NEOPHILOLOGICA 2019, T. 31, 303—319
ISSN 2353-088X
DOI: http://doi.org/10.31261/NEO.2019.31.18

Agnieszka Palion-Musioł
Universidad de Bielsko-Biała
Polonia
https://orcid.org/0000-0003-4244-7866

Traduce el hombre y traduce la máquina El esbozo histórico de la traducción automática

Translating human and translating machine
Historical development of machine translation

Abstract

The paper focuses on various automatic translation systems from the historical perspective in order to show three major directions of its development. The oldest ones, translating word by word, the subsequent ones — based on interlingua, and the newest ones — which use neural networks and machine learning. The paper presents various linguistic and interdisciplinary points of view, methodological proposals and prototypes, which have been made from the 17th to the 21st century.

Keywords

Machine translation, computational linguistics, neural networks, machine learning

1. Introducción

La traducción automática conforma la rama de la Lingüística Computacional la cual se enfoca en crear, implementar, evaluar y aplicar los sistemas y programas informáticos para traducir textos de la lengua de origen a la lengua de llegada. Como muestran los últimos resultados obtenidos por los líderes en la traducción automática actual — Google y DeepL — cada día nos acercamos más al sueño de muchos investigadores, lingüistas e ingenieros que intentaban proponer tal modelo y algoritmo a la máquina para que esta generase textos de alta calidad, comprensibles y comparables a los textos traducidos por los traductores de carne

y hueso. El objetivo del texto es la revisión de los sistemas de traducción automática en su perspectiva histórica y mostrar así sus tres principales ejes de evolución.

2. Desarrollo de la traducción automática desde la perspectiva histórica

Los primeros intentos de traducción automática pueden observarse ya en el siglo XVII cuando los lingüistas propusieron el uso del diccionario mecánico para solucionar los problemas de la lengua. Descartes y Leibniz presentaron entonces teorías sobre la confección de diccionarios basados en códigos numéricos universales (véase, J.A. Alonso Martí, 2003: 96—97). Asimismo, a mediados de dicho siglo, Athanasius Kircher, Cave Beck y Johann Becher elaboraron ejemplos reales que estaban inspirados en la idea de crear una lengua universal, no ambigua, que se centrara en símbolos icónicos y principios lógicos para posibilitar la comunicación global sin problemas ni confusiones. Como escriben Hutchins y Somers, la interlingua de mayor alcance y la más conocida fue la propuesta por John Wilkins en 1668 en el trabajo *Essay towards a Real Character and a Philosophical Language* (J. Hutchins, H. Somers, 1995: 30). Otro ejemplo de interlenguaje internacional fue el esperanto, pero los intentos y esfuerzos dedicados en aquel tiempo a la traducción automática como tal fueron poco visibles, descritos y emprendidos. Sin embargo, en 1933, de forma paralela y autónoma surgieron dos patentes: — una de ellas elaborada en Rusia por Petr Smirnov-Troyanski y otra en Francia, por George Artsouni. El modelo de Artsouni radicaba en un diccionario bilingüe digital y fue en 1937 cuando dio a conocer su prototipo al público. Por otro lado, el proyecto de Troyanski, que fue más significativo e importante para la traducción automática, basó su modelo en tres etapas. Hay que subrayar que su patente se refería únicamente a la máquina descrita en la segunda fase y las otras dos se quedaron en la periferia de su proyecto. Asimismo, el autor ruso sostenía que el análisis lógico de la primera fase podría automatizarse por completo. Por ende, en la primera, habría un editor responsable del análisis lógico de las palabras en la lengua de origen, la única que este conocía, que se ocupaba de reducir estas palabras a sus formas y funciones sintácticas básicas. En la segunda etapa, la máquina transformaba las secuencias de formas básicas y las funciones en sus equivalentes en la otra lengua y, durante la última etapa, el siguiente editor, que conocería únicamente la lengua meta, convertiría el resultado obtenido en las formas comunes propias de dicha lengua. A pesar del carácter innovador y la relevancia del trabajo de Troyanski, cabe destacar que sus investigaciones y proyectos fueron desconocidos en otros países, y no obtuvieron partidarios ni seguidores. No obstante, a finales de la década de los 40, observamos movimientos de mu-

cha importancia para la traducción automática y la lingüística computacional que también tuvieron sus inicios en aquella época. En aquel tiempo, Warren Weaver de la Fundación Rockefeller que era también eximio experto en los trabajos de criptografía, y su compañero Andrew D. Booth, cristalógrafo inglés, diseñaron la posibilidad de usar los ordenadores para traducir (J. Hutchins, H. Somers, 1995: 30; K. Sparck Jones, 1992: 54). En su regreso a Londres, Andrew Booth centró sus trabajos en la automatización de un diccionario bilingüe e inició una colaboración con Richard Richens de Cambridge, quien utilizaba tarjetas perforadas en sus investigaciones y proyectos para hacer traducciones toscas, palabra por palabra, de resúmenes científicos. Sin embargo, los esfuerzos de estos autores no desempeñaron un papel tan crucial como el Warren Weaver, que en julio de 1949 presentó en su informe titulado *Translation*, o más conocido como el *Weaver's Memorandum,* la idea de la Traducción Automática y la necesidad de investigar sobre la polisemia de las unidades lingüísticas, los análisis estadísticos, el uso de técnicas y métodos criptográficos, la relevancia de la teoría de la información de Shannon[1], el estudio de la lógica subyacente y las características universales de la lengua (J. Hutchins, H. Somers, 1995: 30). Estos proyectos animaron a otros investigadores y centros académicos a iniciar los estudios sobre la traducción automática como por ejemplo The Massachusetts Institute of Technology (MIT), la Universidad de Harvard o la Universidad de California. En 1951, se nombró en el MIT al primer lingüista con dedicación exclusiva a la Traducción Automática: Yehoshua Bar-Hillel, conocido por sus trabajos y estudios sobre las gramáticas categoriales. Al año siguiente, el mismo investigador organizó el primer simposio, el *Primer Congreso Internacional sobre Traducción automática*, en el que se discutieron cuestiones de mucha necesidad entre las que se encuentra el análisis sintáctico y morfológico, la necesidad de la intervención humana en la preedición y postedición de textos, la idoneidad de crear sublenguajes, el uso de las técnicas automáticas durante la consulta en los diccionarios, la construcción de una interlingua caracterizada en términos de un sistema de representación abstracta del significado lingüístico y por último, el estudio y la resolución del complejo dilema que presentan la polisemia y la homografía (J. Hutchins, H. Somers, 1995: 30).

Otras universidades, como la de Georgetown, también desarrollaban al mismo tiempo investigaciones en esta misma materia. Por ello, en 1954, Leon Dostert, que colaboró con International Business Machines (IBM), mostró públicamente un programa de traducción automática. El programa tradujo 49 oraciones seleccionadas de un texto especializado del ruso al inglés con el objetivo de confeccionar un diccionario basado únicamente en solamente seis reglas gramaticales y en

[1] Claude Shannon y Warren Weaver en 1949 publicaron su libro *The Mathematical Theory of Communication*, en el que se muestran las relaciones entre las leyes matemáticas y las ciencias de la computación que se ocupa de investigar la información y todo lo que está relacionado con ella — canal, criptografía y compresión de datos.

un vocabulario muy limitado de 250 palabras. La traducción fue realizada sustituyendo palabra por palabra, lo que causó que, finalmente, el proyecto no tuviera mucho valor científico, pero su carácter práctico garantizó el apoyo financiero de muchas empresas y el gran interés sobre la TA en todo el mundo, con mayor relevancia en Estados Unidos y Rusia (J. Hutchins, H. Somers, 1995: 31; J.A. Alonso Martí, 2003: 97). Como observa M. Kay, el método y carácter del programa estaban basados en la perspectiva científica de sus autores que fueron criptógrafos y matemáticos y por eso, su objetivo era codificar un texto inteligible en otro incomprensible y de igual modo, transformar durante la descodificación un signo enigmático en un signo lingüístico (M. Kay, 2003: XVIII).

Todas estas iniciativas fueron apoyadas en 1954 cuando se publicó *Mechanical Translation*, la publicación periódica dedicada a la Traducción Automática. Con el tiempo, en 1965, el título fue cambiado por *Mechanical Translation and Computational Linguistics*, después, en 1974, en *American Journal of Computational Linguistics* y finalmente, a partir de 1980 se publicó bajo el título *Computational Linguistics* (consúltense, M. Kay, 2003: XVII).

Tanto las primeras investigaciones como los programas de traducción automática, los congresos y las publicaciones influyeron sin duda en el desarrollo de la traducción automática, pero nos parece importante señalar unos ejemplos de este período para mostrar la calidad de las traducciones realizadas, en las se prestó muy poca atención a la sintaxis, la semántica y la ambigüedad léxica de la lengua natural.

Aunque los ejemplos presentados señalan faltas y defectos del programa, es cierto que esta muestra llamó la atención e interés de diferentes instituciones financieras en todo el mundo. Del mismo modo, este interés se vio reflejado también en el número de conferencias organizadas sobre el tema, entre otras, en 1956 se organiza el *Segundo Congreso Internacional sobre Traducción Automática*, en ese mismo año también se celebra el *Congreso Internacional de Washington sobre Ciencias de la información* y en 1961 se inicia el *Congreso Internacional de Teddington sobre Traducción automática y Lingüística aplicada* (véase, K. Sparck Jones, 1992: 54).

En la década siguiente, los investigadores se dividieron en dos grupos: unos que buscaban obtener resultados y efectos prácticos en poco tiempo y con un aparato teórico muy limitado, y otros que se centraban en estudios teóricos bien estructurados. Los primeros representaban el método lexicográfico de la Universidad de Washington en Seattle, que después fue utilizado en un programa de traducción del ruso al inglés para las Fuerzas Aéreas de los Estados Unidos por IBM, en el método de ingeniería estadística preparado por RAND Corporation y en los métodos del Institute of Precision Mechanics en la Unión Soviética y en el National Physical Laboratory en Gran Bretaña. Los investigadores más significativos trabajaron en la Universidad de Georgetown, donde se diseñó un sistema de traducción palabra por palabra del ruso al inglés. Los segundos pertenecían

a centros científicos tales como el MIT, a las universidades de Harvard, Berkley y Texas, al Instituto de Lingüística en Moscú, la Universidad de Leningrado, la Universidad de Milán y Grenoble y a la Unidad de Investigación del Lenguaje en Cambridge. Cabe señalar que en el MIT y en la Unidad de Investigación del Lenguaje en Cambridge se ensayaron los primeros sistemas de interlingua y transferencia, a diferencia de los partidarios del enfoque pragmático quienes abogaban por la traducción directa — palabra por palabra (J. Hutchins, H. Somers, 1995: 31).

Las investigaciones de este período influyeron de manera significativa tanto en la Traducción automática como en la inteligencia artificial y la lingüística computacional, y en consecuencia, en diccionarios automáticos y técnicas de análisis sintáctico. No obstante, a pesar de este progreso, los investigadores no consiguieron traducciones fiables y de calidad. Por eso, como se recoge en J. Hutchins y H. Somers, en 1960 el padre de la traducción automática Bar-Hillel, dudó de la realización de traducciones totalmente automáticas de alta calidad que imitaran el lenguaje humano (J. Hutchins, H. Somers, 1995: 32). Además, sugirió la dificultad de superar las barreras semánticas al no poder integrar las enormes cantidades de conocimiento enciclopédico del mundo que poseen los traductores humanos.

La calidad de los primeros textos traducidos era baja no solamente por la capacidad y potencia de los ordenadores disponibles en esa época, sino también por la naturaleza de la programación, que estaba destinada a trabajar con números y no con el lenguaje natural.

Como podemos observar, las carencias y limitaciones de los primeros programas no se pueden negar, aunque los lingüistas contemporáneos como por ejemplo K. Sparck Jones o M.A. Martí Antonín y I. Castellón Masalles indican en esta época los siguientes logros (K. Sparck Jones, 1992: 55; M.A. Martí Antonín, I. Castellón Masalles, 2000: 12):

a) darse cuenta de lo complicado que es la lengua humana y de la escasez de dichos recursos informáticos;
b) los primeros intentos del procesamiento computacional del lenguaje, sobre todo, de los analizadores sintácticos y parsers, así como las investigaciones sobre la ambigüedad léxica y los métodos a través de interlingua, transferencia o traducción directa.

En 1965, el lingüista estadounidense Noam Chomsky da a conocer su estudio principal titulado *Aspectos para una teoría de la sintaxis* que provoca mucha resonancia en la Lingüística Teórica y la Psicología. Cabe subrayar que, tal y como indican R. Grishman o M. Kay, su influencia en la Lingüística Computacional no fue desdeñable, sobre todo en los años sesenta cuando la Gramática Generativa Transformacional tuvo gran repercusión, pese a que su autor no era partidario del empleo de dicha teoría en la Lingüística Computacional (consúltese R. Grishman, 1986 [1991]: 50—81; M. Kay, 2003: XVIII).

En su obra *Estructuras sintácticas* de 1957 N. Chomsky se opuso a la aplicación de las gramáticas independientes del contexto o las gramáticas de estados finitos para analizar el lenguaje en su complejidad (N. C h o m s k y, 1957 [1974]: 32—66). Por eso, en su obra de 1965 presentó un modelo gramatical con transformaciones que podían exponer las complejidades del lenguaje, aunque según R. Grishman, el precursor de dichas transformaciones fue Zellig Harris, su mentor (R. G r i s h m a n, 1986 [1991]: 49—50). Estas propuestas fueron adquiridas por la Lingüística Computacional, pero en corto tiempo fueron rechazadas, mostrando el fracaso de la Gramática Generativa Transformacional en la Lingüística computacional y el regreso a las gramáticas de estados finitos en los años setenta y después, en los años novent, dando lugar a la revolución de la teoría probabilística.

Los estudios y las investigaciones de aquella época se concentraron en el desarrollo teórico de la disciplina y también, en los efectos prácticos obtenidos. Sin embargo, los investigadores buscaban respuestas sobre las perspectivas y los retos de la TA en el futuro. Con este objetivo, en 1964, se formó el grupo ALPAC (de Automatic Language Processing Advisory Commitee) al que pertenecían los expertos de la Academia Nacional de Ciencias (Natural Academic of Sciences), entre los que se encontraba David Hays, quien inventó el término de *Lingüística computacional* como sugiere M. Kay en su trabajo, que valoró los proyectos y estudios dedicados a la traducción automática en los Estados Unidos (M. K a y, 2003: XVII). En el conocido informe ALPAC publicado en 1966, los autores declararon que los resultados obtenidos no eran suficientes ni precisos, y que resultaban dos veces más caros que la traducción humana, además de no existir una expectativa previsible o inmediata de la utilidad de la traducción automática (J. H u t c h i n s, H. S o m e r s, 1995: 32). Del mismo modo, se rechazó realizar nuevas inversiones y proyectos sobre la TA recomendando el apoyo de herramientas informáticas (diccionarios electrónicos) que facilitaran el proceso de traducción a los profesionales. Asimismo se sugería un mayor estudio de la Lingüística computacional.

A pesar de los resultados presentados en el informe ALPAC, la crítica existente y muchos defectos de los programas creados, se siguieron desarrollando proyectos tanto científicos como comerciales. Merece la pena destacar que se desarrolló sobre todo la parte teórica de dichos proyectos, causando un mayor número de aplicaciones. En consecuencia, en 1970, Woods sugirió aplicar de manera computacional las *redes de transición aumentadas* (Augmented Transition Network). Al mismo tiempo otros autores intentaron adaptar las propuestas de Chomsky a la Lingüística Computacional. En 1965 cuando Chomsky publicó su segunda obra, los expertos de MITRE — Zwicky, Hall, Friedman y Walker — trabajaron con una gramática generativa transformacional que, en lo relativo a la producción, contenía 275 reglas sintagmáticas para el componente base y 54 reglas transformacionales, y en caso del reconocimiento, se presentaron alrededor de 550 reglas sintagmáticas para analizar la estructura superficial y 134 reglas transformaciona-

les inversas, es decir, reglas que desde el análisis de la estructura superficial, eran capaces de establecer su estructura profunda. En 1966, en el grupo MITRE participaron Walker, Gross, Geis y Chapin quienes ofrecieron una nueva aplicación de la Gramática Generativa Transformacional que residía en desechar los análisis sintácticos inútiles y, con ello, acelerar todo el proceso de análisis. Grishman, en su libro, muestra que con una oración de 12 palabras se podían generar 48 análisis posibles que eran capaces de ser procesados en diferentes fases de análisis y que podían ser rechazados solamente en la última fase (R. Grishman, 1986 [1992]: 65). Dos años más tarde, en 1968, Charles Fillmore publica su *Gramática de casos,* otro enfoque teórico. La gramática de Fillmore, como la de Chomsky, surge en el ámbito de la Lingüística Teórica e intenta ser aplicada en la Lingüística Computacional como es el caso de los traductores automáticos AtlasI y AtlasII (consúltese M.A. Martí Antonín, I. Castellón Masalles, 2000: 14, 33). También en este año Quillian y en 1975 Woods presentan respectivamente la representación de nódulos que conformaba la estructura o jerarquía semántica entre las palabras y que fue denominada red semántica (M.A. Martí Antonín, I. Castellón Masalles, 2000: 15, 136—38). En cuanto a la Inteligencia Artificial hay que destacar el importante trabajo de 1972 de Winograd, *Understanding Natural Language,* en el que se expusieron las bases para la comprensión del lenguaje natural y otras propuestas también importantes por parte de Minsky o de Schank y Abelson para destacar el papel de la información que proveniente del conocimiento no lingüístico o enciclopédico y que se emplea en el procesamiento del lenguaje. Minsky, en su artículo publicado *A framework for representing the knowledge* en 1975, aseguraba que el conocimiento humano está guardado en la memoria en forma de datos denominados *marcos* o *esquemas* (del inglés *frame*), que representan situaciones estereotipadas. Entre el año 1972 y 1975, Schank inventa el término *dependencia conceptual* según el cual cada oración refleja diferentes conceptos, que guardaban relaciones con otros conceptos más generales que intentaban representar los primitivos semánticos que de acuerdo con la idea de Schank eran imprescindibles para una comprensión correcta de los diferentes enunciados (M.A. Martí Antonín, I. Castellón Masalles, 2000: 15). En 1977 Schank y Abelson publican *Scripts, plans, goals and understanding* en el que dan a conocer el concepto de *guión* (del inglés *script*) que se parece mucho al concepto de *marco*, salvo por su naturaleza dinámica y secuencial: los sucesos dentro de las situaciones estereotipadas están bien delimitados.

Las últimas investigaciones y proyectos desarrollados permitieron estrechar las relaciones entre la Lingüística Teórica y la Lingüística Aplicada y, además, en otros ámbitos tales como la Psicolingüística y la Lingüística Computacional. Tanto la gramática de Chomsky como la de Fillmore fueron creadas en el marco de la Lingüística Teórica, y asimismo, la segunda radicaba su interés tanto en la Psicolingüística como en la Lingüística computacional. Lo mismo observamos con las propuestas de Schank y Abelson o de Minsky que fueron creadas en el

marco de la Inteligencia Artificial, pero con influencia en la Lingüística teórica o en la Psicolingüística.

Entre los proyectos de esta etapa, implementados en la Lingüística Computacional, vale la pena indicar tres sistemas — *Eliza, Lunar y Shrdlu* — basados en la colaboración hombre-máquina a través del uso del lenguaje natural. Como nos informan M.A. Martí Antonín y I. Castellón Masalles, los sistemas se referían a temas muy delimitados e intentaban imitar diálogos naturales de la vida cotidiana (véase M.A. Martí Antonín, I. Castellón Masalles, 2000: 15, 52). *Eliza* fue diseñada en 1699 por Weizenbaum en el MIT y su objetivo principal era simular la conversación entre un psicoterapeuta representado por el ordenador, y el paciente, interpretado por el traductor humano. El tipo de interacción entre *Eliza* y el hombre se denomina como pattern-matching y consistía en palabras claves que estaban jerarquizadas. Según las palabras claves que aparecían en el comunicado del paciente, la jerarquía seleccionaba el tema del diálogo. Si no aparecían dichas palabras, el programa hacía preguntas adicionales para obtener más información, como por ejemplo *Continúa, Háblame más de eso*, etc. (véase, J. Allen, 1995: 6—9). Además, las palabras clave se referían a patrones concretos que se cotejaban con el enunciado anterior del paciente. En caso de la concordancia entre el enunciado y el patrón, se generaba una respuesta asociada a dicha palabra clave. Para ver los ejemplos y el amplio análisis, consúltese Martí Antonín y Castellón Masalles o Allen (M.A. Martí Antonín, I. Castellón Masalles, 2000: 12—13; 1995: 7 respectivamente). Los dos sistemas restantes, *Lunar y Shrdlu*, fueron proyectados en base a las propuestas de Minsky, Schank y Abelson o Winograd. El primer sistema, *Lunar*, tiene su origen en 1973 y fue desarrollado por Woods y diseñado sobre el uso de datos de las muestras geológicas de la luna destinados a los científicos que trabajaban con el programa espacial Apolo. Una vez más, los diálogos en lenguaje natural entre la máquina y los investigadores estaban muy restringidos, esta vez al tema de la geología lunar, para posibilitar esta comuniación y la interacción mutua. Cabe destacar que este sistema poseía más recursos que por ejemplo *Eliza*, porque su autor creía así que el sistema desarrollaría una verdadera habilidad de comprensión de las preguntas formuladas. También se emplearon en el sistema redes semánticas que organizaban la información y que fueron propuestas por Quillian. La representación tenía lugar a través de nudos y representaba entidades, y mediante arcos se representaban las relaciones entre las entidades de los nudos que se expresaban mediante las fórmulas: *ser clase de, ser parte de,* etc. Para ilustrar el éxito del programa y su importancia, cabe señalar que el sistema interpretaba correctamente casi el 90% de los enunciados producidos por el traductor humano (consúltense, M.A. Martí Antonín y I. Castellón Masalles, 2000: 15). Los trabajos de J. Allen y el anteriormente mencionado de M.A. Martí Antonín y I. Castellón Masalles aportan más información sobre este tema (véase, J. Allen, 1995: 419; M.A. Martí Antonín, I. Castellón Masalles, 2000: 15, 136—138).

El último sistema que posibilitaba la comunicación entre máquina y hombre era *Shrdlu* de Winograd, propuesto en 1972. Este programa utilizaba la gramática sistémica de Halliday y se inscribía dentro del enfoque teórico de la Inteligencia Artificial. El marco de actuación de este sistema abarcaba una realidad muy limitada referida a colores, formas determinadas y dimensiones. Este mundo se reflejaba en una red semántica en la que se caracterizaban las propiedades y la naturaleza de los objetos así como la localización de dichos objetos en el sistema (consúltense, K. Sparck, 1992: 56; M.A. Martí Antonín, I. Castellón Masalles, 2000: 15). Además de esta realidad enmarcada, el ordenador realizaba peticiones. Lo que merece la pena destacar es que el sistema era capaz de resolver algunas ambigüedades lingüísticas y solicitar información y datos para solucionarlas. *Shrdlu* era un sistema bastante restringido por su limitado mundo de bloques, pero su progreso tanto en el ámbito semántico como sintáctico y pragmático causó que desempeñara un papel importante en la Lingüística Computacional.

Para ser precisos, tal y como requiere el título del presente artículo, queremos describir, aunque de manera general, otros sistemas de aquel tiempo que a veces quedan olvidados y que tenían como objetivo la interacción entre la máquina y el hombre o la comprensión del lenguaje natural.

En 1975, Colby desarrolla su propuesta de un sistema de interacción máquina-hombre denominado *Parry*. El objetivo de este sistema era inventar y formular situaciones que no habían pasado en el mundo real. Este sistema actuaba mediante el pattern-matching que también usaba *Eliza*. Podía hacer análisis sintácticos y morfológicos pequeños y asimismo sus conocimientos pragmáticos le permitían entender las implicaciones de un enunciado. Otro aspecto interesante de este programa es que podía imitar estados emocionales humanos como por ejemplo el temor, la vergüenza o la ira. Para ampliar información, consúltense las obras de M. Meya y W. Huber y M.A. Martí Antonín y I. Castellón Masalles (M. Meya, W. Huber, 1986: 159—160; 166—171; M.A. Martí Antonín, I. Castellón Masalles, 2000: 15—19, 43—44, 52—53).

Ese mismo año, en 1975, Schank diseña otro sistema denominado MARGIE (del inglés Memory Analysis. Response Generation in English). Su objetivo era la comprensión del lenguaje natural basado en el modelo de dependencia conceptual propuesto por Schank. De acuerdo con su objetivo, el autor intentó crear un sistema capaz de comprender el lenguaje natural y en la práctica, capaz de parafrasear de diferentes maneras los enunciados y sacar tanto el significado literal como el relacionado con nuestra visión y conocimiento del mundo, pues el sistema podía hacer inferencias. Por eso, el programa realizaba análisis de las oraciones y producía representaciones que eran dependencias conceptuales. Estas representaciones le servían para percibir las posibles inferencias de la oración. Por ejemplo, en la oración *María dio a comer a su hijo*, el sistema podía interpretar que María creía

que su hijo tenía hambre, que el hijo no había comido por algún tiempo, que el hijo pedía algo de comida y que el hijo comería algo.

Otros programas parecidos a MARGIE, que utilizaban en sus proyectos los conceptos de guión y planes de Schank y Abelson, surgieron en el año 1978, — el sistema SAM, — proyectado por Cullingford y — el sistema PAM, — de Schank, Abelson y Wilensky perteneciente a los años 1977 y 1978. Cabe mencionar también los sistemas desarrollados a partir el uso del concepto de marco de Minsky, en 1977, denomiado MS MALAPROP y diseñado por Charniak así como el de 1980, TDUS, creado por Robinson. Para profundizar en el tema, sugerimos las obras citadas de Meya y Huber y Martí Antonín y Castellón Masalles (M. Meya, W. Huber, 1986: 159—160, 166—171; M.A. Martí Antonín, I. Castellón Masalles, 2000: 15—19, 43—44, 52—53).

Por deber científico, no podemos quedarnos sin comentar brevemente también otros sistemas como por ejemplo HAM-ANS (Hamburg Application-oriented Natural-language System) o TEAM y otros parecidos. El primer sistema fue proyectado entre 1978 y 1983 por Wahlster y sus colegas con el objetivo de comunicarse con el ordenador para comprobar las reservas disponibles en el hotel. Hay que mencionar que este tipo de programas fueron preparados tomando en cuenta diferentes problemas y complejidades del diálogo real. En consecuencia, el programa era capaz de entender órdenes, preguntas directas e indirectas, peticiones del usuario humano e incluso, cumplir dichas peticiones. Su actitud era cooperativa y activa y por eso, podía entender el tema de la conversación y los referentes conversacionales. Por otro lado, estos sistemas tenían una selección limitada del léxico, pero el léxico incluido codificaba la información semántica, sintáctica y morfológica, por para solucionar las redundancias de la lengua y manejar adecuadamente los modalizadores del lenguaje y para tomar en cuenta el contexto situacional en que el diálogo tenía lugar. Así pues, el sistema incorporaba modelos de discursos. Además, merece la pena subrayar que HAM-ANS era un sistema apto para generar elipsis y resolver anáforas. En cuanto a las elipsis, el sistema al principio eliminaba de la estructura sintáctica las ramas idénticas en la representación arbórea y después, verificaba que la estructura semántica recibida fuera similar a la representación recibida en la oración emitida anteriormente por el usuario humano. Para resolver la anáfora, el sistema determinaba el antecedente de un pronombre según su grado de tematicidad.

También Bobrow, Kaplan y Kay diseñaron el sistema parecido a HAM-ANS en 1976 que llamaron GUS. El programa estaba destinado a entablar diálogo con el usuario humano para que este pudiera hacer la reserva de un avión.

Con el tiempo surgieron otros sistemas de este tipo — FAS-80, proyectado en 1980 por Witchas, Zänker y Heibig para obtener información de los documentos disponibles en una biblioteca de programas o en el mismo año, el programa KOPRO que desempeñaba el papel de un tutor de estudios o de 1981 el sistema ATN-BIC propuesto por Metzing que ofrecía diferentes posibles viajes a realizar.

A continuación mencionamos el sistema TEAM (Transportable English Data Manager) que fue creado por Grosz, Appelt, Martin y Pereira y que representaba un sistema de consulta abarcando tres etapas — adquisición del conocimiento, procesamiento del lenguaje (llamado DIALOGIC) y finalmente, la consulta de los datos del sistema.

Otro sistema de comprensión del lenguaje natural fue diseñado en 1976 por Meehan y capaz de generar historias coherentes. Con el tiempo, en 1988, Hovy propone otro sistema denomiado Pauline que ofrecía un abanico de versiones de un mismo texto según las necesidades y los objetivos del usuario. En 1981, Sager crea el sistema LSP que era capaz de analizar documentos médicos con el fin de almacenar la información recibida de estos informes en una base de datos a través del proceso de traducción.

Las repercusiones del informe ALPAC fueron muy negativas para la situación científica y económica de la Traducción Automática. A muchos centros e investigadores se les limitó la ayuda financiera causando que la TA tuviera un desarrollo más lento. Tal situación provocó que los esfuerzos llevados a cabo se observaran sobre todo fuera de Estados Unidos, en Europa Occidental y Canadá, y no ganaron mucho interés por parte de los investigadores y científicos. Los proyectos estadounidenses se habían enfocado en textos técnicos traducidos del ruso al inglés. Los europeos, que tenían unas expectativas y necesidades diferentes a las de los canadienses, requerían la traducción de textos técnicos, científicos, legales y administrativos de todas las lenguas, mientras que los canadienses, dada su situación bicultural, requerían de la traducción inglés-francés. Esta demanda del mercado canadiense animó a un grupo de investigadores de TAUM de Montreal (Traduction automatique de l'Université de Montréal) a crear un programa de traducción inglés-francés para manuales de aeronáutica, pero el proyecto fracasó y, en consecuencia, en 1976, se construyó otro programa *Météo* del Centro Meteorológico Canadiense que, a día de hoy, traduce alrededor de 37.000 palabras meteorológicas de emisión diaria, con muy buena precisión llegando a alcanzar el 90%. Los ejemplos de la traducción de este programa y el análisis detallado del sistema se encuentran en los trabajos de autores conocidos en este artículo, J. Hutchins y H. Somers y, en el campo polaco, en Ł. Bogucki (J. Hutchins, H. Somers, 1995: 287—304; Ł. Bogucki, 2009: 44—45).

Ese mismo año, la Comisión de las Comunidades Europeas decidió explotar un programa, Systran, de traducción inglés-francés proyectado por Peter Toma para hacer traducciones del ruso al inglés. El programa fue diseñado para las Fuerzas Aéreas estadounidenses y en Europa se utilizaba desde 1970. Con el tiempo, al programa se le añadieron otras traducciones francés-inglés, inglés-italiano e inglés-alemán destinadas a la Comisión Europea. También en los años setenta, se decidió desarrollar un programa multilingüe, *Eurotra,* basado en los últimos avances en Lingüística Computacional y Traducción Automática que abarcaba todos los idiomas de la Comunidad. *Eurotra,* en su proyecto, contiene muchas

soluciones y propuestas llevadas a cabo en Grenoble y Saarbrücken. Es decir, en los años sesenta, los expertos de Grenoble (la corriente perfeccionista) habían diseñado un sistema de interlingua para realizar traducciones entre el francés y el ruso. Hay que señalar que el sistema creado no era exactamente un sistema de interlingua ya que empleaba la transferencia léxica bilingüe. Sin embargo, los resultados obtenidos no cumplieron con los objetivos previos y los investigadores dejaron de trabajar en el proyecto. Su atención se dirigió hacia llamó otro sistema de la misma universidad — *Ariane* — desarrollado por Bernard Vauquois y basado en la transferencia. *Ariane* era un sistema de segunda generación denomiado también GETA de Groupe d'Études pour la Traduction Automatique. Los franceses tuvieron mucha esperanza en *Ariane*, pero la valoración no fue satisfactoria. Por otro lado, los expertos alemanes de Saarbrücken al final de los años sesenta desarrollaban un proyecto de transferencia multilingüe denominado *Susy* que formaba la sigla de Saarbrücken Übersetzungssystem. El sistema hacía traducciones ruso-alemanas, pero después, se añadieron otros idiomas como el inglés, el francés y el esperanto. Durante su existencia, casi 20 años, el programa fue desarrollado y ha mejorado mucho. La descripción del programa junto con un detallado análisis se puede consultar en Hutchins y Somers (J. Hutchins, H. Somers, 1995: 267—285).

Cabe destacar que en aquella época, los investigadores estaban de acuerdo de que los mejores resultados para mejorar los sistemas de TA residían en los programas de transferencia (consúltense, J. Hutchins, H. Somers, 1995: 32—33; Ł. Bogucki, 2009: 46—47). A la misma conclusión llegaron los especialistas del LRC (Linguistics Research Center) en Austin, Texas, haciendo pruebas con un programa interlingüe. Por lo tanto, dirigieron su atención y empezaron a trabajar con el programa *METAL* basado en la transferencia. En aquel entonces, los programas basados en la transferencia eran muy populares y abarcaban casi todas las lenguas, lo que podemos apreciar en el ejemplo del sistema *Mu* creado y desarrollado en la Universidad de Kyoto en Japón para traducción japonés-inglés.

Queremos destacar que esta etapa es muy importante para la lingüística computacional dado que se aplican lenguajes de programación de alto nivel. Según Martí Antonín y Castellón Masalles una de las barreras computacionales reside en los lenguajes de programación, que en su mayoría estaban destinados a procesar datos simples, a los que pertenecen los números, y no datos simbólicos, que representan el lenguaje natural (M.A. Martí Antonín, I. Castellón Masalles, 2000: 14). Así, desde entonces, los sistemas están diseñados con lenguajes de programación, de alto nivel, algo que se puede observar en numerosos ejemplos, algunos ya mencionados como *Susy* basado en este tipo de lenguaje de programación FORTRAN (véase, J. Hutchins, H. Somers, 1995: 267).

En los años ochenta, los programas con interlinguas se han ido acercando a los sistemas basados en la transferencia. Merece la pena destacar el papel de la Universidad de Carnegie Mellon en Pittsburgh y su investigación sobre los siste-

mas basados en el conocimiento y que se basan en el desarrollo de sistemas de comprensión del lenguaje natural. Los investigadores subrayan que es necesario que la Traducción Automática abarque no solamente la información lingüística que representan la sintaxis y la semántica, sino también el entendimiento del contenido y el sentido de los textos, es decir, que se refiera a lo real, al mundo real. Este método entiende la traducción mediante elementos extralingüísticos. Del mismo modo, dos propuestas holandesas se han basado en la interlingua, aunque las estrategias no se basaban principalmente en la Inteligencia Artificial. En el caso del proyecto *Rosetta* desarrollado por Jan Landsbergen en Philips, en Eindhoven, entre los años 1980 y 1985, se aplicó la semántica de Richard Montague, que sirve de base a una interlingua, y en el segundo caso del traductor automático, *DLT* (*Distributed Language Translation*) de Utrecht desarrollado por la empresa *Buro voor Systemontwikkeling* se apoyó en un método de interlingua basado en una versión del esperanto (J. Hutchins, H. Somers, 1995: 397—398). Además, en 1982 la Comunidad Europea encargó el sistema de traducción EUROTRA que abarca las lenguas de los países miembros (J. Hutchins, H. Somers, 1995: 327—328).

Con el tiempo los lingüístas han creado sistemas destinados a la traducción automática del habla, algo que a muchos investigadores les parecía imposible en aquella etapa. Las investigaciones dedicadas a la síntesis y el reconocimiento de la voz humana han dado pie al diseño de nuevos proyectos en Gran Bretaña y en Japón. En Japón se han iniciado los trabajos con ATR, es decir, Advanced Telecommunications Research, y en Gran Bretaña, se ha desarrollado British Telecom.

En los años ochenta abunda el nacimiento de gramáticas denominadas de unificación y estas propuestas han animado a otros centros y empresas para volver a investigar y construir este tipo de sistemas de traducción automática. Actualmente, los estudios más avanzados se observan en los centros de IBM en Nueva York.

Para trazar una visión completa hay que mencionar también la obra de Chomsky que en 1981 publicó su *Lectures on Government and Binding* cuyos conceptos se enfocaron en las cuestiones de pronombres, anáfora y categoría vacía, sin embargo, su influencia fue muy limitada en la Lingüística computacional (A. Moreno Sandoval, 1998: 43—44). El apoyo empieza con la obra editada por Joan Bresnan en 1982 titulada *The Mental Representation of Grammatical Relation* con la Gramática Léxico-Funcional (*Lexical Functional Grammar*). Después, Gazdar, Klein, Pullum y Sag publican su propuesta denominada *Gramática de Estructura Sintagmática Generalizada* (*Generalised Phrase Structure Grammar*) en 1985. Y en 1987 Pollard y Sag dan a conocer la teoría que es una reformulación de la anterior teoría de Gazdar llamada *Gramática de Estructura Sintagmática Generalizada* (*Head-driven Phrase Structure Grammar*).

Es necesario que mencionemos que no todas las teorías y gramáticas adaptadas por la Lingüística Computacional fueron gramáticas de unificación, sino

que existen otras concepciones y propuestas que a pesar de tener poca relevancia para la Lingüística Teórica sí tuvieron mucha importancia para la Lingüística Computacional. Ejemplos de este tipo son la gramática de Joshi conocida como *Gramática de Adjunción de Árboles* (*Tree-Adjoining Grammar*) de 1984 y el *Linguistic String Project* iniciado ya en los años sesenta (M.A. Martí Antonín, I. Castellón Masalles, 2000: 17). Según la gramática de Joshi, dos estructuras sintagmáticas pueden dar lugar a otra estructura sintagmática, nueva, a través de la adjunción o de la substitución (elementos no adjuntos) (véase, A.K. Joshi, Y. Schabes, 1997: 70—78). En cuanto al proyecto *Linguistic String Project,* se caracterizó por la gramática referida al enfoque estructuralista propuesto en los años sesenta por Zelling Harris y que contenía un analizador. En este proyecto, la gramática constaba de 200 reglas de reescritura y 200 restricciones y además de un lexicón que incluía información sintáctica y semántica (M.A. Martí Antonín, I. Castellón Masalles, 2000: 17).

El avance tecnológico de este último período sugiere que los problemas de la Lingüística Computacional no eran técnicos, es decir, no estaban relacionados con el procesamiento o la capacidad de memoria de los ordenadores, sino que eran problemas lingüísticos.

Entre los éxitos más importantes de los avances de este tiempo hay que destacar el papel de los analizadores sintácticos denominados *parsers*. Antes no se distinguía entre las reglas gramaticales y el procedimiento empleado para usar estas reglas en el análisis. Es decir, se distingue entre la información procedimental (los parsers) y la información declarativa (la gramática) (A. Moreno Sandoval, 1998: 42—43).

La separación de la gramática y del parser permite que la gramática de una lengua dada sea aplicada al sistema de traducción automática y además, a otros sistemas que sirven para extraer la información o los sistemas de diálogos. También, esta distinción genera la posibilidad de aplicar una misma gramática tanto en el análisis como en la generación. En consecuencia, en los programas de diálogos se puede usar la misma gramática y durante la generación o el análisis se cambiará la forma en la que se use este conocimiento gramatical, es decir, el conocimiento declarativo permanece, pero el conocimiento procedural no.

Los años noventa se caracterizan por un mayor acceso y oferta de los traductores automáticos destinado a los usuarios privados como por ejemplo *PC-Translator*, *Trados*, *SDLX*, *Déjà-Vu*, *Transit*, *Wordfast, Translation Manager* y muchos otros. Los programas disponibles generan los resultados que requieren una revisión humana, pero cabe subrayar que facilitan mucho el trabajo al traductor y le permiten ser competitivo en el mercado gracias a su tiempo y eficacia.

En los años noventa se introducen dos modelos: el modelo probabilístico, que incluye conocimientos estadísticos; y el modelo biológico, que incorpora lo relacionado con la Neurología o Biología. Frente a los modelos simbólicos, estos no ofrecen la resolución de los problemas lingüísticos solamente en base al conoci-

miento lingüístico, sino que intentan hacerlo con la ayuda de otras herramientas científicas como la Biología o la Neurología anteriormente mencionadas (véase, A. Moreno Sandoval, 1998).

En el siglo XXI predominan las redes neuronales y el aprendizaje automático (*machine learning*). El sistema neuronal aprende la lengua humana imitando la manera en la que lo hace la gente. La máquina se construye sobre el modelo de la lengua y analiza las oraciones como unidades enteras. Este avance permite eliminar la mayoría de errores que se cometían en los sistemas estadísticos que analizaban las oraciones por palabras o por fragmentos máximos de cinco unidades lexicales.

Últimamente, en 2016 la compañía *Google* anunció su nuevo sistema de traducción automática que ha estado basado en las redes neuronales artificiales. A esta gran novedad se la ha denominado *Neural Machine Translation System*. El proyecto utiliza lo que se llama una *comunidad de traductores,* que permite que el sistema — la red neuronal — interprete toda la frase en su contexto y tenga en cuenta su estructura gramatical. Este retorno de la traducción palabra a palabra a la traducción global basada en el contexto de la oración da muy buenos resultados. El avance observado tiene también otra consecuencia que consiste en eliminar de la traducción la lengua "puente" que era el inglés, es decir, todas las traducciones hechas en la lengua de origen se traducía al inglés y después se buscaba un equivalente del término inglés en la lengua de llegada. En el proyecto se experimentó con realizar una traducción entre dos lenguas que el sistema neuronal no conocía y el efecto fue sorprendente. La máquina traducía correctamente, es decir, conectaba palabras y conceptos que había visto por primera vez en el texto traducido. Como resultado, el sistema había desarrollado su propio lenguaje con que se permitía representar el texto y después traducirlo a la lengua meta. Las consecuencias del proyecto fueron muy importantes y muestran que las enormes bases de datos que tiene y que emplea *Google* permiten no solamente el *machine learning*, sino también, como en este caso, el *deep learning*[2].

Otro proyecto importante y bastante reciente ya que surgió en 2017 como respuesta al traductor *Google Neural Machine Translation,* es el sistema DeepL, de una pequeña compañía alemana. Este proyecto, aunque no tiene unas bases de datos tan potentes como tiene Google, obtiene mejores resultados según muchos especialistas[3] al emplear un diccionario multi-idioma que se conoce comúnmente como *Linguee*. Este último invento detecta y traduce con más precisión y sentido las unidades fraseológicas, idiomáticas y frasemas.

Actualmente, los dos líderes en traducción automática incorporan nuevas lenguas en sus sistemas para que sus alcances se extiendan y atiendan a cada vez

[2] https://www.bbva.com/es/avances-traductor-google-nos-dicen-futuro-inteligencia-artificial/ (consulta: el 10 de abril de 2019).

[3] https://www.xataka.com/servicios/deepl-vs-google-translate-quien-gana-batalla-traductores-online (consulta: el 10 de abril de 2019).

mayores pares de lenguas habladas en el mundo. En el año 2019 las estadísticas de Google indican 103 lenguas que es capaz de combinar y traducir y mientras que DeepL cuenta con 9.

3. Conclusiones

La revisión de los sistemas de traducción automática desde la perspectiva histórica nos muestra tres tendencias en su evolución y desarrollo. Empezando con la más básica, en la que se traducía palabra por palabra, otra basada en interlingua que presenta una representación del contenido y finalizando con la última tendencia, la más avanzada, estadística que utilizan la Inteligencia Artificial. Los neurones artificiales obtienen los mejores resultados ya que analizan toda la frase, permitiendo muchas veces desambiguar el sentido de la oración y sus componentes. Actualmente, los trabajos van más allá y se intenta que el sistema neuronal traduzca todo el apartado para que sus efectos sean más precisos y tengan en cuenta un mayor contexto.

Referencias citadas

Allen J., 1995: *Natural Language Understanding*. Second edition. Benjamin/Cumming.

Alonso Martí J.A., 2003: "La traducción automática". En: M.A. Martí Antonín, coord.: *Tecnologías del lenguaje*. Barcelona: Editorial UOC.

Bogucki Ł., 2009: *Tłumaczenie wspomagane komputerowo*. Warszawa: Wydawnictwo Naukowe PWN.

Chomsky N., 1957 [1974]: *Estructuras sintácticas*. Méjico: Siglo XXI.

Grishman R., 1986 [1992]: *Introducción a la lingüística computacional*. Madrid: Visor.

Hutchins J., Somers H., 1995: "Introducción a la traducción automática". *Lingüística y conocimiento*, 21 [Madrid, Visor Distribuciones, S.A.].

Joshi A.K., Schabes Y., 1997: "Tree-Adjoining Grammars". In: G. Rozenberg, A. Salomaa, eds.: *Handbook of Formal Languages*. Vol. 3. Berlín: Springer, 69—123.

Kay M., 2003: "Introduction". In: R. Mitkov, ed.: *The Oxford Handbook of Computational Linguistics*. Nueva York y Oxford: Oxford University Press, XVI—XIX.

Martí Antonín M.A., Castellón Masalles I., 2000: *Lingüística Computacional*. Barcelona: Ediciones Universitat de Barcelona.

Meya M., Huber W., 1986: *Lingüística computacional*. Barcelona: Teide.

Moreno Sandoval A., 1998: *Lingüística computacional*. Madrid: Síntesis.
Schank R., Abelson R., 1977a: *Scripts, Plans, Goals and Understanding*. Hillsdale (N.J.): Lawrance, Erlbaum.
Schank R., Abelson R., 1977b: *Guiones, planes, metas y entendimiento*. Barcelona: Paidós Ibérica Ediciones S.A.
Sparck Jones K., 1992: "Natural Language Processing". In: W. Bright, ed.: *International Encyclopedia of Lynguistics*. Vol. 3, 53—59.

https://www.bbva.com/es/avances-traductor-google-nos-dicen-futuro-inteligencia-artificial/ (consulta: el 10 de abril de 2019).
https://www.xataka.com/servicios/deepl-vs-google-translate-quien-gana-batalla-traductores-online (consulta: el 10 de abril de 2019).