

TRADUCCIÓN AUTOMÁTICA E TRADUCCIÓN ASISTIDA POR ORDENADOR: ASPECTOS TERMINOLÓXICOS E TIPOLOXÍA

Javier Gómez Guinovart
Universidade de Vigo

Unha das causas do escepticismo con que moitas persoas interesadas neste campo contemplan as posibilidades da traducción automática (TA) e da traducción asistida por ordenador (TAO) radica nos pobres resultados ofrecidos polos programas informáticos comerciais autodenominados de “traducción automática”, e na xeneralización inxustificada desta percepción negativa ó conxunto de aplicacións informáticas especificamente deseñadas para a súa incorporación no proceso da traducción humana. O obxectivo deste artigo é presentar unha perspectiva global das posibilidades da informática aplicada á traducción, ofrecendo ó mesmo tempo os instrumentos terminolóxicos e conceptuais necesarios para a súa sistematización.

Con este fin, cómpre establecer, en primeiro lugar, certas distincións terminolóxicas necesarias que considero de grande utilidade, inspiradas nalgunhas ideas elaboradas na recente introducción á traducción automática de John Hutchins e Harold Somers¹, xa expresadas con algunhas diferencias por Martin Kay² a principios da década de 1980. Estas distincións terminolóxicas aplícanse ós diversos termos utilizados para referirse ás distintas modalidades de traducción que resultan dos diferentes graos de colaboración entre as persoas e os ordenadores, e poden resultar moi útiles para non caer nas confusións derivadas da ambigüidade semántica do termo “traducción automática”.

O termo “traducción totalmente automática de calidade alta” (TTACA)

1. Hutchins, W. J. e H. L. Somers, *An Introduction to Machine Translation*, Academic Press, Londres, 1992.

2. Kay, M., “Machine Translation”, *American Journal of Computational Linguistics*, 8 (2), 1982, pp. 74-78.

é a miña versión galega do termo orixinal inglés “Fully Automatic High Quality Translation” (FAHQT), cuñado en 1960 por Yehoshua Bar-Hillel³ para referirse ó obxectivo final e ideal da investigación sobre traducción por ordenador. Con este termo aludiremos a un programa informático, aínda inexistente, capaz de traducir calquera texto de calquera xénero entre dúas linguas, sen importa-la dificultade do texto orixinal, nin a distancia cultural entre as linguas implicadas. Se desexasemos equipara-lo termo “traducción automática” coa “traducción totalmente automática de calidade alta”, poderíamos afirmar sen ningunha dúbida que a “traducción automática” neste sentido non existe, nin é probable que vaia existir nun futuro máis ou menos próximo.

Por esta razón, entre outras, o emprego máis común do termo “traducción automática” non adoita referirse á TTACA, senón que alude ó tipo de procesamento informático que vou denominar “traducción semiautomática con intervención humana” (TSIH). Os programas informáticos de traducción semiautomática son programas capaces de ofrecer unha traducción do texto orixinal que debe ser controlada pola persoa que supervisa o seu funcionamento para acadar unha calidade de traducción similar á profesional humana. En xeral, estes programas producen de maneira automática unha primeira versión en borrador do texto, que debe ser corrixida a conciencia para acadar unha calidade estándar no mercado actual da traducción. Algúns dos programas máis coñecidos que englobaríamos nesta categoría son *Systran* e *Metal*, para estacións de traballo, e *Language Assistant* e *Power Translator*, para ordenadores persoais⁴. Dado que non existe TTACA na maioría dos contextos reais de traducción, parece lóxico que o termo TA se aplique a esta categoría de programas de TSIH capaces de ofrecer unha traducción sen intervención humana, a pesar de que a calidade dos seus resultados non sexa aceptable para a meirande parte dos actuais consumidores de traduccions.

Con todo, antes de poder avaliar con xustiza a calidade destes programas de TSIH, non hai que perder de vista o destino que se lle vai dar á traducción obtida. Evidentemente, o grao de esixencia será distinto para unha traducción de uso interno nunha empresa, que para unha traducción de gran tirada que se deba vender nos quioscos. Segundo os datos recompilados por Muriel Vasconcellos⁵, a Comisión das Comunidades Europeas traduce ó redor de 30 millóns de palabras ó ano co programa *Systran*, o que representa

3. Bar-Hillel, Y., “Some reflections on the present outlook for high quality machine translation”, *Advances in Computers*, 1, Academic Press, Nova York, 1960, pp. 91-163.

4. Para unha análise comparativa dos programas de TSIH para ordenadores persoais véxanse Gómez Guinovart, J., “El inglés más cerca: Cinco programas de traducción inglés-castellano para PC”, *Binary*, 53, 1993, pp. 51-67; Gómez Guinovart, J., “Programas de traducción para Windows”, *BYTE España*, 1, 1994, pp. 132-145; e Díez, J., “Traductores para Windows”, *PC World*, 115, 1995, pp. 330-340.

5. Vasconcellos, M., “The Current State of MT Usage”, *MT News International*, 6, 1993, pp. 12-17.

un 15% do total das traducións realizadas por este organismo comunitario. Os documentos traducidos por *Systran* non reciben ningunha revisión, xa que son de uso interno e teñen unha finalidade puramente informativa. O exemplo de *Systran* ilustra algunhas das motivacións do postulado que Juan C. Sager defendeu recentemente no seu libro *Language Engineering and Translation*: “machine translation can, under certain conditions and in certain circumstances, be an acceptable substitute for human translation”⁶.

Outra das posibles motivacións deste razoamento de Juan C. Sager proviría da existencia de programas capaces de traducir cun nivel de fiabilidade similar ó humano en dominios do coñecemento moi restrinxidos. Por exemplo, o programa *Taum-Météo* é capaz de traducir con alta calidade os partes meteorolóxicos do inglés ó francés sen a penas intervención humana (cálculase que os textos que necesitan ser revisados non chegan ó 5% do total dos textos traducidos por este sistema). O Departamento de Medio Ambiente de Canadá utiliza *Taum-Météo* desde 1977, e pasou de traducir 8,5 millóns de palabras ó ano en 1984 a traducir 17 millóns de palabras ó ano na actualidade⁷.

En resumo, se ben é certo que a calidade dos resultados da TSIH non sería aceptable pola maioría dos compradores actuais de traducións, o rendemento destes produtos informáticos é moito máis satisfactorio cando se trata de producir borradores de tradución urxentes e a baixo prezo e, sobre todo, cando o programa de tradución está especificamente deseñado para traducir un tipo de textos similar ó dos partes meteorolóxicos, cun léxico moi restrinxido e unha sintaxe limitada a un número mínimo de construcións gramaticais.

Ademais dos programas de TSIH, existe unha ampla gama de aplicacións informáticas que, a pesar de non estaren concibidas especificamente para o labor da tradución, ocupan un lugar privilexiado entre as ferramentas utilizadas polas persoas que se dedican a esta actividade. Estas aplicacións, que se adoitan englobar dentro do termo “tradución humana con axuda do ordenador” (THAO), poden provir de tecnoloxías tales como a ofimática, a telemática, a xestión documental e as aplicacións lingüísticas da informática.

A aplicación ofimática máis utilizada polas persoas que traducen é sen dúbida o procesador de textos, aínda que neste labor se poden empregar tamén con moito proveito outras aplicacións ofimáticas como o recoñecemento óptico de caracteres⁸ ou os sistemas de dictado para a conversión da

6. Sager, J. C., *Language Engineering and Translation*, John Benjamins, Amsterdam, 1994, p. 261-262.

7. Datos recompilados do artigo xa citado de Muriel Vasconcellos, e de Isabelle, P. e L. Bourbeau, “*Taum-Aviation: Its Technical Features and Some Experimental Results*”, *Computational Linguistics*, 11 (1), 1985, pp. 18-27.

8. Para unha análise comparativa de diversos programas de recoñecemento óptico de caracteres, véxase Eglowstein, H., “*Reconozca su carácter*”, *BYET España*, 2, 1994, pp. 90-94.

voz en texto (como *Personal Dictation System*⁹). Por outra banda, as redes telemáticas substitúen con vantaxe ó correo postal como medio de recepcións dos orixinais e de envío das traducións, ademais de permiti-lo acceso á información espallada nas que hoxe se denominan “superautoestradas da información” ou “ciberespacio”¹⁰. Por poñer só algúns exemplos, a través destas autoestradas electrónicas, podemos consultar inventorios bibliográficos de traducións (como o *Index Translationum* da Unesco), repertorios lexicográficos (como o *Webster’s*) ou bancos de datos terminolóxicos (como *Eurodicautom* da Comisión das Comunidades Europeas¹¹), ademais de poder entrar a formar parte de diversos grupos de discusión internacionais para intercambiar coñecementos, ideas ou experiencias con profesionais da tradución de practicamente todo o mundo.

Por outra parte, os sistemas de xestión documental (tamén denominados sistemas de xestión de bases de datos documentais ou sistemas de recuperación da información) son programas informáticos que permiten realizar consultas sobre unha base de datos constituída por un conxunto de documentos que poden ser artigos de revistas científicas, noticias dos xornais, informes técnicos, libros, traducións ou calquera outro tipo de texto. Cando se dispón dun número considerable de documentos electrónicos, en moitos casos é importante podelos manexar coa flexibilidade e eficacia que permiten as ferramentas de recuperación da información incorporadas nos sistemas de xestión documental (como *Personal Librarian* ou *Idealist*, ámbolos dous para ordenadores persoais)¹².

Por último, entre as diversas aplicacións lingüísticas da informática que adoitan ser utilizadas no proceso da tradución, destacarémo-los programas de corrección ortográfica, sintáctica e estilística –habitualmente integrados no contorno de traballo dos procesadores de textos–, os dicionarios electrónicos monolingües, plurilingües e de sinónimos, e os sistemas de xestión de terminoloxía (como *MultiTerm* ou *Term-PC*, para ordenadores persoais), tamén denominados sistemas de xestión de bases de datos terminolóxicos, e que

9. Véxase Diehl, S., “Escribir al dictado”, *BYTE España*, 1, 1994, pp. 42-44, para unha descrición do programa de dictado *Personal Dictation System*.

10. Para unha primeira aproximación ó uso das redes informáticas por parte dos tradutores, véxase Obenaus, G., “The Use of Computer Networks for Translators”, *Language International*, 5 (1), pp. 6-10.

11. Véxase Román, E., “Eurodicautom: La base de datos terminológica de las Comunidades Europeas”, *Enseñanza de lenguas y traducción con ordenadores*, Ediciones Pedagógicas, Madrid, 1995, pp. 127-168, para unha descrición das características e funcionamento de *Eurodicautom*.

12. Para unha excelente introducción ós sistemas de xestión documental, véxase Codina, L., “Sistemas de gestión documental”, *Binary*, 62, pp. 114-119; 63, pp. 92-100; e 64, 1994, pp. 106-112.

serven para crear, administrar e consultar vocabularios plurilingües electrónicos que normalmente poden conter gráficos e ilustracións¹³.

Tamén existe un determinado tipo de aplicacións informáticas, como *TranslationManager* ou *Translator's Workbench*, que configuran un contorno integrado de ferramentas para a tradución¹⁴. Estas aplicacións integrais adoitan incluír un procesador de textos especialmente deseñado para traducir, un sistema de administración de proxectos de tradución, un conxunto de dicionarios bilingües acompañados de ferramentas de xestión das bases de datos léxicas e unha "memoria de tradución". A memoria de tradución é unha base de datos onde se almacenan a versión orixinal e traducida de cada unha das frases que traduce o usuario. Cando o usuario está a traducir unha frase, o programa detecta automaticamente se esa mesma frase ou outra frase similar xa foi traducida con anterioridade, de xeito que se poida reutiliza-la tradución sen necesidade de reescribirla completamente, facendo as modificacións que se consideren máis oportunas.

As aplicacións de THAO, xunto coas aplicacións de TSIA, acostuman a englobarse dentro do termo "tradución asistida por ordenador" (TAO), unha designación moito máis cerca da realidade actual da tradución có termo TA, que en rigor debería reservarse para a tradución totalmente automática, obxectivo último da investigación. Se, como escribiu o profesor Valentín García Yebra, o propósito dunha boa tradución é "decir todo lo que dice el original, no decir nada que el original no diga, y decirlo todo con la corrección y naturalidad que permite la lengua a la que se traduce"¹⁵, non hai dúbida de que á tradución automática aínda lle queda un longo camiño que percorrer. Pola súa parte, a TAO ofrece cada día novas utilidades informáticas que, ben aproveitadas, poden proporcionar importantes melloras ás condicións de traballo dos profesionais da tradución que saiban sacarles partido.

13. Para unha análise das características do programa de xestión de bases de datos terminolóxicos *MultiTerm* para Windows, véxase Gómez Guinovart, J., "MultiTerm: Terminología multilingüe en un entorno gráfico", *Information World en Español*, 33, 1995, pp. 12-14. Para unha descrición de *Term-PC*, véxase Camón, J. B., "Gestión de bancos de datos terminolóxicos", en Gómez Guinovart, J. (ed.), *Aplicaciones lingüísticas de la informática*, Tórculo, Santiago de Compostela, 1994, pp. 51-62. Véxase tamén Verhaest, F., "...valuation de cinq questionnaires de glossaires", *Terminologie et traduction*, 1, pp. 277-292.

14. Véxase Gómez Guinovart, J., "Herramientas para la traducción", *BYTE España*, 1, 1994, pp. 146-148, para unha presentación de *TranslationManager* para Windows; e Berry, M., "The Trados Translator's Workbench II", *Proceedings of the 33rd Annual Conference of the American Translators Association*, 1992, pp. 285-292, para unha descrición de *Translator's Workbench*.

15. García Yebra, V., *Teoría y práctica de la traducción*, Gredos, Madrid, 1982, p. 43.

