

Reptes i oportunitats del català en el programari a distància d'aprenentatge de llengües

Balanç de la situació del català en les empreses tecnològiques

La pandèmia de la COVID-19 va empènyer molta gent a buscar fórmules alternatives per seguir el seus processos d'aprenentatge després que les restriccions sanitàries de mobilitat impedissin l'assistència presencial a cursos i tallers. En aquest àmbit, el coneixement de llengües hi té un paper determinant, especialment fora del conjunt d'estudis impartits de sèrie en instituts i universitats. La facilitat i la flexibilitat han estat dos valors clau perquè aquests programes a distància permeabilitzin la societat: lliçons curtes i concretes, interfícies gràfiques de navegació senzilles i amables, així com una disponibilitat integral des de l'escriptori, el mòbil o el portàtil. És per aquesta raó que, durant aquests primers anys de la dècada del 2020, l'ús i la presència de programes a distància d'aprenentatge de llengües han crescut exponencialment.

Ara bé, quina és la situació del català en aquest programari? La seva presència permet aprendre diverses llengües des del català? I, pel contrari, es pot aprendre català des de diverses llengües? Quin paper hi tenen els conjunts de model de llengua massiu (LLM) i la intel·ligència artificial generativa? Per donar resposta a aquests interrogants convé fer una repassada als principals programes del sector i avaluar-ne les fortaleses i debilitats però, sobretot, els reptes i oportunitats que facin plantejar o preveure la incorporació de la llengua catalana.

Duolingo

En primera posició i d'acord amb el volum de descàrregues mundials de setembre de 2023, l'empresa estatunidenca assoleix una quota de mercat del 65 % i és, amb diferència, la gran dominadora del sector, després d'ampliar encara més el marge capdavanter que tenia el febrer de 2022 del 48 %. Amb aquesta carta de presentació, poques coses fan pensar que la incorporació del català no seria possible. El seu catàleg comprèn l'aprenentatge des de 25 llengües d'origen i 39 de destí. Això vol dir que algunes llengües de destí —fins i tot fictícies com el klingon o l'alt valirià!— només es poden aprendre des de l'anglès. En el cas del català, la llengua només es pot aprendre des del castellà, però en cap cas es pot fer servir el català per aprendre una altra llengua. Les nombroses gestions realitzades per Plataforma per la Llengua amb l'empresa han portat a una mateixa resposta: de moment no es preveuen noves incorporacions. Tanmateix, l'èxit del ChatGPT ha fet moure els fonaments de l'empresa i ha signat un acord de col·laboració

amb OpenAI per tal que la versió Duolingo Max estigui vinculada amb el GPT-4. Aquest fet obre la porta a la millora i ampliació del ventall lingüístic gràcies a la cessió de dades, tot i que, de moment, només s'ha consumat en iOS per aprendre francès i castellà des de l'anglès.

Lingokids

En segon lloc, trobem a l'abast un programa destinat a un segment de població molt determinat: tal com indica el nom, és el dels infants. Concretament és una aplicació destinada a un públic infantil comprès entre els 2 i els 6 anys, que el 2022 es va convertir en el programa d'aprenentatge líder entre els més joves. D'origen anglosaxó, aquest programa compta amb un contingut de solvència contrastada, ja que es nodreix del fons editorial d'Oxford University Press. Una altra característica pròpia que el diferencia de gran part de programes és que té una única llengua de destí: l'anglès. D'aquesta forma, es consolida com a eina de referència per a infants no anglòfons d'arreu del món, a partir de les 16 llengües d'origen disponibles. Malauradament, el català no forma part d'aquest grup.

Buddy.ai

En tercer lloc, ens trobem davant d'un programa molt similar a Lingokids: comparteix l'objectiu d'ensenyar anglès com a única llengua de destí i està pensat per a un públic infantil. Ara bé, no compta amb opcions de llengua d'origen perquè només opera en anglès i el públic diana és una mica més gran d'edat: entre els 4 i els 10 anys. Aquest programa, creat el 2017 a San Francisco, té la característica que integra la intel·ligència artificial a través d'un bot conversacional en forma de dibuix animat, que permet a l'usuari poder aprendre a través de diàlegs i jocs basats en la veu. En aquest sentit, no és un producte objectivament interessant per integrar-hi el català, però sí que pot servir com a exemple de bones pràctiques a l'hora de crear un programa anàleg.

Momo

En quart lloc, tornem a trobar un programa per tal que els més petits de la casa aprenguin anglès. Seguint un criteri pedagògic similar a Lingokids, fa del joc la millor eina per engrescar els infants en l'ús i coneixement de l'anglès. Així mateix, absorbeix de Buddy.ai la filosofia monolingüe: és millor aprendre anglès directament de l'anglès i aquesta ha de ser l'única llengua present. Com a dada significativa, es tracta d'una aplicació per a Android que ha irromput en el sector amb molta força en poc més d'un any.

EWA

En cinquè lloc, hi ha un programa destinat per a tots els públics, que té per objectiu l'ensenyament de la llengua anglesa i, parcialment amb cursos, també del francès i el castellà. En el cas de l'anglès, disposa d'un ventall molt ampli de llengües d'origen, concretament 39, però el català no hi és present. Per redreçar la situació, Plataforma per la Llengua ens hem posat en contacte amb l'empresa de Singapur, Lithium Lab. Pte. Ltd., perquè incorporin el català, tant al programa com al web, si bé la integració segueix en tràmit.

Babbel

En sisè lloc, trobem un dels programes degans del sector de l'aprenentatge a distància de llengües, que encara es manté en la classificació dels més descarregats. L'empresa alemanya, domiciliada a Berlín, segueix el patró comú de lliçons curtes i senzilles per a consum ràpid. des d'un punt de vista de l'oferta, té una assortiment de 9 llengües d'origen per aprendre fins a 13 llengües de destí, tot i que no totes 9 permeten estudiar les 13 i cap d'elles no permet aprendre català. Plataforma per la Llengua ens hi hem posat en contacte en més d'una ocasió i no hem rebut resposta per iniciar un acord de col·laboració per tal de poder sumar-hi el català com a llengua d'origen i de destí. No obstant això, des de novembre d'enguany, Babbel ha començat a fer difusió per les seves xarxes socials de la pluralitat lingüística de l'Estat espanyol i, davant de la petició d'usuaris interessats en cursos de català, ha respost que estudiarà la possibilitat d'incloure'l a la seva oferta.

ELSA

En setè lloc, i reulant lleugerament en nombre de descàrregues respecte de l'any passat, hi ha aquest programa de pagament per subscripció, que també té per objectiu l'ensenyament únic de l'anglès a partir de moltes altres llengües. Concretament, aquest nombre ascendeix fins a 59 i hi inclou el català, si bé la interfície es manté en tot moment en llengua anglesa. Més enllà d'aquest èxit, el programa conté un altre valor afegit: la intel·ligència artificial com a eina de perfeccionament de la pronunciació. A partir de l'ús de paràmetres de l'alfabet fonètic internacional, tot fa pensar que, en un futur, els seus límits podrien anar més enllà de l'ensenyament exclusiu de l'anglès i incorporar-hi el català.

Mondly

En vuitè lloc, tot i que perdent un 50 % del nombre de descàrregues de l'any 2022, ens trobem amb el programa de llengües d'abast mundial amb la millor representació del català com a llengua d'aprenentatge. Disposa d'un ric catàleg format per 33 llengües d'origen per a l'estudi de 41 llengües de destí. El paper que hi té el català és molt satisfactori perquè es pot aprendre des de totes les 33 llengües d'origen. I més encara si tenim en compte que disposa d'un assistent en llengua catalana basat en el sistema de bot de conversa. El repte pendent és incloure el català com a llengua d'origen perquè els catalanoparlants puguem aprendre una de les 41 llengües de destí.

Baicizhan

En novè lloc i com a novetat a la classificació, irromp el programa desenvolupat per l'empresa xinesa Chaoyouai Tech. Co. Ltd. Destinat a un públic molt determinat, té la funció de dotar d'eines la comunitat sinoparlant perquè pugui aprendre anglès, seguint els mateixos criteris de comoditat i disponibilitat que imperen en el sector. Com a tret característic, destaca que es pot seleccionar el xinès tradicional o el xinès simplificat com a llengua d'origen. Així mateix, gaudeix d'un fons molt complet, amb més de 200 tesaurus autoritzats per poder donar cobertura, mitjançant plans d'estudi, a tots els nivells d'ensenyament d'anglès oficialment reconeguts.

Busuu

En desè lloc, i amb un tímid increment de descàrregues, trobem aquesta empresa amb seu a Londres. El programa té a disposició 14 llengües d'origen i 14 llengües de destí, però no hi ha una correlació directa entre unes i altres: és a dir, si ets parlant d'una de les llengües d'origen no podràs aprendre totes i cadascuna de les 14 llengües de destí. I davant de la manca de canvis des de la seva estrena, el català segueix sent llengua destacada d'absència.

Altres programes

Fora d'aquesta classificació dels deu programes de llengua més descarregats al món, n'hi ha d'altres que han esdevingut rellevants en el seu moment, però, al llarg dels anys, la seva importància ha anat decaient.

En aquest subgrup, convé destacar un dels programes degans, Rosetta Stone, que integra 28 llengües d'origen i de destí sense que el català hi sigui present. Com a punt destacable, fa uns anys va engegar el programa Llengües en perill, per tal de nodrir amb contingut llengües ameríndies minoritzades com el chitmachá, l'imuktitut, l'inupiak, el navajo o el mohawk.

Un altre programa de renom és Memrise, que segueix mantenint grans expectatives tot i no colar-se al capdavant. Un motiu destacat per pensar en el seu èxit és la vinculació del programa amb el GPT-3 d'OpenAI i, per tant, pot obtenir grans avenços qualitius i quantitius gràcies al conjunt de dades utilitzat en intel·ligència artificial. D'aquest programa en destaca la capacitat d'aprendre català des de l'anglès.

Un altre programa, ferm candidat a incorporar el català en els propers mesos, és Drops. L'empresa suecojaponesa Kahoot! té un ric assortiment lingüístic: 36 llengües d'origen i 49 de destí. Algunes d'aquestes llengües de destí, amb menys parlants que el català, són l'esperanto, el samoà, el hawaià, l'ainu i la llengua de signes anglesa. Com a dada que cal tenir en consideració, l'empresa té habilitat un espai al web perquè la comunitat presenti demandes de cursos en noves llengües.

En darrer lloc, un programa per tenir present és Pimsleur, desenvolupat als Estats Units d'Amèrica. Disposa de 51 llengües d'origen i 51 llengües de destí, però el català no n'és una. Tanmateix, té en catàleg cursos elementals de llengua de signes anglesa, ojibwa i

crioll haitià, fet que fa pensar que el criteri quantitatiu de parlants no és el motiu d'ingrés de cursos de noves llengües. Com a fet destacable, alguns dels cursos del programa estan dotats d'entrenadors de veu que permeten millorar la pronunciació.

Conclusions

En aquesta anàlisi s'han pres de referència els resultats globals de descàrrega, però per fer una aproximació més acurada caldria projectar els resultats dels territoris de parla catalana. En aquesta línia, el patró de comportament és molt similar, ja que impera el criteri continental europeu: en queden exclosos els programes d'aprenentatge únic de l'anglès, tant per a infants com per a comunitats de parlants extracomunitaris, especialment la comunitat xinesa.

En general, en aquests darrers anys s'ha viscut un increment de l'oferta de cursos en català i per a catalanoparlants, però proporcionalment han anat perdent incidència en el sector, com és el cas de Memrise. Aquesta tendència, però, no ha anat acompanyada d'un augment de l'oferta en català a les empreses líders del programari a distància d'aprenentatge de llengües. En aquest sentit, Duolingo es manté estàtica en l'augment de la seva oferta i no permeabilitza la inclusió del català.

La multiplicació de l'oferta de programes a distància ha donat peu també a la creació de programes per a llengües minoritzades. Més enllà dels exemples esmentats anteriorment, el sector públic i privat ha trobat en aquestes iniciatives un mecanisme de promoció del patrimoni cultural no utilitzat fins al moment. A tall d'exemple, a l'Estat francès, el programa e-Reo s'ha creat per facilitar l'aprenentatge de llengües d'ultramar o de les antigues colònies, com ara el crioll tahitià. En aquest sentit, el català no n'és una excepció: la Generalitat de Catalunya ha creat el programa Qüestiotat per aprendre tot els nivells oficials, mitjans i superiors de català (B1, B2, C1, C2 i LA) i Plataforma per la Llengua ha creat Parauletes amb l'objectiu d'ajudar els catalanoparlants residents a l'estranger a ensenyar la llengua als seus fills des de ben xiquets.

Finalment, i com a nota més destacada per a aquest proper 2024, Google ha començat a aplicar, des del novembre d'enguany, un sistema d'aprenentatge de llengües a través del seu cercador, gràcies als models de llengua massius LaMDA i PaLM. Concretament, permet aprendre anglès a partir de la traducció del resultat i ha entrat en funcionament a l'Índia, Indonèsia, Argentina, Colòmbia, Mèxic i Veneçuela. Dit això, el projecte d'expansió

de Google s'ha iniciat al continent africà amb la generació de bases de veu de llengües en perill d'extinció. Un cop aquest sistema arribi a Europa en els propers mesos, el català hi serà present i el sector viurà una revolució important, tenint en compte la gran quota de mercat de què disposa la *big tech* de Palo Alto, Califòrnia.

N'estarem al cas i treballarem per incidir en la plena presència del català en aquestes eines tan importants en el context actual d'aprenentatge de llengües.