

## Itzultzaile automatikoak osasun-arretan: azterketa bibliografikoa

### *Automatic translators in healthcare: a bibliographic revision*

Adrian Unai Zelaia Arieta-Araunabeña<sup>1</sup>, Aitor Montes Lasarte<sup>2</sup>, Unai LarrinagaTorrontegui<sup>3</sup>

<sup>1</sup> Osakidetza, Araba ESI,

<sup>2</sup> Osakidetza, Debagaraia ESI,

<sup>3</sup> Osakidetza, Debabarrena ESI,

adrianunai.zelaiaarietaraunabena@osakidetza.eus

### Laburpena

---

**Aurrekariak:** Gure asmoa itzultzaile automatikoen erabilerari buruz existitzen den ebidentzia ebaluatzea da. Osakidetza (Euskal Osasun Zerbitzua) euskal hiztunak diren pazienteei osasun-arreta elebiduna eskaintzeko horrelako tresna bat garatzen ari da.

**Metodoak:** Osasun-sistemetan itzultzaile automatikoen erabilerak ebaluatzen dituzten ikerketak bilatu ditugu 2 datu-basetan: Pubmed eta ClinicalKey. Bederatzigarren artikulu jo ziren egokitzat azterketarako, 8 ikerketa eta aditu-iritzi bat hain zuzen ere. Emaitzen neurketak doktore, paziente eta itzultzaile profesionalen egindako inkesta gisa azaltzen dira, ikerketa batek bere emaitzak itzulpen-akats (dentsitate lexikoak doituak) eran eta beste batek itzulpen-itei esleitutako BLEU eta PER emaitzetan.

**Emaitzak:** Ezein aplikaziorik ez luke itzultzaile profesional bat ordezkatu behar, emaitza positiboak lortu arren. Aplikazio hauek jardun kliniko eraginkor baterako mugak dituzte. Ez da aplikazio hauetan soilik oinarrituriko egoera kliniko bat ezartzen, eta beren egokitasuna eta erabilgarritasuna beste inongo tresnarik ez dagoen aldi batera mugatzen dute.

**Eztabaida:** Artikulu hauen ebidentziaren kalitatea baxua da. Ez dago saio klinikorik, eta denak dira *in vitro*, 3 ikerketa salbu. Baieztapen eta ondorio denak zehaztasun onartu behar dira, eta tresna gehigarri gisa soilik erabili. Itzulpen automatikoa ez daiteke gaitasun edota segurtasun kulturalaren estandarretan barnebildu, azken aukera gisa baino.

**Ondorioak:** Itzultzaile automatikoen ez dute osasun-arretan beren kontura ezartzeko eraginkortasun, zehaztasun edo egokitasunik erakutsi. Hizkuntza-mugak direla eta, egungo *goldstandard*-ahizkuntza-parekotasuna da itzulpena baino. Ikerketaren etikak soilik baimentzen du hau hobetu edota berdintzeko asmoa duten alternatibak bilatzea.

**Gako-hitzak:** itzultzaile automatikoa, makina bidezko itzultzailea, aplikazioa, hizkuntza-muga, osasun-arreta

### Abstract

---

**Background:** We mean to assess the existing evidence of automatic translators. The Basque Health Organisation (Osakidetza) is developing such a tool.

**Methods:** We searched 2 databases for studies evaluating the use of automatic translators in healthcare systems: Pubmed and Clinical Key. 9 articles were considered suitable for the review, 8

Adrian Unai Zelaia Arieta-Araunabeña, Aitor Montes Lasarte, Unai LarrinagaTorrontegui

*studies and an expert opinion. The main outcome measures were shown as surveys on doctors, patients and professional translators, with a study that reported its outcomes as number of translation errors (adjusted by lexical density) and another one as the BLEU and PER score assigned to the translation items.*

*Results: No app should replace a professional translator, regardless their positive results. Apps are limited for an effective clinical activity. None establish a clinical situation based only in the evaluated apps, keeping their suitability and applicability for when no other tool is available.*

*Discussion: The quality of the evidence of these articles was low. No clinical trial was carried out, and only 3 studies are not in vitro. Conclusions should be cautiously accepted, and apps only used as supplementary tools. Automatic translation cannot be included among culturally competent or safe standards, but as a last option.*

*Conclusion: Automatic translators have not shown efficiency, accuracy and suitability enough to be implemented on their own in healthcare. Language parity is the gold standard regarding language barriers rather than human translators. Research ethics demand alternatives that only aim to equal or improve it.*

*Keywords: automatic translator, machine translator, application, language barrier, healthcare*

Bidalia: 352 - 21.01.06

Onartua: 2021.04.24

<http://doi.org/10.26876/osagaiz.1.2021.352>

## 1. Sarrera

---

Hizkuntza-dibertsitatea gizateriaren ezaugarria izan da beti, eta hala aitortu izan da beti. Fenomeno hori ez da hizkuntzalarien afera soilik, ezta osasun-arretatik kanpo kokatzen den zerbait ere. Izan ere, osasun-arreta jendearen osasuna zaintzean datza, indibiduo gisa edo kolektiboki, eta jendea edota gizartea ezin uler daitezke hizkuntzarik gabe.

Hizkuntza ofizialen kasuan, arautzat jo ohi da tokiko hizkuntzaren edo hizkuntza nagusiaren derrigorrezko ezagutza, maila edo gaitasun aurreratu egiaztatuaeren terminoetan (C1 edo C2 maila, Europako Hizkuntzen Erreferentzien Marko Komunikazioaren sailkapenean) (Espainiako 1978ko Konstituzioa, 3.1 artikulua: Espainiar denek dute espainiar hizkuntza jakiteko betebeharrak eta erabiltzeko eskubidea; Espainiako Gobernuaren 55/2011 Errege Dekretua, zeinak osasun-arreta espezializaturako espainiarren C1 edo C2 maila derrigorrezkotzat ezartzen duen).

Komunikazio-arazoak eragin desiragaitzen %69ren arduraduntzat jotzen ditu AEBko Joint Commission-ek, argi utziz ingeles-gaitasun baxua (IGB) duten pazienteek eragin desiragaitz gehiago pairatzen dituztela ingeles-hiztunek baino.

Hizkuntza-mugek gaixotasun kronikoen maneio eta jarraipena oztopa dezakete (diabetesa, min kronikoa eta abar), baita tratamenduen betetzea ere, ospitaleratze-tasahandiagoekin eta egonaldi luzeagoekin<sup>(1)</sup>.

Pazienteen konfiantza- eta asebetetze-mailak baxuagoak dira hizkuntza-diskordantziak tarteko direnean, osasun-arreta hizkuntza ezberdin batean jasotzea norberaren identitatearekiko mehatxu gisa edo inposizio gisa ulertzeraino<sup>(2)</sup>.

Hizkuntza-ezagutza falta da osasun-arretako mugarik garrantzitsuena osasun-profesional eta iritsi berri diren horien arabera<sup>(3)</sup>. Bertako osasun-sistemaren irisgarritasuna eta unibertsalitatea zalantzan jar ditzake horrelako afera batek, non osasun-arreta publikoaren estaldura falta egon daitekeen<sup>(4)</sup>.

Paziente eta langileen arteko komunikazioa oinarritzakoa da, ahal den heinean biek hizkuntza bera hitz egitea lehenetsiz, are gehiago pazientearengan zentratutako arreta bilatzen badugu, Ohtaniri jarraiki.



Hizkuntza-parekidetasuna eta konpetentzia kulturala urrezko estandar<sup>(5,6)</sup> kontsideratu izan dira, itzulpen profesionala baino haratago. Lehenetsitako hizkuntza jasotzea ezinezkoa denean eta hizkuntza-parekidetasuna ziurtatu ezin denean, itzultzaile profesionalak ordezeko egoki izan daitezke. Zerbitzu horiek sarri ez daude eskuragarri, garestiak dira, edota ez dira gai behar diren hizkuntza guztietan jarduteko, bereziki urgentzietan. Itzulpengintzan trebatu gabeko langile elebidunei ere sarritan eskatzen zaie momentuko itzulpenak egiteko, edo 24 orduko itzultzaile telefoniko garestietara jo daiteke, esku gabeko ekipamendua eskatzen duena.

Aurrekoetako bat ere izan ezean, paziente eta hornitzaileek anabasa zeharkatu edota, sarri gertatzen den bezala, kontsulta utzi behar izaten dute. Hornitzaileek askotan jotzen dute entrenatu gabeko senide edo boluntarioengana itzultzaile-lanak egiteko, zeinak konfidentzialtasun-, inpartzialtasun- eta itzulpen-akatsak sor ditzakeen<sup>(7)</sup>.

Hizkuntza-teknologiak, marrazki eta sinboloekin batera, arazo horietariko batzuei aurre egiteko modu bat izan daitezke. Hala ere, Google Translate-ren bitartez itzulitako osasun-hizkerako esaldien ia %40 gaizki itzultzen da<sup>(8,9)</sup>, eta itzulpenak ez du zehaztasun nahikoa edo alderagarritasun nahikoa itzultzaile profesional batekin, ezta hizkuntza-parekidetasunezko ingurune batekin ere.

Artikulu honen helburua dela eta, gure asmoa itzultzaile automatikoen (edota makina itzultzaileen) inguruan existitzen den ebidentzia ebaluatzea da, Euskal Autonomia Erkidegoko osasun-sistema publikoan mota horretako tresna baten inplementazioa martxan baita. Euskal Osasun Erakundea (Osakidetza) ezaugarri horietako tresna bat garatzen ari da paziente euskaldunei arreta elebiduna eskaintzeko<sup>(10)</sup>, eta, beraz, gai horren harira planteatzen da azterketa bibliografiko hau.

Euskal Autonomia Erkidegoak espainieraren gaitasun ia absolutua duen errealitate soziolinguistikoa erakusten du, euskal hiztunak %36 inguru direlarik. Ingurune horretan ez dago ia hizkuntza-mugarik bertako (eta espainiar komunitatean jaiotako) paziente eta osasun-langileen artean, alfabetizatu gabeko pertsona heldu edo umeak salbu, non senide edo gurasoek jarduten duten itzultzaile gisa.

## 2. Helburuak

---

Laburbilduz, azterketa honen helburua honakoa da:

1 – Itzultzaile automatiko edota makina itzultzaileen inguruko ebidentzia ebaluatzea, jakinda Osakidetzan (euskal osasun-zerbitzuan) mota horretako tresna bat garatzen ari direla.

## 3. Material eta metodoak

---

Osasun-sistemetako itzultzaile automatikoen erabilera aztertzen zuten ikerketak bilatzeko 2 datu-base ezberdinetan oinarritu ginen: Pubmed eta ClinicalKey.

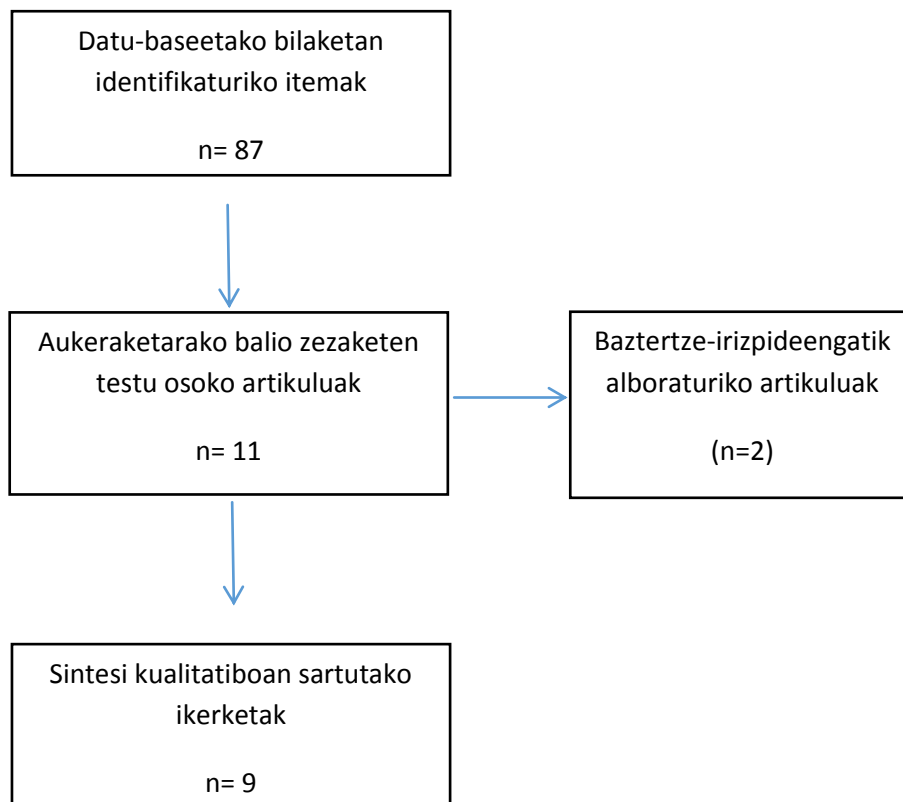
Ikerketa-itea MESH terminoetan “(Automatic[AllFields] OR Machine[AllFields] OR App[AllFields]) AND (Translator[AllFields] OR Interpreter)” zen.

Azterketan kontuan hartzeko, ikerketek honako baldintzak bete behar zituzten (1. irudia):

- Ingelesez publikatua egotea.
- Osasun-arretako ingurunean itzultzaile automatiko edo aplikazioen eragina dokumentatzea.

Baztertze-irizpidea, aldiz, honakoa zen:

- Itzultzaile automatikoen erabilera paziente gabeko ingurunean ebaluatzea.



### **1. irudia.** Bilaketatik sintesi kualitatiborako pausoak.

Aplikazio klinikoan gai honek duen ebidentzia urria ikusita, ez zen inongo bazterketa-irizpiderik ezarri populazio, interbentzio, konparazio, emaitza edo diseinu aldetik. Beraz, alborapen-arriskua bera ere indibidualki aztertu da.

Literatura-bilaketak PubMed-en 81 artikulua eta publikazio erakutsi zituen eta 6 ClinicalKey-n. Titulu eta abstract horiek iragazi ziren ikusteko ea ikerketaren objektua itzultzaile automatikoen osasun-arretako ingurune erabilera edo aplikazioa ote zen.

11 artikulua betetzen zituzten ikerketan onartzeko baldintzak: 1 alboratu zen oraindik garatu gabeko interbentziozko entsegu baten protokolo bat planteatzen zuelako. Makina bidezko itzulpenen errore motek postedizioko esfortzuan zuten eraginari buruzkoa zen beste ikerketa bat ere baztertu zen, hasieran osasun-inguruneetan makina bidezko itzultzaileek izan dezaketen aplikazioaz teorizatzen duen arren, testu osoa osasun-arretako inguruneetatik at kokatzen delako.

Gainerako 9 ikerketak azterketarako egokitzat jo ziren, 8 ikerketa eta aditu baten iritzi bat. Azterketa honen emaitzak haieran oinarritzen dira.

Ikerketa horien datuak ezaugarri hauen arabera bildu genituen: autore, publikazio-urte, ikerketa mota, artikuluan aztertutako itzultzaile automatiko, aplikazio edo tresna, sistema kopuru, esaldi, paziente, azterturiko abstract eta autoreek emaitzak publikatzeko emaniko neurriaren arabera.

Emaitzak publikatzeko neurririk ohikoena medikuek, pazienteek eta itzultzaile profesionalek betetako inkestak ziren, ikerketa bakarrarekin zeinak bere emaitzak itzulpenean emandako akats kopuruetan (dentsitate lexikoak doitua) eta beste batekin itzulpen-itemei esleituriko BLEU (Bilingual Evaluation Understudy, hitz komun kointzidentzia kopurua neurtzen du) eta PER (Position-Independent Word



ErrorRate, itzulpen egokirako beharrezko aldaketa kopurua neurtzen du) puntuazioetan erakusten zituztenak.

Ondorioz, ikerketa ezberdinen emaitzak ezin ziren konbinatu, ezta inongo analisi kuantitatibo gehigarriak egin. Bakoitza modu independentean erakutsi eta azaldu behar da.

#### 4. Emaitzak

Guztira 9 ikerketa errebisatu genituen. Bat 10 itzulpen-aplikaziori eta horiek zuteneguneroko komunikazio klinikorako egokitasunari buruzkoa zen. 6 ikerketa kualitatibo zeuden itzulpen-tresna edo -aplikazioaren gaineko bukaerako inkesta banarekin. Azken biak interbentzio-ikerketak ziren, domeinu medikoan itzulpen-kalifikazio eta erroreak kalifikatzen zituztenak.

Ikerketa kualitatiboetan azterturiko aplikazio edo tresnak Google Translate, iTranslate, BabelDr, Clicker 5 software interaktibo aplikazioa, Komunikazio Digitalerako Asistentzia Tresna (KDAT edo DCAT artikulua originalean) eta MedSLT Open Source ziren. Artikulu guztiek azpimarratzen dute azterturiko tresnen helburua: hizkuntza-mugak gainditzea, bereziki populazio migratzaileen kasuan, bertako hizkuntzarekiko gaitasuna duten pazienteekin alderatuta, itzultzaile profesionalekin alderatuta kostuefektiboa izango den alternatiba bat eskaintzeko asmoz.

Diotenez, hornitzaileek itzultzaile profesionalak behar baino gutxiago erabiltzen dituzte, baita eskuragarri dituztenean ere, nahiz eta beraiekin lan egitearen onurak ulertu. Medikuek sarri beren pazienteen hizkuntza-gaitasunen aurreikuspen handiegia egiten dute, edo kontsultan “aurrera egin” dezakete, itzultzaileen erabilerak kontsulta bera luzatu eta konplikatu duela asumituz.

Google Translate eta Moses ziren interbentzio-ikerketetan azterturiko Makina Itzulpen (MI edo MT) tresnak.

Panayiotouet *al.*-ek aditu-iritzi ikerketa bat egin zuten Apple-ren AppStore-n bilaturiko 10 itzulpen-aplikaziori buruz, eta datuak publikatu, determinatzeko tresna horiek izan zezaketen egokitasuna itzultzaile bat beharrezko izango ez zen osasun-arretako ingurune bat ahalbidetzeko <sup>(11)</sup>.

Itzulpenetan eta osasun-arretako kultura arteko komunikazioan adituak ziren 2 profesional independentek ebaluatu zituzten aplikazioak.

10 aplikazioetatik, itzulpen-aplikazioen edukia dela eta, %73ko (44/60) adostasuna lortu zen ebaluatzaileen artean. Bat ere ez zen guztiz egokia.

Ohartarazten dute teknologia hauekiko gehiegizko dependentziak eta konfiantzak eragin negatiboa ekar dezaketela kalitatezko osasun-arretaren horniduran eta horrekiko adostasunean.

Ondorioztatzen dutenez, ikerketak ez zuen argitu azterturiko aplikazioen itzulpen-zehaztasuna eta egokitasun kulturala, osasun-arretako inguruneetan erabili nahi izatekotan kontu handiz egin behar dela identifikatuz, itzultzaile profesionalen ordezko egokiak ez direlako, eta egungo osasun-politikek partzialki edo bere osotasunean debeka ditzaketelako.

Starlanderet *al.*-ek segurtasunerako kritiko diren aplikaziorako Open Source-ko hizkera medikoaren itzulpen-sistema bat (MedSLT) deskribatu eta ebaluatzen dute <sup>(12)</sup>. Ahozko galderak ingelesetik frantsesera, japonierara eta finlandierara itzultzen ditu, 3 azpidomeinumedikoren barruan (buruko mina, bularreko mina eta min abdominala).

Tresna horren helburua da esaldi finkoen itzulpena, batetik, eta hizkuntza motibaziozko gramatikan oinarrituriko prozesamendua, bestetik, modu azkarrean uztartzen saiatzea eta SLTrekin (*SpokenLanguage Translator*, ahozko hizkuntza-itzultzailea) konparatzea, oinarrian esaldien itzultzaile gisa jokatu. Aldiz, ikerketaren helburua da Hizkuntza Eredu Estatistikoa (SLM, *Statistical Language translation Model*, ahozko itzultzaile gehienek erabilia) Hizkuntza Eredu Gramatikal (GLM) batekin konparatzea da, guztira 870 esaldi erregistraturekin.

Adrian Unai Zelaia Arieta-Araunabeña, Aitor Montes Lasarte, Unai LarrinagaTorrontegui

Frantseserako eta japonierarako itzulpenen onargarritasuna hiztun natiboak ziren epaileek aztertu zituzten: 6 epaile frantses eta 3 japoniar. Epaileek itzulpenak “onak”, “txarrak” edo “balekoak” gisa sailkatu behar zituzten.

Eredu gramatikalak errekonozimendu-akats gutxiago egiten dituela ondorioztatzen dute (%54,6) eredu estatistikoarekin alderatuta (% 59,8), itzulpen on gehiagorekin (% 34,4 eta % 30,8rekin alderatuta frantsesez eta % 36,4 eta % 32,8 japonieraz).

Müllereta *al.*-ek Komunikazio Digitalerako Asistentzia Tresna (DCAT) bat garatu zuten 19 hizkuntza eta dialektorentzat, eta ekintzara bideraturiko hurbilketan oinarrituriko erabiltzailean zentratutako aplikazio interkultural gisa definitu<sup>(13)</sup>. Alemanian kokaturiko 36 paziente errefuxiaturekin probatu zuten aplikazioa.

DCATek pazienteak osasun-arazo ezberdinen inguruko galderen algoritmo batean zehar gidatzen ditu, eta osasun-hornitzailea ikusi aurretik pazienteek itxarongelan erabil dezaten diseinaturik dago. Esaldi finkoak audiozko laguntzarekin eta piktogramekin erabiltzen ziren, testu librea sarritan luzeegi kontsideratzen baitzen.

Inkesta baten bitartez pazienteei DCAT ondo erabiltzeko gai (% 80k baietz, % 16,7k partzialki, % 2k ezetz) izan ote ziren erantzuteko eskatu zitzaizen eta ea beren oinarritako informazio medikoa sartzeko gai izan ote ziren (% 72,2k baietz, %0k partzialki eta % 8,3k ezetz).

Aplikazioaren onargarritasuna eta erabilgarritasuna altua izan zen, eta osagarri egoki gisa agertzen da kontsultan pazienteak itzultzaile (profesionala izan zein ez) batez lagundurik dagoenean.

Somerset *al.*-ek oinarri digitaleko komunikazio laguntzaile bat aurkezten dute, zeina garatu eta aztertu duten langileak aukeraturiko ingelesezko galderak pazientearen hizkuntzara itzultzeko testu eran zein hizkera digitalizatuan, sinboloek eta ingelesezko testuak lagunduta<sup>(14)</sup>.

Komunikazio Alternatibo eta Gehigarriak (KAG edo AAC ingelesez) informazioa adierazteko edo jasotzeko zailtasunak dituzten indibiduen gaitasun komunikatiboak hobetzeko helburua du.

Asma zuten paziente somaliarrekin 26 kontsulta simulatutan zentratu ziren, asmaren tratamenduan esperientziadun 3 mediku jeneralista eta 6 erizainekin lan eginez. Marrazki eta sinbolo espezifikoak diseinatu zituzten, kulturalki eta linguistikoki ingurunera egokituak.

Ikerketarako inplementaturiko CrickSoftware-ren Clicker 5-a erabili zen, testua, grafikoak eta hitzak eskaintzen dituen software interaktiboa erabili nahian.

Langileak oso edo nahikoasebeteak zeuden kontsulten %65ean (17/26), pazienteak berdin sentitu zirelarik kontsulta guztietan (26/26).

Klinikoek konbentzimendu gutxiago adierazten dute pazienteen kezka argitu izanaz edota itzultzaile batekin baino haratago iritsi izanaz. Baliteke pazienteak bereziki optimistak izatea soilik norbait haien arazoaz interesatu zelako. Kontuan hartu behar da ingeles hiztunak diren pazienteekin ere ez dagoela %100eko asebetetze-mailaren ziurtasunik, non komunikazioa (artikulu honetan ulertzen den gisan) ez den gai nagusia. Ikertzaileak kritikatzeko dituzte gai honekiko eginko hurbilketa klinikoetan eta teknologian oinarrituegia dagoelako.

Ingelesetik italierarako dohainezko online itzultzaile batek testu medikoak itzultzeko duen errendimendua kuantitatiboki neurtzen saiatu zen Daniele<sup>(15)</sup>. Gai horri heltzeko, 111 abstract ebaluatu zituen.

Abstract guztiak ingelesetik italierara itzuli ziren *Google Translate* erabiliz. Itzulpenetako akatsak zuzenean ikerketaren autoreak aztertu zituen, *Text Analyzer* erabiliz testu-estatistika zehatzetarako.



Itzulpen-akats kopurua 10-38 artean zegoen, batez beste 21. Akats kopuru totalaren ehunekoak % 6,3- % 34,8 artean zegoen, batez beste % 15. Itzulpen-akatsen eta lexiko dentsitatearen arteko korrelazio zuzena aurkitu zuen.

*Google Translate*-k abstract medikoen gisako idatzi oso akademikoen itzulpenenean errendimendu nahiko ona zuela ondorioztatu zuten

Spechbach *et al.*-ek BabelDr deritzon ahozko esaldi finkoen itzultzaileari buruzko ikerketa publikatu zuten, zeina, gainerako esaldi finkoen itzultzaileen gisan, aurretiaz definituriko aurreitzulitako esaldien zerrenda batean oinarritzen den<sup>(16)</sup>.

Frantses hiztunak ziren 12 medikurekin ikerketa bat burutu zuten Genevan. BabelDr tresna erabiliz, 2 paziente arabiar estandarizaturiko diagnostiko bat egitea eskatu zitzaizkien medikuei.

Orokorrean, parte-hartzaileek gehiago jardun zuten ahoz idatziz baino, tresna hauetan ahozkotetasunak duen funtzionalitatea frogatuz.

Urgentzia-ingurune batean, itzultzailearik ez dagoenean, ahozko esaldi finkoen itzultzaileak pazientearen informazioa jasotzeko alternatiba egokia izan daitezkeela ondorioztatzen dute.

Chen *et al.*-ek, 2016an *Google Translate*-k diabeteserako heziketa-materiala itzultzeko zuen eraginkortasuna ebaluatzeko, 6 itzultzaile kualifikatu errekrutatu zituzten (3 espainiararako eta 3 txinerarako)<sup>(17)</sup>.

Cronbach alfa balioek ebaluatutako bi taldeen emaitzen baliokidetasun altua erakutsi zuten: 0,919 espainiarazko ebaluatzaileentzat eta 0,972 txinerazkoentzat. *Google*-k esaldi sinpletarako itzulpen zehatzak eskaintzen ditu. Hala ere, itzulpen ezegokirako aukerak igo egiten ziren ingelesezko esaldiek ulermen-maila altuagoa eskatzen zutenean.

Artikuluak itzulpenen estandarizaziorako 10 gako eskaintzen ditu, *Somerset et al.*-en gisan, tresna mota hau garatzeko prozesuan hizkuntza-komunitatea kontuan hartzea, hizkuntza eta kultura laguntza talde bat izatea eta gai horretan adituak direnen zein itzultzaile profesionalen laguntza izatea.

Chen *et al.*-ek baita, 2017an *iTranslate*-ren kalitatea aztertu zuten, 2016ko material eta helburu berekin (diabeteserako heziketa-materialarekin): itzulpenetarako ahots-errekonoizimenduko tresna bat aztertzeko ikerketa pilotu baten bitartez ikerketa-hutsunea betetzea<sup>(18)</sup>.

Cronbach alfa balioek ebaluatutako bi taldeen emaitzen baliokidetasun altua erakutsi zuten: 0,920 espainiarazko ebaluatzaileentzat eta 0,971 txinerazkoentzat. *iTranslate*-kor okorrean giza itzultzaileen pareko itzulpen-zehaztasuna zuela aurkitu zuten, esaldi sinpleen kasuan bederen. Esaldi zailtan akats gehiago zituen.

Ikerlarien aburuz, mugikorretako itzulpen-aplikazioak tentuz erabili behar dira, gehigarri gisa, giza itzultzaileak ordezkatu gabe, erabiltzaileei arreta-neurri egokiak ematearekin batera.

Pecina *et al.*-ek domeinu medikoko hizkuntza gurutzaturiko informazio-berreskurapenean (IB) itzulpen automatikoak duen egokitzapena aztertu zuten<sup>(19)</sup>. Moses zen haien itzulpen automatikoko (IA) sistema eta informazio-berreskurapenerako sistema Lucene bilaketa-tresnan oinarritzen zen. Esperimentuak 3 hizkuntza-bikotetan egin eta ebaluatu zituzten: txekiera-ingelesa, aleman-ingelesa eta frantses-ingelesa. Beren sistemak hasierako erreferentzia hobetzen duela diote, *Google Translate* eta *Microsoft Bing* itzultzaileekin batera, hizkuntza-bikote guztietan eginiko konparaketa zuzenetan. Oinarritako BLEU kalifikazioa 26,59tik 41,45era igo zen txekiera-ingelesetan, 23,03tik 40,82ra aleman-ingelesetan eta 32,67tik 40,82ra frantses-ingelesetan. Batez beste % 55eko hobekuntza da. Test multzo honen IB errendimenduaren aldetik, PER eskalaren bitartez neurtua, oinarriaren hobekuntza esanguratsua lortu da frantses-ingelesetan. Txekiera-ingelesetan eta aleman-ingelesetan, IA kalitate-igoerak ez dakar IB emaitza hobetzerik.

Adrian Unai Zelaia Arieta-Araunabeña, Aitor Montes Lasarte, Unai LarrinagaTorrontegui

Badirudi itzulpen hobeak ez dakarrela derrigor informazio-berreskurapen hoberik. Azaltzen dutenez, sistema hauek normalean ez dute hizkera natural edo gramatikalki egokia duen inputik eskatzen. Hitzen ordena eta funtzioak IBrako hutsalak izan daitezke.

Ikerketen ezaugarriak 1. taulan ageri dira.

**1. taula.** Azterketan sartu diren ikerketen ezaugarriak.

Autorea k	Ur-tea	Ikerketa mota	Tool / applicationte sted	N	Emitza-neurria
Starland eret al.	2005	Ikerketakualitat iboa. Inkesta	MedSLT Open Source(GLM vs SLM)	12 ingeles hiztun Grabaturiko 870 enuntziatu	Itzulpen ona/nahikoa/txarra
Somerset al.	2007	Ikerketa kualitatiboa. Inkesta	Clicker 5 software aplikazio interaktiboa	Somaliako paziente asmatikoak	Oso/nahiko/asebeteta/asegabe/bat ere ez
Pecinaet al.	2014	Interbentziozkoa	Moses	8 IA sistema-eredu	Bleuscore. PER score
Chen et al.	2016	Ikerketa gurutzatua Inkesta kualitatiboa	Google Translate	6 esaldi	1-5 artean: jarioa, egokitasuna, esanahia eta akatsaren larritasuna
Chen et al.	2017	Ikerketa gurutzatua Inkesta kualitatiboa	iTranslate	3 esaldi	1-5 artean: jarioa, egokitasuna, esanahia eta akatsaren larritasuna
Spechbach	2019	Ikerketa kualitatiboa Inkesta	BabelDr	12 mediku, 2 paziente arabiar	(ez)adostasuna 1-5 artean, erantzun positibo edo negatiboak
Daniele	2019	Interbentziozkoa	Google Translate	111 abstract	Itzulpen-akats kopurua
Panayiotou et al.	2019	Adituaren iritzia	Apple store itzulpen-aplikazioak	10	Egokitasun kliniko altua/baxua
Müller et al.	2020	Ikerketa kualitatiboa Inkesta	DCAT	36 paziente errefuxiatu	Bai/partzialki/ez Erabilgarritasun- eta onargarritasun-maila

GLM: Gramatika Estatistikoaren itzulpen Eredua

SLM: Statistical Language Translation Model

IA: itzultzaile automatikoa

Ebaluaturiko tresnen emaitzak 2. taulan ageri dira.



**2. taula. Ebaluaturiko tresnen emaitzak.**

Autorea	Tresna Aplikazioa	Hizkuntza kopurua	Idatziz/ahoz/irudiak/sinboloak (testu librea, esaldi itxiak, bai/ez formatua)	n	Emaitza
Starlandere <i>t al.</i>	MedSLT Open Source	3	Hizkera-itzulpena. GLM vs SLM	870 esaldi	GLM eraginkorragoa (% 36,4 ondo)
Somerset <i>al.</i>	Clicker 5 interactive software application	2	Testutik testura + ahozkoa + sinboloetara	26	Asebetetze altua (% 65-100)
Pecinaet <i>al.</i>	Moses	4	Idatziz	8	Batez beste BLEU +8 PER frantsesez bakarrik
Chen <i>et al.</i>	Google Translate	3	Idatziz	6	Txineraz gizakia > Google Espainieraz gizakia dif. ez Google-rekin
Chen <i>et al.</i>	iTranslate	2	Ahozkotik testura	3	Gizakia difez. iTranslate-rekin
Spechbach	BabelDr	2	Ahozkotik testura + ahozkoa	12	19 (-) 116 (+)
Daniele	Google Translate	2	Testutik testura	111 abstract	Batez beste NTTE 21/ BbPTTE % 15.
Panayiotou <i>et al.</i>	Apple storetranslationApps		Ahozkotik ahozkoa +ahozkotik testura + testutik testura	10	Egokitasun altua 2/10
Mülleret <i>al.</i>	DCAT	19	Idatzizko esaldi finkoak + audio-euskarria + piktogramak	36	Erabiltzeko gai(% 80), input egokia (% 72)

GLM: Gramatika Estatistikoaren itzulpen Eredua

SLM: *Statistical Language Translation Model*

IA: itzultzaile automatikoa

NTTE: *number of translation errors*, itzulpen-akatskopurua

PTTE: *percentage of translation errors*, itzulpen-akatsen ehunekoa

±: eta, emaitza positiboa

–: emaitza negatiboa

≥: hobe

**5. Eztabaida**

Sei artikuluk beren emaitzak inkesta gisa publikatu zituzten, ingurune klinikoan emaniko egoera gutxirekin (36 paziente gehienez) edota ingurune klinikotik kanpoko erantzun ugariarekin (itzulpena ebaluatzen du inkestak eta ez arreta).

2 ikerketak itzulpen automatikorako aplikazioen esku-hartzea lantzen dute, itzulpenaren kalitatea ebaluatzeko, pazienterik gabe.

Azken ikerketak 15 aplikazioen aplikazio klinikoaren egokitasuna ebaluatzen du.

Azterketako 8 ikerketak emaitza positiboak lortu zituzten beren ikerketetan, aditu-iritziaren artikuluan salbu, zeinak argi eta garbi ondorioztatzen duen azterturiko aplikazioak ez direla osasun-arretarako egokiak.

Adrian Unai Zelaia Arieta-Araunabeña, Aitor Montes Lasarte, Unai LarrinagaTorrontegui

Itzulpen klinikorako aplikazio ezberdinen paziente potentzialetako erabilerak izan dezaketen onargarritasuna eta asebetetzea aztertzean, 3 aplikaziok (Babel Dr, Clicker 5 eta DCAT) asebetetze altua, erabilerarako gaitasun altua eta input egokia erakutsi zuten.

Pazienteek app itzultzaile automatikoekiko asebetetze-maila altua adierazi zuten, segur aski aurretiazko hizkuntza-mugen eta komunikazio-paradigma berriaren arteko konparazioarengatik. Migratzaileentzat osasun-langileekin modu eraginkor batean komunikatzea eta haien arretan modu egokian eragitea zein zaila den adierazi dute.

Profesionalek hizkuntza-mugekiko ezagutza gutxiago dute eta hizkuntza-parekotasuneko inguruneetara daude ohituak. Horrela, aplikazio hauek osasun-arretan dituzten mugez errazago jabetzen dira, eta haien erabilera eta inplementazioari buruzko ikuspuntu errealista eta pesimistagoa dute. Honakoa behin eta berriz azpimarratzen da ikerketetan: tresna hauek benetakokontsultetan erabilgarriak izan daitezke, betiere itzultzaile profesionalik eskuragarri ez badago.

Google Translate eta iTranslatedirela eta, tresna hedatuagoak eta merkeagoak direnak, maila akademiko altuko idatzien itzulpen-maila nahiko ona erakutsi arren, esaldi labur batzuetan agerturiko akatsek iradokitzen dute mugikorrorako itzulpen-aplikazio horiek tentuz erabili behar direla, sekula giza itzultzaileak ordezkatu gabe<sup>(20,21)</sup>.

Emitza positiboak izan zein ez, ezein aplikaziok ez lukeela itzultzaile profesionalik ordezkatu behar da ikerketa guztien abiapuntua. Denek aitortzen dute zein garrantzitsuak diren hizkuntza-mugak osasun-arretan, eta zein mugatuak dauden aplikazio hauek jardun kliniko eraginkorrerako. Are gehiago, haietariko inor ez da ausartzen soilik aplikazio hauetan oinarrituriko egoera klinikorik ezartzera, eta haien egokitasuna eta ezargarritasuna beste tresnarik ez dagoenerako gordetzen dute.

Artikulu hauen ebidentziaren kalitatea baxua da, eta, beraz, ikerketa hauen baieztapen eta ondorio denak zuhurtziaz onartu behar dira, eta tresna gehigarri gisa soilik erabili. Itzulpen automatikoa ez daiteke, beraz, gaitasun edota segurtasun kulturalaren estandarretan barnebildu, hizkuntza-mugei aurre egiteko azken aukera gisa baino<sup>(22,23)</sup>.

Tresna hauek eremu elebidun batean izan dezaketen egokitasuna dela eta, ez da inongo informaziorikagertu. Ikerketa guztiak populazio migratzaileak aztertzen dituzte, eta itzultzaile automatikoek nola lagun dezaketen hizkuntza-mugei aurre egiten. Segurtasun kulturalaren eta hizkuntza-parekidetasunaren kontzeptuak ez dira inongo artikulutan agertu<sup>23</sup>. Osasun-erakunde eta autoritateak dira kulturalki segurua den arreta eskaintzearen arduradunak, pazienteek eta haien komunitateek hala definituta, eta bidea eginez neurtuta osasun-ekitatea lortu arte. Horretarako, ezarriak dauden botere-harremanak kritikatzeko prest egon behar dute, gainerakoek kulturekiko "konpetentzia" garatzea lehenetsi baino gehiago, beren sistema kulturala eta kultura bera zalantzan jartzeko prest egon behar dute.

Gizarte elebidun batean hizkuntza minorizatua hitz egiten duen pazienteak ez du komunikatzeko gaitasun eta erraztasun bera edukiko osasun-hornitzaile ez-elebidun batekin edo bere komunitateko kide batekin (edo gutxienez bere lehen hizkuntza dakien norbaitekin)<sup>(24)</sup>.

Komunitateak beren osasun-arretako hizkuntza-beharren identifikazio eta konponbideen planifikazioan inplikatu behar dira<sup>(25)</sup>, segurtasun kulturalaren eta eskaintza aktiboaren ereduari jarraitu nahi bazaio. Bestela, pazienteen ongizatea osasun-arretaren horniduran inplikaturiko industria ezberdinen (informatikoa, farmazeutikoak, etab.) lehenetsun ekonomiko edo bestelako lehenetsun teorikoek ordezkatu dezakete, pazientearengan zentratutako ereduak aldatuz. Osasun-ikerketek "goldstandard"-en arabera eta haiei jarraiki funtzionatu beharko lukete, eta ez metodologia korapilatsu edo ez-zientifikoaren arabera.

Komunitate elebidun ezberdinetan gerta daitezkeen identifikatu gabeko beharren diagnosi eta planifikazioa eskatzen du ikerketaren etikak, soilik egungo "goldstandard"-ak hobetu edota berdintzeko asmoa duten alternatibak eskaintzeko. Bere burua errespetatu nahi duen edozein



hizkuntza- edo kultura-programak ezin dezake bere burua hizkuntza-mugen etsipenezko konponbideak bilatzera mugatu, egun gaiaren inguruan dagoen ebidentzia kontuan hartu gabe.

Gai honi buruz gehiago ikertu behar da, eta ebidentzia gehiago eraiki implementazio-fase hipotetiko bat helburu bada. Baina berriro ere, azterketa honek argi erakusten du egungo paradigma eta “goldstandard”-ak oso urruti daudela ingurune elebidun batean itzultzaile automatiko bat ezartzetik.

## 6. Ondorioak

---

Makina bidezko edo automatikoak diren itzultzaileek ez dute osasun-arretan beren kontura ezartzeko eraginkortasun, zehaztasun edo egokitasunik erakutsi, nahiz eta hizkuntza-mugen kasuan tresna osagarri gisa erabil daitezkeen.

Itzulpena ez da aukera egokia ingurune elebidunetarako, non ez dagoen hizkuntza-mugarik. Hizkuntza-parekotasuna da komunikaziorako kulturalki segurua den espazioak sortzeko bidea, non itzulpena azken aukera gisa gorde beharko litzatekeen.

Hizkuntza-mugak direla eta, egungo “goldstandard”-a hizkuntza-parekotasuna da giza itzultzailea baino, bitartekari gabeko komunikazio zuzen eta osoa ematea ahalbidetzen baitu osasun-langileen eta pazienteen artean.

Osasun-sistemek konpetentzia kulturala, segurtasun kulturala, pazientearengan zentratutako arreta eta hizkuntza-parekidetasuna txertatu beharko lituzkete abangoardiako kontzeptu gisa hizkuntza-mugei aurre egitean.

Ikerketa gehiago egin behar dira gai honen inguruan, ebidentzia gehiago behar dira, implementazio-fase bat eraiki nahi bada. Egungo paradigma eta “goldstandard”-a oso urruti daude ingurune elebidun batean itzultzaile automatiko bat inplementatzetik.

## 7. Interes-gatazkak:

---

Lanaren egilea eta zuzendariak eremu elebidun bateko osasun-sistema publikoko langile euskaldunak dira, eta, beraz, ikerketa honen emaitzak beraien egunerokotasunean eragina izateko aukerak interes-gatazkak sor diezazkieke.

Autoreek ez dute inongo interes ekonomikoren gatazkarik ikerketa honetan.

## 8. Bibliografia

---

- 1- Kripalani S, LeFevre F, Phillips CO, Williams MV, Basaviah P, Baker DW. Deficits in communication and information transfer between hospital-based and primary care physicians: implications for patient safety and continuity of care. JAMA. 2007 Ots 28;297(8):831-41.
- 2- Institute of Medicine (US) Committee on Understanding and Eliminating Racial and Ethnic Disparities in Health Care. Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care. Smedley BD, Stith AY, Nelson AR, editors. Washington (DC): National Academies Press (US); 2003.
- 3- Taylor G. Migrants and refugees. Hemen: Papadopoulos I, ed. Transcultural health and social care: development of culturally competent practitioners. Edinburg [etab.]: Churchill Livingstone Elsevier, 2006: 45-63.

Adrian Unai Zelaia Arieta-Araunabeña, Aitor Montes Lasarte, Unai LarrinagaTorrontegui

- 4- United Nations High Commissioner for Refugees (UNHCR) (ed.). UNHCR Statistical Yearbook 2012 [Internet]. Geneva: UNHCR;2013 [Kontsulta: 2021-02-22]. 172 or. Eskuragarri: [https://www.unhcr.org/static/statistical\\_yearbook/2012/2012-12\\_SYB12\\_FULLBOOK\\_low.zip](https://www.unhcr.org/static/statistical_yearbook/2012/2012-12_SYB12_FULLBOOK_low.zip)
- 5-Tervalon M, Murray-Garcia J. Cultural humility versus cultural competence: a critical distinction in defining physician training outcomes in multicultural education. *J Health Care Poor Underserved* 1998; 9:117-25.
- 6-Betancourt et al., Defining Cultural Competence: A Practical Framework for Addressing Racial/Ethnic Disparities in Health and Health Care. *Public Health Reports / July-August 2003 / Volume 118*. Association of Schools of Public Health. P 293- 303. USA.
- 7-Meeuwesen L, Twilt S, ten Thije JD, Harmsen H. "Ne diyor?" (What does she say?): informal interpreting in general practice. *Patient EducCouns*. 2010 Aza;81(2):198-203.
- 8- Patil S, Davies P. Use of Google Translate in medical communication: evaluation of accuray. *BMJ*. 2014 Abe 15;349:g7392.
- 9- Bouillon P, Gerlach J, Spechbach H, Tsourakis N, Halimi S. BabelDr vs Google translate: a user study at Geneva University Hospitals (HUG) [Internet]. Hemen:20th Annual Conference of the European Association for MachineTranslation (EAMT). Praga:EAMT, 2017 [Kontsulta: 22-ots-2021]. 7 or. Eskuragarri: <https://archive-ouverte.unige.ch/unige:94511/ATTACHMENT01>
- 10- Osaraba [Internet]. Vitoria-Gasteiz: Osaraba. Itzulbide proiektuaren aurkezpena Araba ESiko profesionaleri; 2019 [kontsulta: 2021-2-22]. 1 or. Eskuragarri:<http://osaraba.eus/eu/itzulbide-proiektuaren-aurkezpena-araba-esiko-profesionalei/>
- 11- Panayiotou A, Gardner A, Williams S, Zucchi E, Mascitti-Meuter M, Goh AM, You E, Chong TW, Logiudice D, Lin X, Haralambous B, Batchelor F. Language Translation Apps in Health Care Settings: expert Opinion. *JMIR MhealthUhealth*. 2019 Api 9;7(4):e11316.
- 12- Starlander M, Bouillon P, Rayner M, Chatzichrisafis N, Hockey BA, Isahara H, Kanzaki K, Nakao Y, Santaholma M. Breaking the language barrier: machine assisted diagnosis using the medical speech translator. *Stud Health Technol Inform*. 2005;116: 811-6.
- 13- Müller F, Chandra S, Furaijat G, Kruse S, Waligorski A, Simmenroth A, Kleinert E. A Digital Communication Assistance Tool (DCAT) to Obtain Medical History from Foreign-Language Patients: Development and Pilot Testing in a Primary Health Care Center for Refugees. *Int J Environ Res Public Health*. 2020 Feb 20;17(4):1368
- 14- Somers HL, Caress AL, EvansDG, Johnson MJ, LovelHJ, Mohamed Z. A computer-basedaid for communication between patients with limited English and their clinicians, using symbols and digitisedspeech. *Int J MedInform*. 2008 Abu;77(8):507-17.
- 15- Daniele F. Performance of an automatic translator in translating medical abstracts. *Heliyon*. 2019 Oct 21;5(10):e02687. doi: 10.1016/j.heliyon.2019.e02687. Somers et al., A computer-based aid for communication between patients with limited English and their clinicians, using symbols and digitized speech. *International Journal of Medical Informatics* 77 (2008) 507-517.
- 16- Spechbach H, Gerlach J, MazouriKarker S, Tsourakis N, Combescure C, Bouillon P. A Speech-Enabled Fixed-Phrase Translator for Emergency Settings: Crossover Study. *JMIR Med Inform*. 2019 May 7;7(2):e13167.
- 17- Chen X, Acosta S, Barry AE. Evaluating the Accuracy of Google Translate for Diabetes Education Material. *JMIR Diabetes*. 2016 Eka 28;1(1):e3
- 18- Chen X, Acosta S, Barry AE. Machine or Human? Evaluating the Quality of a Language



Translation Mobile App for Diabetes Education Material. JMIR Diabetes. 2017 Eka 29;2(1):e13.

19- Pecina P, Dušek O, Goeuriot L, Hajič J, Hlaváčová J, Jones GJ, Kelly L, Leveling J, Mareček D, Novák M, Popel M, Rosa R, Tamchyna A, Urešová Z. Adaptation of machine translation for multilingual information retrieval in the medical domain. ArtifIntell Med. 2014 Uzt;61(3):165-85.

20- Moberly T. Doctors are cautioned against using Google Translate in consultations. BMJ. 2018; 363: k4546.

21- Betancourt JR, Green AR, Carrillo JE, Ananeh-Firempong O 2nd. Defining cultural competence: a practical framework for addressing racial/ethnic disparities in health and health care. Public Health Rep. 2003 Uzt-Abu;118(4):293-302

22- Anderson LM, Scrimshaw SC, Fullilove MT, Fielding JE, Normand J; Task Force on Community Preventive Services. Culturally competent healthcare systems. A systematic review. Am J Prev Med. 2003 Api;24(3 Sup):68-79.

23- Curtis E, Jones R, Tipene-Leach D, Walker C, Loring B, Paine SJ, Reid P. Why cultural safety rather than cultural competency is required to achieve health equity: a literature review and recommended definition. Int J Equity Health. 2019 Aza 14;18(1):174.

24- Czapka EA, Gerwing J, Sagbakken M. Invisible rights: barriers and facilitators to access and use of interpreter services in health care settings by Polish migrants in Norway. Scand J Public Health. 2019 Aza;47(7):755-764

25- Lowell A, Kildea S, Liddle M, Cox B, Paterson B. Supporting aboriginal knowledge and practice in health care: lessons from a qualitative evaluation of the strong women, strong babies, strong culture program. BMC Pregnancy Childbirth. 2015 Ots 5;15:19.

## 10. Eranskinak

---

### 1. eranskina:

Bilaketa: (Automatic[All Fields] OR Machine[All Fields] OR App [All Fields]) AND (Translator[All Fields] OR Interpreter). Pubmed: 81 10 --> 8

4. A Digital Communication Assistance Tool (DCAT) to Obtain Medical History from Foreign-Language Patients: Development and Pilot Testing in a Primary Health Care Center for Refugees.

6. [Performance of an automatic translator in translating medical abstracts.](#)

12. A Speech-Enabled Fixed-Phrase Translator for Emergency Settings: Crossover Study.

13. [Language Translation Apps in Health Care Settings: Expert Opinion.](#)

32. [Machine or Human? Evaluating the Quality of a Language Translation Mobile App for Diabetes Education Material.](#)

36. [Evaluating the Accuracy of Google Translate for Diabetes Education Material.](#)

43. [Adaptation of machine translation for multilingual information retrieval in the medical domain.](#)

59. Breaking the Language Barrier: Machine Assisted Diagnosis Using the Medical Speech Translator

Adrian Unai Zelaia Arieta-Araunabeña, Aitor Montes Lasarte, Unai LarrinagaTorrontegui

Excluded studies: Identifying the machine translation error types with the greatest impact on post-editing effort. Clinical Key. 6<sup>2</sup>

4. A computer-based id for communication between patients with limited English and their clinicians, using symbols and digitised speech