



Lingüística forense en México y su uso en la identificación humana

Dra. Fernanda López-Escobedo
Fecha de publicación: 24 de febrero de 2021

Resumen

La identificación es una tarea que requiere del estudio de características humanas que sirvan para establecer que una persona es la misma que supone ser. Al igual que características como las lofoscópicas o dentales, las producciones lingüísticas también son susceptibles de análisis con un fin de identificación. El área de las ciencias forenses encargada de estudiar este tipo de características es la lingüística forense, a través del estudio del lenguaje evidencial o probatorio. En este trabajo se explica en qué consiste el trabajo que se realiza en esta área y los tres principales supuestos que busca probar: si el o la autora de una muestra dubitada, de autoría desconocida, es el o la autora de una muestra indubitada, de autoría conocida; si el o la autora de una muestra dubitada, es el o la autora de entre un grupo de muestras indubitadas; y el perfilamiento lingüístico cuando existe una muestra dubitada pero ningún sospechoso. La muestra dubitada puede ser una evidencia textual (un correo electrónico, un mensaje instantáneo o una carta) o una evidencia oral (una llamada de secuestro o de extorsión). En ambos tipos de evidencia es posible analizar diferentes características

lingüísticas con un fin de identificación forense.

Palabras clave: lingüística forense, lenguaje probatorio o evidencial, identificación humana, evidencia textual, evidencia oral.

Abstract

Human identification requires the study of human characteristics in order to establish the identity of a person. As with lofoscopic or dental characteristics, linguistic productions are also susceptible to being analyzed for human identification. Forensic linguistics is part of forensic sciences and studies the characteristics of language that can be used as evidence. This paper describes the main goals of forensic linguistics and the three main assumptions that could be proved: whether the author of a sample from an unknown source is the author of a reference sample; whether the author of a questioned sample is the author of a group of known samples; and the profile of a speaker when there is a questioned sample but no suspect. The evidence can be textual (an email, instant message, or letter) or orally produced (via a kidnapping or extortion call). In both

types of evidence different linguistic characteristics can be analyzed for forensic human identification.

Keywords: forensic linguistics, language as evidence, human identification, written evidence, speech evidence.

La identificación humana generalmente se asocia a la práctica forense que se lleva a cabo en la identificación de cadáveres. Por ejemplo, en el “Protocolo para el tratamiento e identificación forense” publicado en 2015 por la entonces Procuraduría General de la República (1), ahora Fiscalía, se detalla el procedimiento que deben seguir los especialistas que intervienen en un lugar ante el hallazgo de cadáveres. Entre las especialidades que se mencionan están la de: Antropología Forense, Criminalística de Campo, Video, Fotografía Forense, Medicina Forense, Dactiloscopia, Genética Forense y Odontología Forense. No es de extrañar que en un país como México en donde las cifras de desaparecidos aumentan día con día, el interés de la identificación humana se centre en identificar cadáveres. Según el Informe sobre fosas clandestinas y registro nacional de personas desaparecidas o no localizadas publicado el 6 de enero de 2020 por el Gobierno Federal, de 1124 cuerpos exhumados entre el 1 de diciembre de 2018 y el 31 de diciembre de 2019, solo el 35.4% ha sido identificado.

La identificación humana, como su término lo indica, consiste en establecer o demostrar que una persona o cosa es la misma que supone ser, es decir, que la identificación humana no se limita a la identificación de cadáveres (2). Cualquier característica humana que sirva para establecer que una persona es la misma que supone ser es susceptible de análisis. Y es ahí donde el análisis del lenguaje de un individuo resulta una herramienta útil en la identificación forense.

Una de las características del ser humano como ser racional, es su capacidad de comunicarse a través del lenguaje. Por medio del uso de signos lingüísticos es capaz de comunicar un mensaje y expresar una idea. Dicha producción lingüística, ya sea textual u oral, es el objeto de estudio de una de las

ramas de la lingüística forense. Esto debido a que una de sus premisas parte del hecho de que los individuos, como usuarios de una lengua determinada, tienen la capacidad de elegir entre una serie de formas lingüísticas que definen su propio estilo y, por lo tanto, un lenguaje propio (3).

El estudio de la lingüística se suele dividir en áreas de estudio: la fonética y la fonología, que tiene como objeto de estudio los sonidos del habla, conocidos como fonemas; la morfología que estudia la estructura interna de las palabras; la semántica, que se encarga del estudio del significado de las palabras; la sintaxis que se encarga del estudio del conjunto de reglas que regulan las combinaciones entre las palabras para producir oraciones; y la pragmática que estudia las intenciones comunicativas del hablante, el contexto y las circunstancias en que se produce el lenguaje.

El uso de todo este conocimiento lingüístico para un fin forense es lo que se conoce como lingüística forense, área que a su vez se divide en tres: el estudio del lenguaje legal, el estudio del lenguaje en el procedimiento judicial y el estudio del lenguaje probatorio o evidencial (4), esta última de interés para lo que se aborda en este texto pues busca identificar a un individuo por medio de una evidencia oral (por ejemplo: una llamada de teléfono, un mensaje de voz, etc.) o por medio de una evidencia textual (un correo electrónico, una carta, una declaración, etc.).

En México, la lingüística forense es un área de reciente desarrollo, comparado con el que ha tenido en países anglosajones. En países de habla hispana, España es pionera en el desarrollo de investigaciones en esta área. La propuesta de técnicas y métodos desarrollados por la lingüística forense se ha enfocado en su aplicación en lenguas como el inglés, el alemán y en mucho menor medida el español.

Al ser la lingüística forense una disciplina que tiene como principal objeto de análisis una producción lingüística, el desarrollo de técnicas para otras lenguas distintas al español resulta un factor a tomar en cuenta. Por un lado, las variaciones fonéticas de las diferentes lenguas del mundo –desde el número de

sonidos con los que cuenta cada una, hasta la entonación- son aspectos que inciden a la hora de realizar una identificación a partir de una muestra de voz. Por otro lado, las diferencias en la sintaxis de las lenguas dan como resultado textos escritos con estructuras más o menos complejas y más o menos flexibles y, por lo tanto, con distintas metodologías para ser abordados. Es por esta razón que el desarrollo de la investigación en lingüística forense en México es fundamental para lograr métodos y técnicas específicas para el dialecto mexicano, lo que necesariamente conlleva a resultados más confiables.

Ahora bien, ¿Cómo la lingüística forense contribuye en la identificación humana? Tanto si partimos de una evidencia oral como de una evidencia textual, los supuestos que nos podría interesar probar son:

si el o la autora de una muestra dubitada, de autoría desconocida, es el o la autora de una muestra indubitada, de autoría conocida;
si el o la autora de una muestra dubitada, es el o la autora de entre un grupo de muestras indubitadas;
cuando tenemos una muestra dubitada pero ningún sospechoso y queremos obtener información sobre las características sociolingüísticas del autor;

En el primer supuesto, la metodología consistirá en analizar las características acústicas y/o lingüísticas presentes en la muestra dubitada y determinar su similitud respecto a las mismas características presentes en la muestra indubitada. En el caso de una evidencia oral es posible analizar, tanto características acústicas, como características lingüísticas. Por ejemplo: el valor promedio del tono fundamental en la vocal /a/ y/o la nasalización de vocales. En el caso de una evidencia textual, la comparación se centrará en el análisis de características semánticas y sintácticas, principalmente. Por ejemplo, la frecuencia relativa de palabras funcionales y de contenido en uno y otro texto o la frecuencia relativa de oraciones subordinadas sustantivas. Con base en la presencia o ausencia de este tipo de características y con el uso de medidas de tipo estadístico es posible definir un porcentaje de similitud entre

la muestra dubitada y la muestra indubitada. Es así como un gran número de herramientas computacionales llevan a cabo comparaciones automatizadas para identificar la autoría de una producción lingüística.

La principal diferencia que se presenta entre un análisis manual y un análisis automatizado en lingüística forense es el número de características que se pueden analizar al mismo tiempo. Mientras que en el primer caso el análisis se hace de una en una, y es posible analizar algunas características que solo el oído o la lectura entrenada de un experto puede percibir; en los análisis automatizados es posible comprar un gran número de características de manera simultánea, y además comprarlas no solo con una muestra indubitada sino con una cantidad ilimitada de muestras indubitadas. Y es aquí donde llegamos al segundo supuesto, donde la comparación ya no es del tipo 1:1 sino respecto a un conjunto de muestras indubitadas. Este conjunto puede llegar a ser tan grande como la comparación que se realiza entre una muestra dubitada y una base de datos nacional. Aquí, el análisis manual se vuelve casi imposible y por esa razón se suelen utilizar softwares que realizan la comparación de características de manera automática. Tanto para voz como para texto existen diferentes softwares comerciales que ofrecen, por un precio en ocasiones bastante elevado, la posibilidad de realizar dichas comparaciones. En los últimos años también es posible utilizar herramientas de código abierto pero que necesariamente requieren del conocimiento de un experto que sea capaz de modificar el código según las necesidades de cada caso.

Es importante mencionar que en este supuesto el desarrollo de bases de datos resulta fundamental para conocer la frecuencia de ocurrencia de una característica, ya sea acústica o lingüística, en una población determinada. Y aquí la población que nos interesará será la de hablantes del español mexicano, donde la frecuencia de una característica forzosamente variará respecto a otra lengua o a otro dialecto del español. Las características acústicas del ruso no presentarán la misma frecuencia de aparición que las del español, por ejemplo; y las carac-

terísticas sintácticas del español argentino tampoco presentarán la misma frecuencia de ocurrencia que las características del español mexicano. De aquí la importancia de llevar a cabo estudios para la variante mexicana.

Además, la posibilidad de calcular la frecuencia de aparición de una característica lingüística permite generar resultados numéricos como el Likelihood Ratio (LR), comúnmente utilizado en diferentes áreas de la Ciencia Forense.

El último de los supuestos es el que se conoce como perfilamiento lingüístico pues se cuenta con una muestra dubitada pero ningún sospechoso. En este supuesto, el único recurso que se tiene es el contenido lingüístico, ya sea oral o escrito, de la muestra dubitada. La información contenida en ella puede, por medio de un experto en lingüística forense, orientar sobre la procedencia, la edad, y el género del autor, entre otras características sociolingüísticas. Un ejemplo de ello es el famoso caso del reportero James Foley, quien fue secuestrado en el norte de Siria en 2012 mientras cubría el levantamiento contra el gobierno de Bashar al-Ásad. En 2014, un video es dado a conocer a los medios en donde un hombre totalmente cubierto del rostro emite una nueva amenaza y decapita al fotoperiodista James Foley. La evidencia apuntaba a un militante yihadista enmascarado del Estado Islámico en Irak y tras el análisis de su discurso, un grupo de expertos concluye que el autor del discurso es, además, un individuo de origen británico con acento tipo cockney propio de una persona de origen extranjero criada en Londres (5).

Es así como la lingüística forense contribuye en el área de identificación humana analizando un rasgo exclusivamente humano y con una riqueza de información aún del todo desconocida para la ciencia forense.

Glosario

Características sociolingüísticas: cualidades que dan cuenta de aspectos sociales que influyen en el uso de la lengua.

Likelihood Ratio (LR): cálculo estadístico que resume en un cociente el respaldo relativo que proporciona una prueba a dos hipótesis. Cuando la razón de verosimilitud es de 1, las pruebas disponibles ofrecen el mismo respaldo a ambas hipótesis, es decir, tanto a la de la fiscalía como a la de la defensa.

Nasalización de vocales: fenómeno fonético en el que un sonido que es oral se pronuncia con el velo del paladar abajo, lo que impide que el flujo de aire salga por la boca y la mayoría escapa a través de la nariz.

Oraciones subordinadas sustantivas: proposiciones que dependen jerárquicamente de una oración principal y que funcionan como sustantivos dentro de la oración.

Tono fundamental: onda sonora más baja de la onda sonora compleja y que corresponde a la frecuencia de abertura y cierre de las cuerdas vocales.

Palabras funcionales: palabras que tienen un significado semántico como los verbos, los sustantivos y los adjetivos.

Palabras de contenido: palabras que no tienen contenido semántico, como los artículos, las conjunciones y las preposiciones.

Referencias

1. Procuraduría General de la República. Protocolo para el tratamiento e identificación forense. 2015. México
2. López-Escobedo F, Castillo ZG, Sanchez JH, Olvera CP. En búsqueda de un lenguaje común entre la ciencia y el derecho: el caso de la ciencia forense en México. Debate Terminológico. ISSN: 1813-1867. 2017(17):2-17.
3. Coulthard M. Author identification, idiolect, and linguistic uniqueness. Applied linguistics. 2004 Dec 1;25(4):431-47.
4. Gibbons J, Turell MT, editors. Dimensions of forensic linguistics. John Benjamins Publishing; 2008 Nov 21.
5. Queralt S. El análisis lingüístico: una ciencia forense. 2017. Recuperado: <https://eapc-rld.blog.gencat.cat/2017/07/06/el-analisis-linguistico-una-ciencia-foren->



Dra. Fernanda López-Escobedo. Es licenciada en Lengua y Literaturas Hispánicas por la UNAM, Maestra en Lingüística Aplicada y Doctora en Ciencias del Lenguaje y Lingüística Aplicada por la Universidad Pompeu Fabra de Barcelona, España. Cuenta con una especialidad en estadística aplicada por el Instituto de Investigaciones en Matemáticas Aplicadas y Sistemas (IIMAS) de la UNAM. Actualmente cursa la licenciatura en Derecho en el Sistema de Universidad Abierta de la UNAM. Ha participado en proyectos relacionados con las tecnologías del lenguaje y la lingüística forense en España y en México. Asimismo, ha organizado diversos eventos académicos dentro de la UNAM y ha participado en conferencias y cursos a nivel nacional e internacional. Cuenta con publicaciones académicas y de divulgación en el ámbito de las tecnologías del lenguaje y la lingüística forense. Actualmente es académica de tiempo completo en la Licenciatura en Ciencia Forense de la UNAM, colaboradora del Grupo de Ingeniería Lingüística del Instituto de Ingeniería de la UNAM, miembro fundadora de la Asociación Mexicana de Procesamiento del Lenguaje Natural y miembro del Sistema Nacional de Investigadores Nivel I. Correo electrónico: flopeze@unam.mx

IDENTIFICACIÓN humana.mx

“Fortalecimiento del Estado de Derecho en México II (FED II)” de la
Cooperación Alemana al Desarrollo Sustentable (GIZ) GmbH
Av. Amatlan 149, Col. Hipódromo Condesa, Del. Cuauhtémoc, C.P. 06000
Ciudad de México, México
E estadodederecho@giz.de
I www.giz.de/mexico-mx
I www.giz.de/en/worldwide/79202.html
Tw https://twitter.com/FED_GIZMX

Diseño editorial : FED II, mayo de 2021, Ciudad de México, México
Diseño de página web: CEOS New Media Agency

Créditos fotográficos
iStock.com/peterschreiber.media

Texto
Dra. Fernanda López-Escobedo