



MODELO MEMORIA **MOVING MINDS**

[DATOS nombre apellidos, institución de destino, país]

Ángela Almela Sánchez-Lafuente
Institute for Linguistic Evidence
Estados Unidos (Georgetown, DE)

Título: Investigación en lingüística forense computacional en español: Hacia una versión en español del método de identificación de autorías SynAID

RESUMEN sobre los resultados principales de la investigación/de formación (máx. 500 palabras)

Mediante el presente proyecto, se ha comenzado la adaptación de la herramienta de lingüística forense para identificación de autorías ALIAS SynAID a la lengua española, ante la creciente demanda de este tipo de servicios dirigida tanto a la solicitante como a la Dra. Chaski, contacto del centro receptor. Por tanto, el proyecto ha tenido como objetivo principal formalizar la marcadez sintáctica en español para el diseño de una versión adaptada del método ALIAS SynAID, desarrollado por la Dra. Chaski en lengua inglesa.

En primer lugar, la solicitante firmó un acuerdo de confidencialidad con el centro receptor al comienzo de la estancia, puesto que la adaptación del método sintáctico para identificación de autorías ALIAS SynAID al español suponía un exhaustivo estudio del método original en inglés protegido por patente, por lo que la Dra. Almela se comprometió a no hacer públicos aspectos científicos de los que no se habían realizado el registro de la propiedad intelectual y se encuentran sometidos al secreto industrial.

Como principal resultado de investigación, se han comprobado las categorías de marcadez sintáctica utilizadas en el método SynAID en lengua inglesa que funcionan también en lengua española. A continuación, se han codificado dichas categorías en ALIAS para la versión española de SynAID. El objetivo siguiente, en el que se continúa trabajando a distancia, es comprobar la precisión con que el método es capaz de asignar los textos a los autores correctos en lengua española.

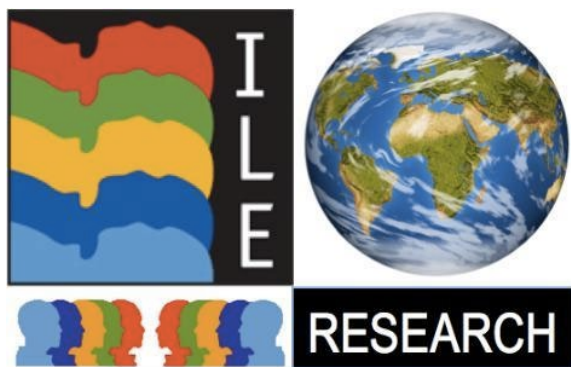
Para ello, se creó una base de datos que recoge una comparativa entre las categorías de marcadez sintáctica utilizadas en el método SynAID en lengua inglesa y lengua española. Para el diseño sintáctico de las estructuras en español, cada núcleo sintáctico se listó con todos sus posibles patrones; a continuación, dichos patrones se subcategorizaron en patrones marcados y no marcados considerando niveles de complejidad y otros factores de acuerdo con la teoría de la marcadez. Tal y como se



esperaba, se encontraron un número de similitudes entre los patrones ya codificados en el método SynAID para el inglés y los nuevos patrones para la lengua española, así como también fenómenos lingüísticos específicos del español que llevan más tiempo codificar. También se codificaron las categorías de marcadez sintáctica en español en ALIAS para la creación de la versión española de SynAID, adaptando los patrones de marcadez obtenidos en el paso anterior al lenguaje de programación Python para el trasvase de información a la herramienta ALIAS Technology LLC. El corpus de estudio lo recopiló la solicitante durante los años 2019-2021 en la Universidad de Murcia a través de la adaptación de un instrumento diseñado por la Dra. Chaski, y durante la estancia se ha digitalizado y subido a la base de datos de ALIAS. Además de ello, para la explotación y manejo de los datos, tanto la Dra. Almela como el Institute for Linguistic Evidence se han cumplido con todos los requisitos éticos para la investigación forense, esto es, anonimidad, confidencialidad, y preservación y acceso restringido a los datos.

Por último, cabe destacar que los resultados obtenidos durante la estancia se presentarán en la próxima edición del congreso anual de la American Association of Forensic Sciences (febrero 2023), y se enviarán para publicación en la revista científica Journal of Forensic Sciences.

[incluir imagen alusiva a la estancia investigadora/de formación]



Murcia, 23 de septiembre de 2022