena, letren % 73,02 aurreztea ahalbidetzen baitigu. Gainontzeko 3 multzoetan, letra kopurua orekatuagoa izan da: testuko 4-6 letrako hitzetan, letren % 34,38 aurreztu dira, 6-9 letra dituztenetan % 25,26 eta 10 letratik gorakoetan % 30,94. Aurrezteko aukerei erreparatuz gero, 4-6 eta 6-9 letrako hitzetan aurreztutako datuak oso antzekoak dira, hurrenez hurren letren % 54,78 eta % 52,66 aurrezteko parada izan baitugu. Haatik, ehunekoa dezente jaisten da 10 letratik gorakoetan kasuan; % 35,75era, hain zuzen. Gainera, testuka gertatu dena begiratzen badugu, argi eta garbi ikusten dugu testuen artean alde esanguratsua egon dela. Emaitzetan ikusten denez, 1. testuan, 10 letratik gorako hitzek 32 letra biltzen dituzte, eta IGARRITZek 14 aurrezteko aukera eman digu (% 43,75), eta, 2. testuan, emaitzak are hobekak dira, 100 letratatik 60 (% 60) aurrezteko aukera ematen baitigu. Baina emaitzak erabat bestelakoak dira beste bi testuetan; batean zein bestean, hitzek 23 eta 52 letra bildu arren, ez dugu izan letraren bat aurrezterik testu-moldaketarik egin ezean. Testuen arteko alde hori oso ondo ikusten da 3. irudian. Hala, ikus dezakegu testuaren letra kopuruak gora egiten duenean IGARRITZek letrak aurrezten jarraitzen duela, hots, asmatze-tasari eutsi egiten dio, forma bereko mendi urdin eta laranja irudikatuz. Alabaina, batzuetan ez da bete hori, eta ikus dezakegu multzo bereko hitzetan grafiko urdin eta laranja

forma bera izatetik erabat bestelakoa izatera igaro dela. Multzo bereko hitzen artean horrelako aldea egoteak ideia bera azpimarratzera garamatza berriro ere: IGARRITZek aurrez duen testuinguruak eta idazten ari garen eduki edo gaiak eragin zuzena dute iragarpenak eraginkorrak izateko orduan.

**3. taula.** Hitz moten arabera IGARRITZekin aurreztu daitekeen letra kopurua, moldaketak kontuan izan gabe.

	1-3 letra		4-6 letra		6-9 letra		10 ≤ letra	
	zenbat	aurreztu	zenbat	aurreztu	zenbat	aurreztu	zenbat	aurreztu
1. testua	12	12	57	25	7	4	32	14
2. testua	15	11	70	39	53	39	100	60
3. testua	32	21	84	53	94	33	23	0
4. testua	4	2	19	9	15	13	52	0
<b>Orotara</b>	<b>63</b>	<b>46</b>	<b>230</b>	<b>126</b>	<b>169</b>	<b>89</b>	<b>207</b>	<b>74</b>
<b>%</b>	<b>9,42</b>	<b>73,02</b>	<b>34,38</b>	<b>54,78</b>	<b>25,26</b>	<b>2,66</b>	<b>0,94</b>	<b>35,75</b>



**3. irudia.** Aurreztutako letrak 4 testuetan.

## 7. ONDORIOAK ETA ETORKIZUNeko LANA

Lan honetan, euskarazko testu-iragarpena egiteko eta begiradarekin idazteko programekin bateragarria den lehen web-ingurune egokitua sortu dugu adimen artifizialeko teknikak erabiliz: IGARRITZ (ikus: <http://igarritz.clariah.eus/>). Testu-iragarpena egiteko, *Transformer* arkitektura erabiltzen du [4], eta RoBERTa kodifikatzaile-arkitekturaren [5] oinarritutako HiTZ/roberta-eus-euscrawl-base-cased hizkuntza-eredua [6]. IGARRITZ behar duen DBHko ikasle batek tresna probatu du eta hobetzen lagundu digu. Ikaslearen beharrezanetara egokitzeko, tresnak darabilen hizkuntza-eredua berariaz sorturiko hezkuntzari buruzko corpusekin berrentrenatu dugu, etengabeko aurretiko entrenamendua (*continual-pretraining*) deritzon teknika erabiliz. Hizkuntza-teknologian oinarrituta dagoen tresna da IGARRITZ. Desgaitasun motorra dutenentzat diseinatu da, euskaraz idazteko; eta begiradaren bidezko komunikazioa errazteko eta areagotzeko sistema bat da, doakoa eta online atzitzen dena. Horrez gain, baliabide interaktiboa da; erabiltzaileari hurrengo hitza idazteko zenbait hautagai iragartzen dizkio IGARRITZek, eta, letrak sartu ahala, idatzitako testuingurura eta sartutako letretara egokitzen ditu iragarpenak.

Ebaluazioari dagokionez, ebaluazio automatikoari esker, IGARRITZek iragarpenik onenak zer hizkuntza-eredurekin egiten dituen ebatzi ahal izan dugu. Emaitzen arabera, hezkuntzako corpusarekin doitutako hizkuntza-ereduak lehenago iragartzen du hitz aproposa, eta, horren ondorioz, erabiltzaileak letra gehiago aurrezten ditu, oinarritzko hizkuntza-ereduak antzeko zehaztasuna izan arren. Bestalde, IGARRITZ-en iragarpen-ahalmena eskuz ebaluatu dugunean, ikusi dugu IGARRITZekin emaitza askoz hobekuntza lortu direla ikasleak orain arte erabilitako Grid3 tresnaren iragarpenekin baino, eta hobekuntza hori, gainera, aztertutako testu guztietan hauteman dugu. Hala ere, eskuzko ebaluazioan, ikusi dugu testu batetik bestera alde dagoela eta alde horren zio nagusiak bi izan daitezkeela: i) aurrez testu zati bat (testuingurua) idatzita izateak aurreikuspenak hobetu egiten dituela, eta ii) testuaren gaiak eta edukiak eragin zuzena dutela aurreikuspenetan. Nabarmendu nahi dugu lan honetan ebaluatu duguna tresnen iragarpen-ahalmena dela, eta ez ikaslearen erabilera erreala. Edonola ere, ikaslearekin izan dugun harreman zuzenari esker, zuzenean behatu dugu erabilera; gainera, zuzenean galdetu diogu iritzia, eta ikasleak erantzun du askoz azkarrago idazten duela IGARRITZekin, nahiz eta oraindik funtzio batzuk egokitu behar diren, erosoago erabil dezan. Horrez gain, aipatu nahi dugu ikaslearen tutoreak ere adierazi digula IGARRITZek asko laguntzen diola eta interesgarria litzatekeela etorkizuneko ikasturteetako eginkizun eta gaietara gehiago egokitzea.



## ESKERTZA

Lan honek Eusko Jaurlaritzako Hezkuntza Sailaren laguntza jaso du. Horrez gain, eskerrak eman nahi dizkiegu Iñaki Orberi eta Lorea Aretxagari, ikerketa hau sustatzeko emandako laguntza guztiarengatik, baita IGARRITZ erabili duen ikasleari eta haren familiari ere. Horrez gain, CLARIAH-EUS azpiegituraren lanen barruan egin da lan hau, eta HiTZ zentroak ere lagundu du, baliabideak eta webgunea eskainiz. Bukatzeko, ISEI-IVEIri, ELKARI eta Gizapediari ere eskerrak eman nahi dizkiegu, haiek sortutako testuak IGARRITZ berrentrentzeko uztearren.

## BIBLIOGRAFIA

- [1] EUSKO JAURLARITZA. 2022. *Hezkuntza-sisteman «komunikazio sistema handigarriak edo/eta ordezkioak» (KSHO) inplementatzeko protokoloa*. Eusko Jaurlaritzako Hezkuntza Saila. [https://www.euskadi.eus/contenidos/informacion/inn\\_edu\\_inc\\_neapoyoeducativo/eu\\_def/adjuntos/Protocolo\\_SAACS\\_e.pdf](https://www.euskadi.eus/contenidos/informacion/inn_edu_inc_neapoyoeducativo/eu_def/adjuntos/Protocolo_SAACS_e.pdf)
- [2] LLOYD, L. & KARLAN, G. 1984. Non-speech communication symbols and systems: where have we been and where are we going? *Journal of mental deficiency research*, 28 (Pt 1), 3-20. <https://pubmed.ncbi.nlm.nih.gov/6232388/>
- [3] KLAUS, B., AIGNER, B., & VEIGL, C. 2019. AsTeRICS Grid—a flexible web-based application for Alternative Communication (AAC), environmental and computer control.
- [4] VASWANI, A., SHAZEER, N., PARMAR, N., USZKOREIT, J., JONES, L., GOMEZ, A. N., KAISER, L. & POLOSUKHIN, I. 2017. Attention is all you need. *Advances in neural information processing systems*, 30.
- [5] LIU, Y., OTT, M., GOYAL, N., DU, J., JOSHI, M., CHEN, D., LEVY, O., LEWIS, M., ZETTMLOYER, L. & STOYANOV, V. 2019. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*.
- [6] ARTETXE, M., ALDABE, I., AGERRI, R., PEREZ-DE-VIÑASPRE, O., & SOROA, A. 2022. Does Corpus Quality Really Matter for Low-Resource Languages?. *arXiv preprint arXiv:2203.08111*.
- [7] LUNA, M. R. 2013. Tecnología y discapacidad: Una mirada pedagógica. *RDU, Revista Digital Universitaria*, 14(12), 1-19, <http://www.ru.tic.unam.mx/handle/123456789/2177> [20XX-XX-X].
- [8] EUSKAL HIZTEGIAREN MAIZTASUN EGITURA (eguneratua: 2024-01-10). Euskal Herriko Unibertsitatea. <https://www.ehu.eus/ehg/ehme/datu2hitz.htm>
- [9] WANG, A. & CHO, K. 2019. BERT has a mouth, and it must speak: BERT as a Markov random field language model. *arXiv preprint arXiv:1902.04094*.
- [10] Ashish VASWANI, Noam SHAZEER, Niki PARMAR, Jakob USZKOREIT, Llion JONES, Aidan N. GOMEZ, Lukasz KAISER, & Illia POLOSUKHIN. 2017. Attention is All you Need. In *NIPS* (pp. 5998-6008).

- [11] GARAY-VITORIA, N., & ABASCAL, J. 2006. Text prediction systems: a survey. *Universal Access in the Information Society*, 4, 188-203.
- [12] ALTUNA, B., IRUSKIETA, M., ESTARRONA, A., FARWELL, A., ARRIOLA, J. M., ALKORTA J., eta ARREGI, X. (2022). CLARIAH-EUS: Building a Cross-border CLARIAH Node for the Basque Language. *Digital Humanities Conference November 23-25*. Budapest.