

Research Article

La irrupción de los traductores neuronales en sistemas de actividad lingüística reglada: Distorsiones, conflictos e innovaciones

The emergence of neural translators in regulated language activity systems: Distortions, conflicts and innovations

Eduardo Apodaka^{1*}, Asier Amezaga² y Auxkin Galarraga³

¹ Universidad del País Vasco, Facultad de Ciencias Sociales y de la Comunicación, Grupo de investigación INNOKLAB. Departamento de Psicología Social. Barrio Sarriena, 48940, Leioa, Bizkaia, España.

² Universidad del País Vasco, Facultad de Ciencias Sociales y de la Comunicación, Grupo de investigación INNOKLAB. Departamento de Sociología y Trabajo Social. Barrio Sarriena, 48940, Leioa, Bizkaia, España.

³ Universidad del País Vasco, Facultad de Derecho, Departamento de Sociología y Trabajo Social. Paseo Manuel Lizardizabal, 2, Campus de Ibaeta, 20018, Donostia, Gipuzkoa, España.

*Correspondencia: eduardo.apodaka@ehu.eus

Resumen: La inteligencia artificial basada en redes neuronales está suponiendo una revolución en la tecnologización de las actividades lingüísticas de todo tipo. El procesamiento o la comprensión del lenguaje natural y el reconocimiento automático de habla y sus aplicaciones van a modificar las relaciones y actividades lingüísticas y la comunicación en general. Presentamos en este artículo el planteamiento y primeros pasos de una investigación en curso sobre el impacto de la Inteligencia Artificial y, en particular, sobre los desafíos y oportunidades de una sus aplicaciones, los traductores neuronales en el caso de una lengua minoritaria en proceso de revitalización, el euskera. El estudio se centra en sistemas de actividad funcionales y reglados que tienen por objetivo expreso tanto las competencias y hábitos lingüísticos como las capacidades colectivas para el uso vivo de la lengua vasca. Se examina el impacto, las distorsiones e innovaciones que esta tecnología está produciendo y las que puede producir en un futuro cercano atendiendo a la tensión entre objetivos, normas y recursos.

Palabras clave: traductores neuronales; euskera; revitalización de lenguas minoritarias; sistemas de actividad reglada.

Abstract: Artificial intelligence based on neural networks is assuming a revolution in the technologization of linguistic activities of all kinds. The processing or understanding of natural language and the automatic recognition of speech and its applications will change language relations and activities and communication in general. In this article we present the approach and first steps of an ongoing research on the impact of artificial intelligence and, in particular, on the challenges and opportunities of its applications, neural translators in the case of a minority language in the process of revitalization, the Basque language. The study focuses on functional and regulated activity systems that explicitly target both language skills and habits and collective capacities for the live use of the Basque language. It examines the impact, distortions and innovations that technology is producing and those that it can produce in the near future in the light of the tension between objectives, standards and resources.

Keywords: neural translators; basque; revitalization of minority languages; regulated activity systems.

1. Introducción

En este artículo presentamos el planteamiento general y los primeros pasos de una investigación en curso sobre el impacto de la Inteligencia Artificial (en adelante, IA) y, en particular, sobre los desafíos y oportunidades de una sus aplicaciones, los traductores neuronales (en adelante, TN). Aunque el objeto general del estudio es el impacto de los TN en el uso y valor (o estatus) social de una lengua minorizada, el euskera, en el contexto de las políticas, prácticas e iniciativas para su revitalización, la investigación se centra en sistemas de actividad funcionales y reglados y examina el impacto, las distorsiones e innovaciones que esta tecnología está produciendo y las que puede producir en un futuro cercano.

Nos preguntamos qué ocurre cuando se introducen los traductores neuronales en sistemas de interacción reglados que tienen como objetivo y parte fundamental de su desarrollo tanto las aptitudes, habilidades o competencias lingüísticas en la lengua minoritaria como el aumento y la estructuración de las capacidades colectivas de uso de la lengua en proceso de revitalización.

Para ello, consideramos prioritario ser sensibles a los cambios prácticos que la interacción con estas tecnologías y su mediación en la comunicación y vida social están produciendo en las competencias, recursos y capacidades lingüísticas. Sobre esa base pretendemos construir un marco de reflexión acerca de cómo afectará a la definición, desarrollo, objetivos, medios, etc., de la revitalización del euskera, tanto en general como en sectores y actividades más concretos (sanidad, educación, etc.).

El proyecto en el que se integra este estudio prevé tres fases: una inicial en la que se realizará una aproximación al estado actual mediante entrevistas, coloquios y debates con expertos en IA que trabajan en el desarrollo de traductores y con profesionales de ámbitos de uso donde se pueden dar los conflictos mencionados (educación secundaria, universidad, sanidad, medios de comunicación, traducción, etc.). En una segunda fase se seleccionarán varios sistemas de actividad y se diseñará, junto con los actores, modelos abiertos de intervención en el uso de los traductores, con el objetivo de experimentar nuevos usos, nuevas reglas y nuevas capacidades. En la tercera fase se realizarán las intervenciones, se analizarán los resultados y se sistematizarán modelos de innovación basados en la investigación.

En este contexto, este artículo expone los resultados iniciales de la primera fase, además de dar cuenta de las reflexiones iniciales acerca del diseño metodológico y teórico. Aunque estemos en los pasos iniciales de la investigación, se ha desarrollado un ejercicio de reflexión sobre las disrupciones o rupturas, tensiones e innovaciones derivadas de los TN, sobre el aumento de capacidades que desbordan los sistemas de actividad, sus normas, objetivos y medios, y sobre las adaptaciones innovadoras de dichos sistemas.

El artículo está estructurado tal y como se detalla a continuación. En el segundo apartado explicaremos brevemente en qué consiste la IA basada en redes neuronales y aprendizaje profundo. En los siguientes apartados, abordaremos el impacto de los TN en las lenguas minoritarias o de bajos recursos, y en el proceso de revitalización del euskera. Por último, discutiremos el enfoque teórico más adecuado para la investigación.

2. Inteligencia Artificial, Redes Neuronales y Procesamiento de Lenguaje Natural

La IA es un amplio campo de investigación científica y aplicación tecnológica que aúna el estudio y la construcción de “entidades inteligentes” (Russell & Norvig, 2010, p. 1). La capacidad de procesamiento de información y de cálculo de la IA ha superado la simple rutina de seguimiento de reglas o heurísticos y el cálculo estadístico, gracias sobre todo al desarrollo de la inteligencia computacional, el aprendizaje automático (machine learning) y el aprendizaje profundo que, en síntesis, permiten el desarrollo de sistemas inteligentes complejos capaces de aprender en y de su propio desempeño. Entre las aplicaciones más relevantes de la inteligencia computacional y el aprendizaje automático encontramos, precisamente, las redes neuronales.

En cuanto al desarrollo tecnológico, la IA es uno de los ejemplos más relevantes de los que Clayton Christensen (1997) denominó innovación disruptiva. Aplicado a las tecnologías,

Christensen se refería a la capacidad de una tecnología para crear valor desde segmentos de mercado infravalorados o incluso de abrir nuevos mercados rompiendo con las previsiones y los marcos comúnmente aceptados. La IA, el aprendizaje profundo y las redes neuronales tienen ese efecto no solo desde el punto de vista económico. Los efectos sociales, psicológicos, culturales, etc., de estas innovaciones tecnológicas están aún muy lejos de ser comprendidos, pero hay pocas dudas sobre la capacidad de la IA para trastocar previsiones y tendencias sobre el cambio social y cultural, y para seguir incrementando su valor e impacto exponencialmente.

Por todo ello, el campo de la IA, más allá del desarrollo tecnológico, implica estudios sobre procesos cognitivos y conductuales, así como numerosas controversias filosóficas, éticas y científicas. Sociólogos y filósofos como Eric Sadin (2017, 2020), Geertze Lovink (2019), Bernard Stiegler (2001, 2015), Yuk Hui (2020) o Domingue Cardon (2018) entre muchos otros, han analizado desde un punto de vista crítico y pesimista los cambios que han derivado del complejo formado por las TIC, Internet y la Inteligencia Artificial. Han dado continuidad a una antigua línea de pensamiento filosófico sobre técnica y tecnología que agrupa a Heidegger, Günther Anders, o Lewis Mumford y nos habla de un cambio de gran escala, un cambio de época o civilización.

Otros han puesto el acento en los cambios que las innovaciones tecnológicas van a producir en las capacidades cognitivas y sociales (Bridle, 2020; Carr, 2019; Lash, 2005; Wolf, 2018). Y otros muchos han publicado estudios globales desde la crítica de la economía política de la IA, la infraestructura terrestre de la nube, los derechos protegidos y privados, el acceso universal, lo común de la lengua, etc. (Srnicek, 2019; Srnicek & Williams, 2017; Steinhopff, 2021).

Los cambios tienen, por supuesto, un carácter ambivalente y dan lugar a valoraciones antagónicas: se puede hacer hincapié en la parte negativa del impacto en las competencias personales o en las formas de vida actuales, o en la adaptación colectiva y personal más o menos resiliente, o en la celebración de las nuevas capacidades, potencias y comodidades tecno-sociales. Se tome la posición que se tome, hay algo que no podemos ignorar: muchas de esas innovaciones tecnológicas se han extendido con una inmensa rapidez y han superado todo tipo de “obstáculos previsibles” gracias a las virtualidades o *affordances* que los usuarios han sabido encontrar desde sus primeras interacciones con esas tecnologías –es el caso de los teléfonos móviles y de su evolución, desde un teléfono portátil al dispositivo multifuncional, híbrido de ordenador y teléfono, que es hoy en día. Esta es la perspectiva inicial con la que nos acercamos a los traductores basados en redes neuronales, dado que las redes neuronales vienen “cargadas” de *affordances*.

Las redes neuronales son sistemas o modelos de computación que desarrollan combinaciones de parámetros, de modo que pueden darnos la combinación más idónea o ajustada al resultado deseado. La unidad básica, denominada metafóricamente “neurona”, es una unidad de cálculo, una suma ponderada de los datos o parámetros de entrada a la que se aplica una función no lineal. Las redes están compuestas por varias neuronas y varias capas (que hacen aumentar la “profundidad” de la red), de manera que se hacen cálculos sobre los resultados de los cálculos anteriores, corrigiendo o ajustando esos mismos resultados. Las neuronas están constantemente colaborando entre sí y mediante el algoritmo de “propagación hacia atrás” o “retro-propagación” pueden volver sobre los cálculos y conexiones realizados en la red y reformularlos para obtener mejores resultados.

La capacidad de computación de las redes hace que mediante combinaciones de los datos iniciales sean capaces de encontrar información “oculta” y puedan “adaptarse” (es decir, usar los nuevos datos en los cálculos posteriores) y mejorar su rendimiento. Por eso mismo se habla de entrenamiento de las redes, de aprendizaje de máquina, de aprendizaje profundo (el aprendizaje automático a través de la retro-propagación en múltiples capas), etc. Es decir, las redes “aprenden” de sus “experiencias”, mejoran su rendimiento y se adaptan a nuevas demandas y situaciones.

Pero como hemos apuntado más arriba, más allá del funcionamiento técnico y de la arquitectura de las redes neuronales, está la pregunta sobre sus utilidades (dejamos a un lado su

historia social, para qué fueron diseñadas en principio y cómo han evolucionado). Las redes neuronales se están aplicando en actividades concretas como diagnóstico médico, cálculo financiero, robótica, control de sistemas automáticos, predicción de procesos, reconocimiento de patrones (de personas o de vehículos), minería de datos, etc. En general, para diagnósticos, pronósticos, o toma de decisión en situaciones complejas. Y, en cuanto a aplicaciones técnicas destacan el procesamiento de imágenes y el procesamiento de lenguaje natural (en adelante, PLN).

El procesamiento de lenguaje natural da nombre a todas las tareas implicadas en la interacción entre máquinas y seres humanos mediante el lenguaje natural. Para que las máquinas puedan usar las lenguas naturales deben ser capaces de procesar y analizar grandes cantidades de “dato lingüístico natural”. Tras dos periodos de PLN en que el procesamiento se basaba en la aplicación de normas y cálculo estadístico, hoy en día el PLN se basa en general en redes neuronales. La extensión y mejora de las redes neuronales está siendo meteórica, en el caso de la traducción, por ejemplo, la primera publicación sobre aprendizaje profundo aplicado a la traducción es de 2014 (Bahdanau, Cho & Bengio, 2014), pronto se popularizó y se hicieron mejoras en su arquitectura, en 2017 se presentó el modelo Transformer de aprendizaje profundo (Vaswani et al., 2017), un año después se presentaba el traductor en euskera MODELA.

El PLN engloba la comprensión y generación de lenguaje natural (NLU y el NLG respectivamente en su acrónimo en inglés). Áreas de este último son los sistemas de reconocimiento del habla (ASR), como los integrados en los asistentes de voz de Google o Amazon o los chatbots, y los sistemas de síntesis del habla capaces de leer un texto (TTS) y los modelos de producción de textos, como el conocido GTP-3 de OpenAI (Generative Pre-trained Transformer 3). Entre las tareas que realizan estos sistemas están la minería de textos, extracción de información, la realización de resúmenes, análisis y clasificación de documentos, creación de diccionarios, el filtrado de informaciones, análisis de sentimientos, inferencia de textos, y muchos más. Y, por supuesto, la traducción entre lenguas naturales.

Toda esta tecnología de IA y su rápido desarrollo han dado lugar a numerosas previsiones y elucubraciones acerca de lo que está por llegar. Los traductores neuronales no se limitarán al modelo actual (aplicaciones integradas o no en servicios de Internet, etc.). Las tecnologías de PLN tienden a ensamblarse en sistemas más complejos con utilidades múltiples como la traducción directa de entrevistas en video-llamada, o subtítulos automática de contenidos audiovisuales, o traducción con la misma voz del locutor en varios idiomas¹.

En ese sentido, se pueden prever varias tendencias: se extenderán los chatbots en numerosos trabajos y servicios, aumentará la comprensión emocional o empática en la evaluación y resolución de situaciones; las interfaces se irán “invisibilizando” e “inmaterializando” gracias a la nanotecnología, el diseño emocional o integrado, o la utilización del cuerpo como interfaz o joystick; los servicios y monitorizaciones se individualizarán, con un posible efecto de hiperindividualización de control de parámetros de salud o biométricos; extensión de la “conversación” entre máquinas, interconexiones en sistemas cada vez más complejos (domotización, automatización de plantas de producción, etc.)

Como tendencia general, Eric Sadin (2020, p. 88) ha señalado que “todos los objetos y todas las superficies nos hablarán” o que cada vez serán más los dispositivos tecnológicos que decidirán por las personas, anticipando incluso sus “deseos”. Los chatbots, los asistentes virtuales o todo tipo de objeto y máquinas dotados de IA mantendrán conversaciones, discusiones, nos aconsejarán, indicarán qué hacer, o simplemente nos prescribirán, con o sin nuestro conocimiento consciente, cursos de acción. Todas esas relaciones, como hemos dicho, se harán en lenguaje natural, lo cual, se pronostica, afectará a nuestras competencias cognitivas y lingüísticas. Este es

¹ Google IA presentó en 2019 Translatotron, programa que traduce reproduciendo la voz original y que se puede aplicar automáticamente a retransmisiones en vivo. Véase: <https://ai.googleblog.com/2019/05/introducing-translatotron-end-to-end.html>

un tema polémico y recurrente desde la integración de la informática y las TICs en la vida cotidiana y en la educación. Son muchas las preguntas que los investigadores se plantean respecto a los cambios que estas tecnologías acarrearán en el futuro: ¿Qué le ocurrirá a la lectura-escritura y a las habilidades cognitivas ligadas a esas competencias lingüísticas? ¿Qué ocurrirá si solo mantenemos relaciones orales con nuestro entorno tecnológico? ¿Cómo afectará a la memoria o al razonamiento el uso intenso de dispositivos y aplicaciones que realizarán las tareas cognitivas por nosotros? ¿Cuáles son, cuáles serán, las nuevas habilidades digitales? ¿Cómo y quién las capitalizará? El alcance de nuestra investigación está lejos de poder responder a estas cuestiones, pero es este tipo de especulación de donde parte la investigación que presentamos.

3. Las lenguas ante la digitalización y la IA

La IA supone nuevas formas de aprendizaje y entrenamiento de las habilidades y competencias cognitivas tenidas por básicas, formas que según algunos autores pueden debilitar esas competencias (Bridle, 2020; Carr, 2011, 2014), reforzar otras e incluso hacer surgir nuevas competencias (Reed, 2020; OECD, 2021). Un ejemplo clásico es el de lectura o la comprensión lectora, asociada a su vez al pensamiento crítico, la empatía o la reflexión. Algunos estudios señalan que en una generación se perderá la capacidad de leer un texto (Crawford, 2015), una tendencia cuyos síntomas se constatarían en el aumento de déficits de concentración o el analfabetismo funcional debidos a la pérdida de la cultura de la lectura (Wolf, 2018, 2008).

De todas formas, la mediación tecnológica o tecnologización de las lenguas no es nueva, no partimos de lenguas “a-tecnológicas”, la lengua natural es de hecho una lengua mediada por tecnologías “socio-lingüísticas” (desde útiles como diccionarios o gramáticas, a espacios e instituciones sociales como los medios de difusión y comunicación que han estandarizado lenguas o jerarquizados modos de hablar...). Así como Walter J. Ong (2002) propuso estudiar las relaciones entre oralidades pre-escritura y nuevas oralidades, creemos que este tipo de estudios deben enfocarse en las innovaciones y disrupciones en los modos de vida tecnológicos (Lash, 2005), sin presuponer estados lingüísticos pre o a-tecnológicos, sino analizando las transformaciones y derivas de la lengua en sus formas de tecnologización.

Nos encontramos, de todas formas, ante una situación de incertidumbre que afecta a las competencias lingüísticas y al valor y uso social de las lenguas particulares. Tal como ocurrió con la imprenta y la cultura del libro, el acceso a la digitalización y la IA se ha convertido en una necesidad para las comunidades lingüísticas. La presunción básica y generalizada es que no habrá vida social o cultural, ni desarrollo económico o tecnológico sin mediación digital y que el acceso a esa mediación es y será un campo de tensiones. La competencia entre lenguas en los mercados culturales audiovisuales, tecnológicos o, incluso, académicos, anuncia una fuerte jerarquización y capitalización de las lenguas, además de dependencias y conflictos multinivel (por ejemplo, acerca de la propiedad de los productos y obras realizadas mediante IA, o sobre el acceso de la ciudadanía a unas tecnologías que serán imprescindibles en muchas facetas de la vida social y personal).

La preocupación por la digitalización no es exclusiva de las lenguas minoritarias. En el caso del castellano, una de las lenguas con más hablantes del mundo, los proyectos e iniciativas inciden en esa línea. En 2019, la RAE en colaboración con Telefónica y las grandes empresa tecnológicas (Google, Amazon, Microsoft, Twitter y Facebook) puso en marcha el proyecto LEIA, para impulsar el «buen uso de la lengua española en el universo digital y, especialmente, en el ámbito de la inteligencia artificial y las tecnologías actuales» (RAE, 2019). En 2021 la Biblioteca Nacional de España y el Barcelona Supercomputing Center (BSC, 2021) presentaron MarIA, el primer sistema de IA masivo y experto capaz de comprender en lengua española. En marzo de 2022 el Gobierno español aprobó el Proyecto Estratégico para la Recuperación y Transformación Económica (PERTE) “Nueva economía de la Lengua” para aprovechar el «potencial del español y de las lenguas cooficiales como factor de crecimiento económico y competitividad internacional en áreas como la inteligencia artificial, la traducción, el aprendizaje, la divulgación cultural, la

producción audiovisual, la investigación y la ciencia». En la memoria ejecutiva se argumenta que cada vez son más las aplicaciones «donde el uso intensivo de las tecnologías de la información y de la comunicación, procesamiento del lenguaje natural y la traducción automática resultan críticas para dotar a la ciudadanía de nuevos servicios avanzados y optimizar procesos productivos» (Gobierno de España, 2020, p. 3).

La tensión entre las oportunidades y los riesgos de la IA y el PLN se agudizan en el caso de las lenguas minoritarias o de bajos recursos. A la falta de recursos ante la globalización de la producción audiovisual o de los mercados culturales en general, se le suman las dificultades para el acceso a la IA o la dependencia de grandes plataformas y empresas tecnológicas. Sin embargo, al mismo tiempo, la IA también puede ayudar en la difusión de productos culturales más allá de la comunidad de hablantes o colaborar en la revitalización lingüística en la propia comunidad.

Estos cambios tecnológicos no se han producido al margen de cambios de carácter, sociales, culturales, económicos, etc. Estos cambios han modificado sensiblemente el contexto en el que durante las últimas décadas se han desarrollado los proyectos e iniciativas de revitalización lingüística (Lewis & Royles, 2018, p. 504), que habían sido pensados en marcos sociales menos tecnologizados y menos individualistas (Lewis & McLeod, 2021). En este contexto, se ha hablado de la muerte digital de las lenguas, una lengua que no tiene cabida en el mundo digital no podrá sobrevivir, perderá toda su fuerza vital y desaparecerá (Kornai, 2013).

Aunque en 2013 András Kornai se referiría sobre todo a la presencia de las lenguas en Internet, los retos de la IA son los mismos: las lenguas que no accedan a la IA corren el peligro de perder funcionalidad, prestigio, o uso y competencia entre las nuevas generaciones. Kornai concluía que la gran mayoría de los idiomas del mundo (hasta un 95 %) ya habían perdido la capacidad de ascender al campo digital (Kornai, 2013). En 2012, el informe de META-NET, Alianza Tecnológica Multilingüe de Europa, preveía que el euskera, junto con otras 21 lenguas europeas, estaban en peligro de extinción digital (Rehm & Uszkoreit, 2012). Sin embargo, desde entonces la IA ha provisto de herramientas que podrían servir para evitar la desaparición de las lenguas ante la presión de la digitalización², al tiempo que se han ido elaborando nuevos indicadores lingüísticos en Internet (Bromham et al., 2021; Pimienta, 2021; Prys et al., 2021).

4. Situación sociolingüística y nivel de digitalización del euskera

La población de habla vasca se extiende por un territorio bajo dos administraciones estatales y varias administraciones sub-estatales con diferentes situaciones, leyes y políticas lingüísticas (Comunidad Autónoma Vasca, Comunidad Autónoma de Navarra, en territorio español y la Comunidad de aglomeración del País Vasco en territorio francés). Según la VI Encuesta Sociolingüística (Gobierno Vasco, 2016) unas 750 000 personas saben euskera, 28,4 % de la población mayor de 16 años, pero solo un 16,5% lo usa intensivamente. En los últimos 25 años el uso ha aumentado 3 puntos, pero en los diez últimos no ha variado. La UNESCO clasifica la lengua vasca como vulnerable en su Atlas de lenguas en peligro (Moseley & Alexandre, 2010)³.

El Gobierno Vasco ofrece una serie de indicadores para hacer un seguimiento a la vitalidad digital del euskera y su uso, así como del desarrollo de la IA en euskera: uso en redes sociales, webs de referencia en euskera, oferta lingüística en aplicaciones móviles, servicios web, herramientas básicas y sistemas operativos, usuarios de Wikipedia en euskera, etc., (Gobierno Vasco, s.f.). Según esos indicadores, en el año 2020 el 14% de los usuarios –especialmente los jóvenes– utilizó el euskera en redes sociales, solo un 5 % de los mensajes de Twitter fueron en

² Véase, por ejemplo, el Kit de herramientas para idiomas minorizados y de bajos recursos de Col·lectivaT y el Centro Sefardí de Estambul, en <https://language-toolkit.readthedocs.io/es/latest/index.html>. Para una síntesis de la situación digital de las lenguas minoritarias y regionales de Europa, veáse Miriam Gerken (2020), *Facilitating the implementation of the European Charter for Regional or Minority Languages through artificial intelligence*, disponible en: <https://rm.coe.int/min-lang-2022-4-ai-and-ecrml-en/1680a657c5>

³ Esa clasificación se debe en parte a la inmersión lingüística en la escuela, ya que vulnerables son aquellas lenguas que son habladas por la mayoría de las generaciones más jóvenes, pero restringidas a ciertas áreas (e.g. la familia o la escuela).

euskera (frente a un 80 % en castellano, un 11 % en francés o un 4% en inglés), Wikipedia cuenta con unos 400.000 artículos en euskera (por número de artículos el euskera ocupa el puesto 34 entre las lenguas de Wikipedia), por último, de los servicios Web, plataformas, APPs, sistemas operativos y herramientas básicas más usadas en Europa la mitad está disponible en euskera.

El volumen de mercado que nos indican los datos anteriores ha hecho que la inversión y los esfuerzos por avanzar en la digitalización del euskera, liderados por las instituciones públicas y el movimiento social y se enmarquen en las políticas e iniciativas por la revitalización del euskera. El desarrollo de la digitalización del euskera ha sido liderado por el dinamismo y la implicación de varios agentes, que han sido los que, entre otras muchas aportaciones, han desarrollado los traductores neuronales. En una breve exposición de la situación del euskera ante la IA y del desarrollo de las tecnologías lingüísticas en euskera en los últimos veinte años, I. Leturia y J. Aztiria analizan varios proyectos⁴, la mayoría financiados por el Gobierno Vasco a través de convocatorias específicas (Eortek, Elkartek, etc.). En todo ese camino, dicen, han colaborado Vicomtech, Tecnalia, los grupos de investigación Ixa y Aholab de la Universidad del País Vasco (UPV/EHU) y la Fundación Elhuyar (Leturia & Aztiria, 2021, p. 84)⁵. Estos son, pues, los principales traductores en el área de Tecnologías de la Lengua del País Vasco:

- Ixa Taldea, grupo de investigación de la UPV/EHU creado en 1988, trabaja en PLN aplicado a la revitalización y normalización del euskera y otras lenguas de bajos recursos. Ha desarrollado numerosos procesadores, recursos y aplicaciones (el corrector ortográfico Xuxen, el traductor automático Matxin, la red lexical Euskal WordNet, etc.)⁶.
- Aholab, grupo de investigación de la UPV/EHU creado en 1995. Trabaja en procesamiento de habla (reconocimiento, conversión a texto, etc.)
- Hitz Zentroa, centro de investigación de la UPV/EHU, fundado en 2019 por Ixa Taldea y Aholab, referente europeo en investigación básica en PLN.
- Vicomtech es un Centro Tecnológico especializados en tecnologías digitales. Es el grupo que, además de numerosos planes y proyectos institucionales, ha desarrollado el traductor neuronal del Gobierno Vasco, Itzuli.
- Elhuyar es desde 2002 una Fundación dedicada fundamentalmente a tecnologías de la lengua, pero tiene sus raíces en el grupo Elhuyar creado en 1972 por alumnado voluntario de la Escuela de Ingeniería de la Universidad de Navarra, en San Sebastián, con el objetivo de desarrollar herramientas para el uso del euskera en ciencia y tecnología. Hoy en día, además del trabajo en ciencia y euskara, ofrece servicios para la aplicación del conocimiento avanzado. Elhuyar es famosa por su el diccionario Elhuyar de 1996 digitalizado en 1999. En tecnologías del habla y traducción ha desarrollado, entre otros proyectos, un sistema de reconocimiento de habla, Aditu, capaz de hacer transcripciones y subtítulos automáticamente, está en trabajando en un asistente inteligente (altavoz inteligente), Mycroft.eus, capaz de interactuar en euskera y el software libre o traductor neuronal Elia. En 2022, Elhuyar ha convertido la unidad de I+D que se encargaba de los servicios lingüísticos en Orai, centro de investigación, desarrollo e innovación especializado en PLN y dirigido a la industria.

⁴ Los proyectos analizados son Hizking 21, Anhitz, Berbatek, Be2Tek, ElkarOla, BerbaOla, Modela, Quales, Módona, Tando, DLNP4 DeepText, que se han desarrollado en colaboración con y con la colaboración con Vicomtech y Tecnalia, los grupos de investigación Ixa y Aholab de la Universidad del País Vasco (UPV/EHU) y la Fundación Elhuyar.

⁵ Recientemente, en julio de 2021, se ha constituido el Centro Vasco de Inteligencia Artificial, BAIC (Basque Artificial Intelligence Center) fundamentalmente un clúster de empresas para la promoción de la investigación aplicada en IA, <https://www.baic.eus/>. Y en abril de 2022 el Gobierno Vasco ha aprobado el Plan de Acción de las Tecnologías de la Lengua 2021-2024, Gaitu, compuesto de proyectos concretos de la Administración.

⁶ Se puede consultar la producción del grupo en <https://www.ixaeus/produktuak?language=eu>

5. Los traductores neuronales en euskera

El desarrollo de la traducción neuronal en euskera se inició en 2015 con un proyecto denominado TAdeep del grupo Ixa Taldea. Anteriormente Ixa Taldea y Elhuyar habían creado Matxin un traductor automático entre euskera y castellano, con las tecnologías disponibles en la época. En 2017, se puso en marcha el proyecto MODELA en el que participaban varios agentes (Ixa Taldea, Elhuyar, Vicomtech, Ametzagaña e Isea Coop. de la Corporación Mondragon) con el fin de aplicar el aprendizaje profundo en un traductor de euskera. En un año, en 2018, se presentó el traductor MODELA, basado en deep learning, que fue el primer servicio de traducción neuronal en euskera abierto al público general (Cortés, 2017; Cortés & Jauregi, 2019). Desde entonces se han puesto a disposición del público cuatro traductores neuronales desarrollados en el País Vasco, con algunas diferencias técnicas, y basados en distintos corpus lingüísticos: Itzuli, Batua y Elia.

- Batua, desarrollado por Vicomtech, 2019, traduce entre cuatro idiomas (euskara, castellano, inglés, francés), basado en tecnologías deep learning y transformer.
- Itzuli, traductor del Gobierno Vasco, también de Vicomtech. Se ha usado como base el corpus digitalizado de la administración vasca. Como el anterior traduce entre los cuatro idiomas mencionados y está basado en mejoras del sistema MODELA.
- Elia, traductor automático multilingüe de Elhuyar, 2019. Trabaja en cinco idiomas (euskera, castellano, francés, inglés y catalán).

En cuanto a los datos de uso, y a modo de ejemplo, en el año 2021 el traductor ELIA tuvo 588.983 usuarios, 2.022.069 visitas, se instaló la aplicación en 18.809 dispositivos móviles, se hicieron 15.188.937 solicitudes de traducción y fueron más de 387 millones las palabras traducidas: un 51,3% de castellano a euskera, un 19,4% de euskera a castellano y un 29,3% entre otras combinaciones (el traductor ofrece traducciones en cuatro idiomas: euskera, castellano, francés e inglés).

Pero además de los traductores desarrollados en el País Vasco, las grandes empresas tecnológicas como Google o Microsoft han ido incluyendo en sus traductores al euskera. Ello ha sido posible gracias a traducciones cruzadas, con otro idioma como mediador, y ha hecho posible extender los traductores tanto en las redes, en Internet, como en aplicaciones y dispositivos que traducen directamente conversaciones y textos.

Como hemos dicho, en los últimos años a medida que los traductores se han ido popularizando han provocado tensiones en muchos sectores. El “don de lenguas” que ofrecen rompe o distorsiona las normas, objetivos, rutinas y estructuras de actividades configuradas para la revitalización de la lengua vasca, actividades y espacios concebidos para vivir en euskera o para su difusión. Durante el último año hemos recogido informaciones y opiniones sobre dichas tensiones en varios sectores de actividad⁷: medios sociales de comunicación en euskera, educación y universidad, traducción profesional o servicios de salud⁸. Exponemos brevemente la información más relevante por sectores.

- Durante las últimas décadas han surgido numerosos medios sociales de comunicación en euskera, con el fin de dotar a la comunidad de hablantes de medios locales, cercanos y en euskera. Muchos de estos medios se encuentran con que sus páginas web o sus redes sociales son traducidas por los usuarios o llegan a los usuarios traducidas, no siempre con la calidad deseable. No es una opción que hayan hecho los propios medios, escapa a su control ya que los traductores están disponibles en aplicaciones e integrados en muchas páginas web o redes sociales (el ejemplo de Twitter es bien conocido), de modo

⁷ En entrevistas con informantes directos, en coloquios y mesas redondas sobre el tema donde hemos participado como ponentes u oyentes, en artículos de prensa, en conversaciones informales, etc.

⁸ Hay otros ámbitos en los que hay un compromiso efectivo o meramente normativo con la revitalización del euskera y en los que se está introduciendo la mediación de la traducción automática: la enseñanza para adultos del euskera, administración de empresas, planes de normalización en el mundo laboral, relaciones con la administración, servicios públicos, etc.

que podemos acceder a páginas que se traducen según las consultamos a la lengua que hayamos elegido o en la que tengamos configurado el dispositivo. Dado que el objetivo fundacional de esos medios era la normalización de la comunicación social en euskera, ¿qué deben hacer? ¿Pone la traducción automática en cuestión, el carácter y objetivos de estos medios? ¿Deben reinventarse para mantener su objetivo?⁹ Josu Aztiria (2022) propone varias hipótesis de trabajo que pueden darnos una idea de las alternativas y de las decisiones que tendrán que ir tomando los medios. Pueden optar por ofrecer todo el contenido del medio en varios idiomas mediante la traducción automática, pueden ofrecer traducciones de contenido de forma asimétrica, puede ofertar productos diferenciados y adaptados en otros idiomas, o pueden incluir subescritura y doblaje masivo de los contenidos audiovisuales en euskera. También pueden optar por generar proyectos comunicativos conjuntos con redes internacionales en la que cada medio comunicaría en su idioma. En definitiva, se trataría de impulsar medios que trabajan originalmente en euskera promoviendo la comunicación en euskera dentro de la comunidad de lengua, pero aprovechando la traducción automática para ampliar su público, sin perder el control sobre el estilo o la calidad de los contenidos traducidos¹⁰.

- En el ámbito educativo, el profesorado de educación media y universitario se ha percatado que es cada vez más habitual que el alumnado redacte los trabajos en castellano y use los traductores automáticos para entregar los trabajos en euskera. Esta tendencia hace temer que pronto, estos trabajos no sólo se traducirán, sino que también podrán ser “confeccionados” en procesadores de texto capaces de generar artículos, informes, etc. El problema es doble. Por una parte, una de las grandes apuestas para la normalización del euskera ha sido su integración en todo el sistema escolar mediante la inmersión lingüística, de modo que la actividad docente también tiene como objetivo la capacitación del alumnado en el uso competente de la lengua. ¿Se deben aceptar o dar por buenos trabajos traducidos automáticamente? ¿En qué casos? ¿Es posible integrar en la formación o en el currículo escolar los traductores y las demás aplicaciones de la IA? Ya que la prohibición de su uso no parece una opción viable, ¿cómo superar la tensión actual entre objetivos y medios? Por otra parte, vista la tendencia al uso de traductores en todos los ámbitos de la vida no parece exagerado afirmar que todos seremos “traductores” o usuarios de dispositivos que realizarán las traducciones directamente. En esos casos, ¿qué sucede con los sesgos ideológicos, o de género que estos traductores introducen? Tanto en un caso como en el otro, el conocimiento crítico de la IA y sus aplicaciones tendrá que ser incluido en el currículo educativo como competencias y habilidades socio-cognitivas y relacionales con las que trabajar (Moorkens, 2018; Sandoval, 2018).
- La traducción automática puede ayudar a mejorar la comunicación entre pacientes y servicios de salud, así como a implantar un servicio que pueda funcionar en varias lenguas independientemente de las competencias lingüísticas de los trabajadores o usuarios del servicio. No hace falta incidir en la importancia de una comunicación clara, sensible o empática en el ámbito de la salud. Los traductores automáticos parecen prometer una solución económica y directa en un mundo en el que cada vez son más las interacciones entre multilingües. Aun así, lo que parece una mediación menos intrusiva

⁹ Son algunas de las cuestiones y opiniones que se plantearon en un coloquio organizado por Hekimen, asociación de medios sociales en euskera, sobre los riesgos y oportunidad de la traducción automática, “Itzulpen Automatikoa: aukerak eta arriskuak” (Andoain, 24 de febrero de 2022). Participaron en el coloquio unos veinte medios (Hekimen lo forman 50 medios de comunicación).

¹⁰ Por ejemplo, el contenido, originalmente en euskera, de las páginas web de las revistas de Elhuyar zientzia.eus y aldizkaria.elhuyar.eus se puede leer en seis idiomas (euskera, castellano, francés, inglés, catalán y gallego). El usuario puede pedir el texto en uno de esos idiomas y el traductor realiza la traducción en el momento. Véase en <https://zientzia.eus/> y <https://aldizkaria.elhuyar.eus/>

—no mediada por una tercera persona— también puede producir un efecto de despersonalización y de estandarización, una comunicación “fría” y una delegación de competencia o una dependencia de la capacidad de la máquina. Actualmente el grupo Ixa Taldea está poniendo en marcha en el Servicio Vasco de Salud (Osakidetza, CAV) un modelo adaptado de traducción automática. Para ello han desarrollado la herramienta Itzulbide, que permite la traducción de texto libre de los informes e historias clínicas (Perez de Viñaspre, 2020). En teoría, con esa herramienta cualquier sanitario podrá escribir y leer los informes en el idioma que desee. El problema, una vez más, es la estandarización de los textos y la competencia real en los idiomas oficiales: el 50% de la plantilla de Osakidetza es bilingüe, pero de esta manera, indican las voces críticas, el peso de la normalización lingüística recae en las elecciones personales de los trabajadores y no se actúa sobre el sistema de relaciones ni se oferta claramente una atención en euskera.

- La traducción es, por supuesto, una de las actividades profesionales más afectadas por los traductores automáticos (Álvarez, 2018). Incluso hay quien considera que acabará con la traducción como actividad profesional. De momento, son quienes se dedican a la traducción quienes están adaptando su tarea a las nuevas herramientas: cursos, talleres y guías de post edición (Hu & Cadwell, 2016; Forcada, 2017; Sánchez & Ric, 2020; Iñaurrieta & Aranberri, 2021), estudios sobre las competencias necesarias para integrar la traducción automática en su labor (Mellinger, 2017; Pym & Torres, 2019), reflexión o investigación sobre el impacto de los traductores en ámbitos concretos como la administración (Arrasate, 2021), Wikipedia (Alegria et al., 2013) o sanidad (Soto et al. 2019), o sobre las diferencias entre la post edición profesional y no profesional (Aranberri et al., 2014; Schumacher, 2020).

6. Orientando la investigación

Los comentarios anteriores están en buena parte realizados desde un marco teórico en construcción que consta de varias capas: primero, tiene en cuenta los modelos o sistemas socio-técnicos segundo, las tensiones entre las competencias y disposiciones de los sujetos, las normas del sistema de actividad y las capacidades de los dispositivos, de las máquinas. Como hemos dicho, se trata de estudiar las innovaciones en cualquier aspecto del sistema de actividad en conjunto o redes de interacciones que han sido explícitamente organizados con atención a unos fines, el aprendizaje o el uso del euskera, y unas normas que regulan los medios para eso fines.

La principal característica de la IA es posiblemente su capacidad de aprendizaje. Los sistemas o redes de IA mejoran su capacitación y, como en el caso de los traductores, se “entrenan” o son entrenados en habilidades concretas y específicas, es decir, para unas tareas y para unos objetivos que o bien han sido explicitados o bien han surgido en las interacciones con usuarios o/y entre máquinas. En este último caso, un modelo socio-técnico bien conocido como la Teoría del Actor-Red (ANT) de Bruno Latour podría guiar la investigación (Latour, 2005). En ese caso nos preguntaríamos por las redes y agenciamientos o ensamblajes de elementos heterogéneos, por las interacciones múltiples, por su capacidad de generar agencia, por los cambios y “traducciones” de y en la red, etc. Pero, como hemos indicado, nos encontramos más bien en el primero de los casos: los sistemas de actividad a los que nos referimos han sido organizados intencionalmente, con unos objetivos y reglas explícitos. Son producto de la intención, la decisión, la planificación y la disponibilidad de recursos.

Nos movemos, por tanto, en la estela de la Teoría de la Actividad de Vigostky y Leontiev, siguiendo de cerca a la Teoría de Sistemas de Actividad de Yrjö Engeström (Engeström, Miettinen & Punamäki, 1999; Engeström, 2005; Daniels et al., 2010; Sannino et al., 2009). Aunque, se hará un uso crítico del modelo desde las aportaciones de la ANT.

En los modelos iniciales de Vigostky y Leontiev la actividad se entendía como un sistema de sujeto, objeto y mediador, un conjunto de acciones interrelacionadas para conseguir un objetivo. El mediador era el instrumento o recurso (el artefacto). Engeström introdujo en el modelo otros

componentes y, principalmente, las interacciones y retro-acoplamientos entre elementos, pero manteniendo la importancia de los motivos y de la conciencia. Las personas no son reducidas a nodos o agentes determinados por las interrelaciones del sistema (Nardi, 1996). Las actuaciones de las máquinas y de las personas no son del mismo orden. Se trata de una cuestión que se distancia de la ontología plana de la ANT, pero no de la idea de que todos los componentes pueden producir efectos no-intencionales e imprevisibles.

Según el modelo de Engeström, el sistema de actividad estaría formado por seis elementos, sus relaciones y sus interacciones: un par sujeto-objeto mediado por el instrumento o artefacto y la comunidad, una serie de normas y una división del trabajo. La interacción entre los elementos produce el resultado intencional por el que se ha puesto en marcha la actividad. Cada uno de los elementos debe ser especificado: el objeto es el sentido y la intención, es decir, el objetivo; el instrumento es cualquier mediación (desde el lenguaje a las máquinas e incluso otros sistemas socio-técnicos); la comunidad es el contexto social con toda su complejidad; la división del trabajo o las normas se refieren a las formas de organización social del sistema de actividad. El sujeto se asimila al individuo, pero debe recordarse que tal como postulaba Vigostky, el individuo "psicosocial" es el resultado de la incorporación y subjetivación de lo social en un cuerpo con necesidades.

Todo eso nos da una foto fija del sistema, un modelo muerto. Son precisamente la orientación intencional y las interacciones las que dirigen la mirada a las dinámicas del sistema en términos de éxito o fracaso, de tensiones y de adaptaciones y mutua transformación de los elementos. En este sentido, del mismo modo que se puede producir un ensamblaje o una adaptación "armoniosa" entre los componentes del sistema, al introducir un nuevo componente o al transformar uno anterior también se pueden dar conflictos, tensiones, fracturas, disrupciones, resistencias y cambios reactivos que reformulen los objetivos, las normas, las competencias, etc. De ahí puede derivarse la reestructuración de todos los componentes y la reorganización de todo el sistema.

Las innovaciones y creaciones se derivan, por tanto, de las relaciones cambiantes internas y externas del sistema. Usando un vocabulario propio, podemos organizar dichas dinámicas en dos ejes: el eje de las disposiciones y el de los dispositivos. En el primer eje encontramos, las interacciones entre las normas-comunidad-división del trabajo, que son los elementos que en el modelo aparecen en la base (de un diagrama triangular) y el instrumento (en el vértice). Hay, por tanto, un sujeto, un elemento activo e intencional que es también un producto o resultado y mediación de la base y el vértice. Es decir, el sujeto es la incorporación de normas, habilidades, estatus, responsabilidades, etc., en una actividad con instrumentos. El segundo eje nos lleva al otro lado del diagrama: en este caso las interacciones entre la base (normas, comunidad, división del trabajo) y el vértice (el instrumento) se realizan en el "objetivo" y se objetivan en el resultado y/o en los dispositivos (que exteriorizan los conocimientos y normas sociales). De este modo, nos acercamos al sistema de actividad como praxis: al actuar intencionalmente, el sujeto no solo modifica un estado del mundo, se modifica a sí mismo y, a su vez, se modifica el sistema de actividad entero. El sujeto es subjetivación de competencias, habilidades y/o disposiciones, y el sistema entero es una objetivación en dispositivos, resultados o productos de relaciones y conocimientos sociales.

Esto nos permite investigar cómo en diferentes sistemas reglados los traductores producen transformaciones disfuncionales y no normativas, cómo se reaprende frente a ellas, y cómo se reorganizan los sistemas de forma proactiva y reactiva integrando el nuevo elemento y desarrollando nuevas competencias, normas u objetivos para trabajar con el elemento disruptivo.

¿Qué sucede cuando el "artefacto" es un traductor neuronal mediador de relaciones entre el sujeto (normativas-fines/comunidad) y el objeto (comunidad-división funcional de trabajo)? ¿Cambiarán las normas? ¿Los sujetos, es decir, las disposiciones subjetivas? ¿El objeto, los productos, los servicios? ¿La comunidad? ¿La organización?

Los dos ejes anteriores se corresponden con dos tipos de conflicto entrecruzados: el conflicto entre, por un lado, los nuevos hábitos (en principio nuevas acciones o usos lingüísticos, de escritura, etc.) y, por otro, las normas funcionales y los objetivos de los sistemas de actividad en cuestión; y el conflicto entre las competencias psicosociales (disposiciones) y los recursos técnicos y estructurales (dispositivos).

Ambos conflictos están interconectados, ensamblados. En nuestro caso, las competencias psicosociales entran en conflicto con los recursos técnicos (con el uso de los traductores) solo porque no se ajustan a las normas y objetivos de los sistemas de actividad que hemos mencionado. Los usuarios, están construyendo, nuevas capacidades socio-técnicas, sistemas socio-técnicos de actividades con nuevas capacidades (en la acepción de capability). Pero hay que tener en cuenta que en el caso de la IA "usuario" es tanto el sujeto como el instrumento: los traductores automáticos y los usuarios humanos se están entrenando mutuamente.

7. Conclusión

Los sistemas de actividad en los que pone su atención esta investigación son un terreno cambiante y fluctuante en el que podemos aventurar que se redefinirán objetivos y normas por la irrupción de un instrumento que demanda nuevas habilidades y nuevas competencias. Ello dará lugar a nuevas capacidades (en este caso, lingüísticas) que afectarán a la vida social de las lenguas, a las relaciones entre ellas, o a la revitalización de lenguas minoritarias. Es de prever que esos cambios serán fruto de conflictos cognitivos, normativos, sociales, etc. Estos conflictos son en sí procesos de transformación, pero no procesos lineales de desarrollo u optimización de los sistemas. La entrada de un nuevo recurso como el traductor neuronal podría producir tres escenarios. En un primer escenario, al entrar en un sistema de actividad previamente organizado y regulado, es de prever que distorsionará el propio sistema y lo transformará hasta reorganizarse un nuevo "conjunto socio-técnico" dotado de nuevas capacidades. La investigación futura tendrá que aclarar cómo se producen esa reorganización. En segundo lugar, se encontrarán resistencias e intentos de hacer perdurar el sistema tal como estaba instituido. Dado que las resistencias no son también nichos de actividad innovadora, la investigación deberá observar qué aportan esas resistencias al desarrollo de los traductores y a la innovación en el uso de lengua. Por último, surgirán propuestas claramente intencionales e innovadoras tendentes a generar nuevas competencias y nuevas capacidades. Al igual que en los anteriores, en este tercer se abrirá una línea de investigación sobre el valor ambivalente de las reacciones ante los traductores y de los cambios multidimensionales que producirán.

Los tres escenarios mencionados dan prioridad a un tipo de desarrollo y reacción y a las consiguientes líneas de investigación. En el primer caso, priman los dispositivos (la IA y los traductores) y su deriva más sistémica, en el segundo, la resistencia reactiva de la institución, de sus normas y discursos; en el tercero, la innovación reflexiva desde el sentido de la actividad: la revitalización del euskera, que hasta ahora pasaba por las disposiciones y competencia lingüística de las personas y de las capacidades de los conjuntos sociales o técnicos.

La investigación pretende, por medio de la experimentación, la especulación y la imaginación, colaborar en el último escenario. Pero es previsible que en la mayoría de los casos se muevan conflictivamente entre los diferentes escenarios. De esta manera, tratará de investigar cómo son y serán los nuevos ensamblajes o agenciamientos técnico-lingüísticos. ¿Qué capacidades socio-técnicas, qué ensamblajes disposición-dispositivo, surgirán en y para la revitalización social/la vida social del euskera y de las lenguas con las que vivirá?

Referencias

- Alegria, I., Cabezón, U., Fernández de Betoño, U., Labaka, G., Mayor, A., Sarasola, K., & Zubiaga (2013). Reciprocal Enrichment between Basque Wikipedia and Machine Translation. En I. Gurevych & J. Kim. (Eds), *The People's Web Meets NLP. Theory and Applications of Natural Language Processing* (pp. 101-118). Springer. https://doi.org/10.1007/978-3-642-35085-6_4
- Álvarez, M. (2018). Los desafíos digitales en el mercado de la traducción. *Caracteres. Estudios culturales y críticos de la esfera digital*, 7(2), 193-215.
- Aranberri, N., Labaka, G., Diaz de Ilarraza, A., & Sarasola, K. (2014). Comparison of post-editing productivity between professional translators and lay users. En S. O'Brien, M. Simard & L. Specia (Eds.), *Proceedings of the Third Workshop on Post-Editing Technology and Practice (WPTP - 3)* (pp. 20–33). Association for Machine Translation in the Americas. <https://aclanthology.org/2014.amta-wptp.2>
- Arrasate, M. (2021). Itzultzaile automatiko neuronalak EAEko Administrazioan, *Senez*, 52, 189-212.
- Aztiria, J. (2022). Itzulpen automatikoa: euskara hutsezko hedabideetatik euskara ardatz duten proiektu komunikatibo eleaniztunetara. *Euskal Hedabideen Urtekaria 2021, Behategia*, 114-132.
- Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural Machine Translation by Jointly Learning to Align and Translate (v1). *arXiv*, 1409.0473. <https://doi.org/10.48550/arXiv.1409.0473>
- Barcelona Supercomputing Center. (2021). *El primer sistema masivo de Inteligencia Artificial de la lengua española, MarIA, empieza a resumir y generar textos*. <https://www.bsc.es/es/noticias/noticias-del-bsc/el-primer-sistema-masivo-de-inteligencia-artificial-de-la-lengua-esp%C3%B1ola-maria-empieza-resumir-y>
- Bridle, J. (2020). *La nueva edad oscura*. Debate.
- Bromham, L., Dinnage, R., Skirgård, H., Ritchie, A., Cardillo, M., Meakins, F., Greenhill, S., & Hua, X. (2022). Global predictors of language endangerment and the future of linguistic diversity. *Nature Ecology & Evolution*, 6, 163-173. <https://doi.org/10.1038/s41559-021-01604-y>
- Cardon, D. (2018). *Con qué sueñan los algoritmos*. Dado.
- Carr, N. (2011). *¿Qué está haciendo Internet con nuestras mentes?* Taurus.
- Carr, N. (2014). *Atrapados. Cómo las máquinas se apoderan de nuestras vidas*. Taurus.
- Carr, N. (2019). *La pesadilla tecnológica*. El Salmón.
- Christensen Clayton, M. (1997). *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Harvard Business Review Press.
- Cortés, I. (2017). Begirada bat itzulpen automatikoko sistemen bilakaerari. *Elhuyar: zientzia eta teknika*, 327, 40-41.
- Cortés, I., & Jauregi, S. (2019). Itzultzaile automatikoa: mehatxua ala aukera? *Senez*, 50, 225-230.
- Crawford, M. (2015). *The World beyond Your Head: On Becoming an Individual in an Age of Distraction*. Farrar, Straus and Giroux.
- Daniels, H., Edwards, A., Engeström, Y., Gallagher, T., & Ludvigsen, S. (2010). *Activity Theory in Practice. Promoting learning across boundaries and agencies*. Routledge.
- Engeström, Y. (2005). *Developmental Work Research: Expanding Activity Theory in Practice*. Lehmanns Media.

Engeström, Y., Miettinen, R., & Punamäki, R.-L. (1999). *Perspectives on Activity Theory*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511812774>

Forcada, M. L. (2017). Making sense of neural machine translation. *Translation spaces*, 6(2), 291-309. <https://doi.org/10.1075/ts.6.2.06for>

Gerken, M. (2020). *Facilitating the implementation of the European Charter for Regional or Minority Languages through artificial intelligence*. Secretariat of the European Charter for Regional or Minority Languages-Council of Europe. <https://rm.coe.int/min-lang-2022-4-ai-and-ecrml-en/1680a657c5>

Gobierno de España. (2020). *PERTE Nueva economía de la lengua. Memoria ejecutiva*. https://planderecuperacion.gob.es/sites/default/files/2022-03/PERTE_Nueva_economia_de_la_lengua_memoria_16032022.pdf

Gobierno Vasco. (2016). Sistema de Indicadores del Euskera. Departamento de Cultura y Política Lingüística. <https://www.euskadi.eus/web01-apeusadi/es/eusadierazle/listaV1.apl?idioma=c&tema=16>

Gobierno Vasco. (2016). *VI Encuesta Sociolingüística*. Gobierno Vasco.

Hekimen, Asociación de Medios Sociales en Euskera. (2022). *Itzulpen Automatikoa: aukerka eta arriskuak* [Video]. <https://hekimen.eus/2022/03/23/itzultzaile-automatikoari-buruzko-hausnarketa-saioaren-laburpena/>

Hu, K., & Cadwell, P. (2016). A Comparative Study of Post-editing Guidelines. *Proceedings of the 19th Annual Conference of the European Association for Machine Translation*, 346–353. <https://aclanthology.org/W16-3420>

Hui, Y. (2020). *Fragmentar el Futuro. Ensayos sobre Tecnodiversidad*. Caja Negra.

Iñaurrieta, U., & Aranberri, N. (2021). Postedizio Lantegia: ikastaroaren gakoak, *Senez*, 52, 165-187.

Kornai, A. (2013). Digital Language Death. *PLoS ONE*, 8(10): e77056. <https://doi.org/10.1371/journal.pone.0077056>

Lash, S. (2005). *Crítica de la información*. Amorrortu.

Latour, B. (2005). *Reassembling the Social. An Introduction to Actor-Network-Theory reassembling the social* oxford. Oxford University Press.

Leturia, I., & Aztiria, J. (2021). Adimen artifiziala eta euskararen erabilera: aukerak eta eragin-eremu berriak. *Hermes*, 69, 84-88.

Lewis, H., & McLeod, W. (2021). Governance, Policy-Making and Language Revitalisation. En H. Lewis & W. McLeod (Eds.), *Language Revitalisation and Social Transformation* (pp. 343-362). Palgrave Macmillan.

Lewis, H., & Royles, E. (2018). Language revitalisation and social transformation: Evaluating the language policy frameworks of sub-state governments in Wales and Scotland. *Policy & Politics* 46(3), 503-529. <https://doi.org/10.1332/030557317X14938075758958>

Lovink, G. (2019). *Tristes por diseño: las redes sociales como ideología*. Consonni.

Mellinger, C. D. (2017). Translator and Machine Translation: Knowledge and Skills Gaps in Translator Pedagogy. *The Interpreter and Translator Trainer*, 11(4), 280-293. <https://doi.org/10.1080/1750399X.2017.1359760>

Moorkens, J. (2018). What to Expect from Neural Machine Translation: A Practical In-class Translation Evaluation Exercise. *The Interpreter and Translator Trainer*, 12(4), 375-387. <https://doi.org/10.1080/1750399X.2018.1501639>

- Moseley, Ch., & Alexandre, N. (2010). *The UNESCO Interactive Atlas of the World's Languages in Danger*. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000189453>
- Nardi, B. A. (1996). Activity theory and human computer interaction. En B. A. Nardi (Ed.), *Context and Consciousness: Activity Theory and Human-Computer Interaction* (pp. 1-8). MIT Press.
- OECD. (2021). *AI and the Future of Skills, Volume 1: Capabilities and Assessments*. OECD Publishing. <https://doi.org/10.1787/5ee71f34-en>
- Ong, W. J. (2002). *Orality and Literacy: The Technologizing of the Word*. Routledge.
- Perez de Viñaspre, O. (2020). Itzulbide: osasun-txostenen itzulpen automatiko neuronal, *Osagaiz*, 4(1), 33-38.
- Pimienta, D. (2021). *Versión nueva y mejorada de un enfoque alternativo para la producción de indicadores lingüísticos en la Internet*. Observatorio de la diversidad lingüística y cultural en la Internet. <http://funredes.org/lc2021/ALI%20V2-ES.pdf>
- Prys, D., Jones, D., Prys, G., Watkins, G., Cooper, S., Roberts, J. C., Butcher, P., Farhat, L., Teahan, W., & Prys, M. (2021). *Language and Technology in Wales: Volume I*. Prifysgol Bangor University.
- Pym, A., & Torres, E. (2019) Efectos de la automatización en las competencias básicas del traductor: la traducción automática neuronal. En A. Vidal & A. Alarcón (Eds.), *Ocupaciones y lenguaje: Indicadores y análisis de competencias lingüísticas en el ámbito laboral* (pp. 475-506). Universitat Rovira i Virgili.
- Real Academia de la Lengua. (2019). *La RAE presenta el proyecto Lengua Española e Inteligencia Artificial (LEIA) en el XVI Congreso de la ASALE*. <https://www.rae.es/noticia/la-rae-presenta-el-proyecto-lengua-espanola-e-inteligencia-artificial-leia-en-el-xvi>
- Reed, S. K. (2020). *Cognitive Skills You Need for the 21st Century*. Oxford University Press. <https://doi.org/10.1093/oso/9780197529003.001.0001>
- Rehm, G., & Uszkoreit, H. (Eds.). (2012). *The Basque Language in the Digital Age*. Springer. <https://doi.org/10.1007/978-3-642-30796-6>.
- Russell, S., & Norvig, P. (2010). *Artificial Intelligence: A Modern Approach*. Prentice Hall-Pearson Education.
- Sadin, E. (2017). *La humanidad aumentada. La administración digital del mundo*. Editorial Caja Negra.
- Sadin, E. (2020). *La inteligencia artificial o el desafío del siglo. Anatomía de un antihumanismo radical*. Editorial Caja Negra.
- Sandoval, E. (2018). Aprendizaje e inteligencia artificial en la era digital: implicancias socio-pedagógicas ¿reales o futuras? *Boletín Redipe*, 7(11), 155-171
- Sánchez, M. D. M., & Rico, C. (2020). *Traducción automática: conceptos clave, procesos de evaluación y técnicas de posesición*. Comares.
- Sannino, A., Daniels, H., & Gutiérrez, K. D. (2009). *Learning and Expanding with Activity Theory*. Cambridge University Press.
- Schumacher, P. (2020). Post-édition et traduction humaine en contexte académique: une étude empirique, *Transletters. International Journal of Translation and Interpreting*, 3, 239-274.
- Soto, X., Perez-de-Viñaspre, O., Labaka, G., & Oronoz, M. (2019). Neural machine translation of clinical texts between long distance languages. *Journal of the American Medical Informatics Association*, 26(12), 1478-1487. <https://doi.org/10.1093/jamia/ocz110>

Srnicek, N. (2019). *Capitalismo de plataformas*. Caja Negra.

Srnicek, N., & Williams, A. (2017). *Inventar el Futuro. Poscapitalismo y un mundo sin trabajo*. Malpaso.

Steinhoff, J. (2019). *Automation and Autonomy. Labour, Capital and Machines in the Artificial Intelligence Industry*. Palgrave-Macmillan.

Stiegler, B (2001). *La Technique et le temps. Tome 3: Le temps du cinéma et la question du mal-être*. Galilée.

Stiegler, B. (2015). *La société automatique. L'Avenir du travail*. Fayard.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention Is All You Need (v.1). *arXiv*, 1706.03762. <https://doi.org/10.48550/arXiv.1706.03762>

Wolf, M. (2008). *Proust and the Squid: The Story and Science of the Reading Brain*. Faber and Faber.

Wolf, M. (2018). *Reader, Come Home: The Reading Brain in a Digital World*. Harper Collins.



© Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0)
<https://creativecommons.org/licenses/by-nc-nd/4.0/>