

# Propuesta de clasificación de un banco de voces con fines de identificación forense

## A proposal to classify a voice database for forensic identification

Fernanda López-Escobedo  
Universidad Nacional Autónoma de México  
flopeze@unam.mx

Julián Solórzano-Soto  
Universidad Nacional Autónoma de México  
jsolorzanos@iingen.unam.mx

### Resumen

---

En este artículo se presenta el proyecto que se desarrolla para proponer una clasificación de un banco de voces con fines de identificación forense. Se expone la manera en que la información lingüística puede ser utilizada en una base de datos para reducir el número de falsos positivos y falsos negativos que resultan cuando se llevan a cabo comparaciones automatizadas para la identificación forense de voz. En particular, se abordan los fenómenos fonéticos que se han propuesto para realizar una clasificación de un banco de voces en este nivel de la lengua. A partir de esta información se describe cómo construir un modelo de base de datos y el tipo de búsquedas que se espera lograr.

La propuesta de generar descriptores lingüísticos para la clasificación de un banco de voces pretende ser una metodología que permita coadyuvar en la impartición de justicia en México y otros países de habla hispana.

### Palabras clave

---

sociolingüística, identificación forense de voz, variación, banco de voces, español

### Abstract

---

This article describes the project developed to classify a voice database for forensic identification. It exposes how the linguistic information can be used in a database to reduce the number of false positives and false negatives when an automatized forensic voice comparison is carried out. In particular, the phonetic phenomena that have been proposed for classifying a voice database in this level of language are reported. From this information it describes how to build a database model and the type of search to be achieved.

The proposal to create linguistic descriptors to classify a voice database is intended as a methodology to assist in the administration of justice in Mexico and other Spanish-speaking countries.

### Keywords

---

sociolinguistics, forensic voice identification, variation, voice database, spanish

### 1 Introducción

---

Actualmente, los laboratorios de análisis forense de voz de los servicios periciales en México cuentan con bancos de voces que están organizados sin tomar en cuenta las variantes lingüísticas de cada hablante. Las voces se clasifican únicamente por género y número de averiguación previa o carpeta de investigación.

Por ejemplo, Aguilar (2011), quien fue Director General de la Coordinación de Servicios Periciales de la la Procuraduría General de la República (PGR) por más de 15 años, menciona que dicha institución cuenta con cuatro bases de datos: huellas dactilares, identificación balística, genética forense y análisis de voz. Respecto a esta última comenta que el reconocimiento se realiza de manera automática con base en las características acústicas de la persona a partir de ondas sonoras, pero no hace referencia a ningún aspecto lingüístico de la voz.

Recientemente, Morrison et al. (2016) realizó una encuesta para conocer cómo las agencias criminales de diferentes partes del mundo llevan a cabo el reconocimiento forense de voz. Uno de los resultados que reporta es que los sistemas automáticos de comparación forense de voz que más se utilizan son BATVOX, seguido de IKAR Lab. En México, las agencias de investigación que cuentan con análisis forense de voz utilizan alguno de estos dos sistemas. Por ejemplo, la PGR cuenta con el sistema BATVOX, mientras que la Procuraduría General de Justicia de la Ciudad de México cuenta con el sistema IKAR Lab.

Ambos sistemas tienen la capacidad de clasificar las muestras de voz asociándoles determinados atributos, lo que facilitaría la aplicación de los resultados obtenidos en este proyecto. No se cuenta con documentación que señale la manera en que las instituciones mencionadas utilizan las ventajas que ofrecen estos sistemas para potenciar las bases de datos con las que cuentan, pues

se trata de información confidencial. Sin embargo, los expertos que en ellas laboran expresan que las muestras de voz no se clasifican a partir de información lingüística. Prueba de ello es que los peritajes que realizan no contienen una descripción lingüística de la voz que se está analizando, y no existen expertos lingüistas laborando en dichas instituciones. El perfil del perito que realiza análisis de voz en México es el de un ingeniero experto en análisis de audio.

Morrison et al. (2016) añade que de las 44 agencias que respondieron la encuesta, 28 cuentan con una base de datos. Agrega que la mayoría contienen grabaciones de sospechosos pero pocas (menos de la mitad) cuentan además con voces de personas condenadas o procesadas por algún delito, así como de grabaciones de la población. Dicho autor también reporta las características técnicas con las que se realizan las grabaciones que están contenidas en las bases de datos (llamadas telefónicas, entrevistas con sospechosos, audios de internet, etc.), el número de lenguas incluidas y el número de hablantes. Sin embargo, no se menciona nada respecto a la clasificación de estas bases de datos a partir de información social o lingüística.

## 2 Reconocimiento forense de voz

Una de las aplicaciones del área de reconocimiento automático de hablantes es la del reconocimiento en el contexto forense. El método que se utiliza consiste en comparar una muestra de voz de un hablante desconocido (dubitada) contra un conjunto de referencia compuesto por muestras de voz de hablantes conocidos (indubitadas); es decir, una base de datos. No obstante, su aplicación en el ámbito forense presenta situaciones complejas. Rose (2002) hace la diferencia entre una base de datos que representa un conjunto de referencia cerrado y una que representa un conjunto abierto. La diferencia se encuentra en que, mientras en el conjunto cerrado se sabe que el hablante de la muestra dubitada es uno de los hablantes del conjunto de referencia, en el conjunto abierto el hablante de la muestra dubitada puede o no estar en el conjunto de referencia. La primera situación representa una tarea más sencilla para el reconocimiento pues es posible determinar que el hablante de la muestra dubitada será el que presente la distancia más pequeña respecto al conjunto de muestras indubitadas. En el ámbito forense esta situación será poco común. Los sistemas automáticos de comparación forense de voz cuentan con un conjunto abierto de muestras de voz indubitadas. Incluso, algunas veces parte de

estas muestras de voz son también dubitadas (por ejemplo, llamadas telefónicas interceptadas). En este último caso, si al realizar la comparación la muestra dubitada resulta similar a una muestra también dubitada del conjunto de referencia, el resultado es útil para reunir información forense de inteligencia pero no para identificar a un hablante, como lo señala Ramos (2007).

Otra característica del reconocimiento automático de hablantes es la dependencia del texto. Rose (2002) menciona que existen sistemas que son dependientes del texto y sistemas que no lo son. Nuevamente, la segunda situación resulta más compleja y es la que se presenta en el ámbito forense. Por ejemplo, las diferencias acústicas entre dos muestras de voz de un mismo hablante diciendo la palabra “mesa”, serán menores a las que presenta en dos muestras de voz produciendo palabras diferentes “mesa” y “silla”.

Ramos (2007) menciona que la producción de habla es un proceso complejo pues depende de muchas variables que incluyen factores sociolingüísticos como nivel educativo, el contexto lingüístico, diferencias dialectales, así como cuestiones fisiológicas. El estado del arte en los sistemas automáticos de comparación forense de voz o reconocimiento de hablantes se ha concentrado en desarrollar sistemas que analizan la señal acústica para realizar la comparación entre la muestra de voz dubitada y el conjunto de referencia. Dichos sistemas se caracterizan por realizar la comparación contra un conjunto de referencia abierto y por realizar el reconocimiento independientemente del texto. No obstante, y a pesar de que la señal acústica proporciona información sobre distintos factores, sus características están directamente relacionadas con cuestiones fisiológicas acerca de la configuración del tracto vocal.

El proyecto que aquí se presenta pretende incluir información sobre diferencias dialectales con el fin de robustecer los sistemas automáticos de comparación forense de voz. Actualmente no existe un análisis lingüístico previo que permita realizar la confronta únicamente contra una parte del conjunto de referencia que se caracterice por contener voces que compartan características lingüísticas con la muestra de voz dubitada. Contar con información que permita generar grupos en el conjunto de referencia de muestras de voz reduciría el número de falsos positivos y falsos negativos que se generan con las confrontas o comparaciones automatizadas.

Por otro lado, debido a que en muchas ocasiones las grabaciones dubitadas se encuentran en formatos digitales comprimidos, las propiedades acústicas se ven muy afectadas, lo que impi-

de llevar a cabo la comparación. La información sobre las variantes lingüísticas que caracterizan las distintas voces contribuiría no solamente a la clasificación del conjunto de referencia, sino también a reforzar los dictámenes que actualmente se realizan pues se podrían llevar a cabo comparaciones no automatizadas basadas en parámetros lingüísticos, aún y cuando las características del formato de audio no lo permitan.

El objetivo de este trabajo consiste en describir los parámetros lingüísticos que se han definido en el nivel fonético para la creación de catálogos de descriptores lingüísticos en un banco de voces. Es importante resaltar que el objetivo de este proyecto es generar el conocimiento para clasificar un banco de voces que sirve como conjunto de referencia en una comparación automatizada, y no así generar el banco de voces, pues como ya se mencionó, las agencias de investigación en México ya cuentan con grandes bases de datos de voz.

### 3 División dialectal del español

Desde 1918 con la publicación del *Manual de Pronunciación española* de Navarro hasta el día de hoy, los estudios lingüísticos sobre la pronunciación del español han tenido grandes avances. El *Manual de Pronunciación española* tenía por objeto “describir breve y sencillamente la pronunciación española” pues “sabido es que la lengua española presenta importantes diferencias de pronunciación, no sólo entre los diversos países en que se habla, sino entre las regiones de un mismo país, y frecuentemente entre las comarcas y lugares de una misma región” (1918, pg. 5).

Moreno Fernández (1993, pg. 15) señala que “los individuos, al hablar entre sí, son capaces de distinguir a los que pertenecen a su misma comunidad de los que son ajenos a ella. Los límites de una comunidad pueden ser locales, regionales, nacionales o incluso supranacionales y sus miembros generalmente conocen o intuyen el alcance de la conducta lingüística que los caracteriza. Esto nos llevaría al concepto de dialecto: los hablantes pueden sentirse miembros de una comunidad dialectal”. Sin embargo, “aunque una persona tenga conciencia de su pertenencia a una comunidad, también es capaz de identificar dentro de ella variantes internas de carácter geolingüístico o sociolingüístico, así como de reconocer cuáles son los usos más prestigiosos de su variedad y apreciar la relación histórica de su habla con otra”.

El primer lingüista en proponer una división dialectal de México fue Henríquez Ureña en

1921. Propone seis grandes zonas dialectales que divide en:

- El territorio hispánico de los Estados Unidos,
- El norte de la República Mexicana,
- La altiplanicie del Centro, donde se ubica la Ciudad de México,
- Las tierras calientes de la costa oriental, en particular Veracruz y Tabasco,
- La península de Yucatán, donde ejerce influencia el maya,
- La América Central, comenzando en el Estado mexicano de Chiapas.

“Lo que a juicio de Henríquez Ureña caracteriza como zona dialectal a cada una de las regiones geográficas enumeradas es el vocabulario, pues reconoce que no hay uniformidad fonética en ninguna de ellas” (Moreno de Alba, 1998, pg. 157). La principal crítica que se le hace a Henríquez Ureña es que se trata de una clasificación basada únicamente en cuestiones léxicas. En este sentido resulta pertinente resaltar que para determinar una zona dialectal se deben tomar en cuenta las características presentes en diferentes niveles: el nivel fonético que estudiará los sonidos de la lengua; el nivel morfológico que estudiará la estructura de las palabras, su constitución interna y sus variaciones; el nivel sintáctico que estudiará la manera en la que se combinan las palabras; y el nivel léxico que estudiará el uso de determinados vocablos. En este trabajo se presentan únicamente los descriptores lingüísticos correspondientes al nivel fonético; no obstante, el objetivo del proyecto es proponer la clasificación de una base de datos con descriptores en todos los niveles, sirviendo el nivel fonético como ejemplo de lo que se espera lograr.

Al trabajo de Henríquez Ureña siguieron otros como *La pronunciación del español de América: Ensayo histórico-descriptivo* de Canfield (1962), *El problema de la división del español americano en zonas dialectales* de Rona (1964) y *Dialectología hispanoamericana: Teoría, descripción, historia* de Zamora & Guitart (1982). Además, se llevaron a cabo investigaciones cuyo objetivo era describir el habla de algunas zonas de la República Mexicana, por ejemplo el trabajo de Matluck (1951) sobre *La pronunciación en el español del Valle de México, Fonología del español hablado en la Ciudad de México: Ensayo de un método sociolingüístico* de Perissinotto (1975), *El habla de Tepetzotlán* de Cortichs de Mora (1951), *El habla de Guanajuato* de Boyd-Bowman (1960),

*El español de Jalisco* de Cárdenas (1967), entre otros.

Sin embargo, el estudio más completo que se ha hecho sobre el español de México es el que comienza el lingüista *Lope Blanch* (1970) con el objetivo de “iniciar una serie de investigaciones, conducentes a reunir los datos lingüísticos –fonéticos, gramaticales y léxicos– necesarios para determinar cuáles son las principales modalidades dialectales existentes hoy en el país” (*Lope Blanch*, 1970, pg. 4). “Los datos recogidos presentaron gran abundancia, variedad y riqueza y en consecuencia el proyecto inicial de delimitar las zonas dialectales fue superado y se transformó en el levantamiento de un Atlas general del español de México” (*Espejo*, 1998, pg. 119). Se estudiaron 193 localidades y se analizaron al menos 3 informantes por localidad, de distinto nivel sociocultural, edad y sexo. Para el levantamiento de los datos se realizaron cuestionarios y entrevistas, algunas de las cuales se grabaron en cintas magnetofónicas. La información recabada se registró en mapas que dan cuenta de la variedad lingüística del español de México.

A pesar de que los datos recopilados en dicho trabajo resultan de gran utilidad en el estudio de las variantes dialectales del español de México, los años han transcurrido y con ello se debe asumir un cambio lingüístico que probablemente ya no esté reflejado en los datos recopilados hace casi más de cuarenta años.

En el trabajo que aquí se presenta se realizó una revisión bibliográfica de los fenómenos fonéticos que se han estudiado, principalmente en el español de México, y de un modo más general en el español de América y en el español peninsular, con el fin de determinar cuáles resultan pertinentes para la clasificación del banco de datos. Una de las principales referencias ha sido el *Atlas lingüístico de México*, sin embargo somos conscientes de la necesidad de corroborar los datos a través del levantamiento de grabaciones en las diferentes regiones a las que se hace referencia.

### 3.1 Descriptores fonéticos

Los fenómenos que se proponen como descriptores lingüísticos en el nivel fonético son 17:

- Articulación apicoalveolar de /s/
- Articulación fricativa de /tʃ/
- Asibilación de róticas
- Aspiración de /s/ implosiva
- Aspiración de /x/

- Aspiración y/o velarización de /f/
- Bilabialización de /f/
- Confusión de líquidas
- Debilitamiento y/o pérdida de /d/ intervocálica y final de sílaba
- Nasalización vocálica
- Realización fricativa de /j/
- Seseo y ceceo
- Variante oclusiva de /b, d, g/ en posición intervocálica
- Velarización de nasal implosiva
- Velarización de róticas
- Vocalización de líquidas
- Yeísmo

Resulta interesante resaltar que en el caso de los fenómenos que afectan a las vocales únicamente se ha incluido el de la nasalización. A pesar de que existen otros fenómenos que se han estudiado ampliamente como la diptongación de hiatos, el debilitamiento vocálico o el cierre vocálico, son fenómenos que podrían resultar complejos en el análisis auditivo hecho por un perito en acústica forense.

Este último criterio ha resultado fundamental en la elección de los fenómenos pues el objetivo es clasificar un banco de voces con fines forenses. Por lo tanto, los usuarios de dicho banco, y a su vez responsables de clasificar las voces, no serán expertos en lingüística sino peritos en acústica forense que, en la mayoría de los casos, tienen una formación en el área de ingeniería. Así, los fenómenos fonéticos que servirán como descriptores lingüísticos deben ser fácilmente identificables en un análisis auditivo por una persona no experta en fonética, condición que consideramos cumplen los 17 fenómenos que aquí se presentan.

### 3.2 Información diatópica y diastrática de los fenómenos

Una vez identificados los fenómenos se elaboró una breve descripción de su realización y se hizo una tabla de cada uno de ellos con información sobre variantes diatópicas y/o diastráticas según se ha descrito en las diferentes referencias bibliográficas. Consideramos importante que los fenómenos describan algunas de estas variantes o ambas pues es la manera en la que se puede caracterizar un individuo no sólo para clasificar su voz en la base de datos, sino también para poder realizar una confronta lingüística, en el caso

Variante diatópica	Variante diastrática	Notas
México: es más frecuente en el noroeste del país; en menor medida, se registra en la región occidental del centro de México (Moreno de Alba, 1994; Lope Blanch, 1990)	Está presente en todos los niveles sociales pero es más frecuente en las mujeres (Lope Blanch, 1990)	
España (Andalucía)	Propia del nivel sociocultural bajo (Jiménez, 1999)	En Cádiz coexisten tanto la africada como la fricativa; sin embargo, entre los hablantes cultos, la segunda es rechazada (Jiménez, 1999)
Chile	Propia del nivel sociocultural bajo (Vivanco, 1999)	Motivó el surgimiento de una nueva variante (Vivanco, 1999)
Panamá: en la capital se registra la variante totalmente fricativa (Aleza & Enguita, 2002)	Es más frecuente en el nivel medio; y parece ser una marca distintiva de sexo (Aleza & Enguita, 2002)	
Puerto Rico: en el área metropolitana se registra la variante fricativa (López, 1983; Navarro, 1948)	Es más frecuente en las mujeres (López, 1983)	Algunos estudios indican que es un fenómeno propio de las generaciones jóvenes (Cedergren, 1973); y otros aseguran lo contrario (López, 1983)

Cuadro 1: Descripción del fenómeno articulación fricativa de /tʃ/

de que las propiedades acústicas de la grabación estén muy afectadas.

Por ejemplo, en el caso del fenómeno “articulación fricativa de /tʃ/” se cuenta con una descripción general que se narra de la siguiente manera: el español cuenta con un solo fonema africado, el linguopalatal sordo /tʃ/. Este se caracteriza por su articulación doble: oclusión + fricción. Generalmente, el período de oclusión es mayor que el de la fricción; sin embargo, hay ciertas áreas dialectales donde se registran variantes (Quilis, 1999). “En la medida en que prevalezca el momento oclusivo sobre el fricativo se tratará de un sonido más tenso; si pasa lo contrario y el momento fricativo predomina, la articulación será menos tensa, más relajada” (Moreno de Alba, 1994, pg. 193).

Respecto a la caracterización diatópica y diastrática del fenómeno se cuenta con la siguiente información. En México este fenómeno se atribuye como característico del habla del noroeste del país: Baja California, Baja California Sur, Sonora, Chihuahua y una parte de Durango (Mendoza, 2004). No obstante, también se registra como poco frecuente en otras regiones de México, por ejemplo: Tacámbaro, Michoacán; Cihuatlán, Jalisco; Tepic y Acaponeta, Nayarit; la región norte de Veracruz (Moreno de Alba, 1994), entre otros.

En el Atlas lingüístico de México (mapa 35) el fenómeno tiene el mayor índice en Baja California Sur (60%) y Guaymas, Sonora (50%); mientras que el resto de los estados del noroeste

presentan un índice menor: Baja California Norte (7.5%), Chihuahua (20%), Durango (5%), etc. Otras localidades de la costa del Pacífico registran lo siguiente: Tepic (7.5%), Tuxpan (30%) y Acaponeta (30%).

Aproximadamente, el 75% de los informantes que registraron la /tʃ/ fricativa fueron mujeres de todos los niveles escolares: analfabeto, semianalfabeto, medio, semiculto y culto.

Asimismo, se cuenta con información sobre este fenómeno en España y otras partes de América Latina. En Andalucía se registra el fenómeno de fricativización conforme se desciende en la escala social o se reduce la edad de los hablantes. Es decir: la articulación africada es más frecuente y prestigiosa. Por lo tanto, penetra en los estratos altos y medios (94% y 90%, respectivamente); mientras que en el nivel sociocultural bajo solo alcanza 68% (Jiménez, 1999). Asimismo, la variante en cuestión es una marca que distingue entre sexos: el alófono fricativo es casi exclusivo del habla masculina (Moya & Weiderman, 1995). En Cádiz coexisten tanto la africada como la fricativa y su índice de aceptación es similar; sin embargo, entre los hablantes cultos la segunda realización no es aceptada.

En Chile, la articulación fricativa es propia de los niveles socioculturales bajos. Por ende, ha sido ampliamente estigmatizada. Este hecho motivó el surgimiento de una nueva variante, como respuesta al rechazo del alófono fricativo. La variante emergente se caracteriza por ser una realización muy marcada de la africada: está ligera-

mente más adelantada que  $/tʃ/$ , tiene un periodo de oclusión más extenso (Vivanco, 1999).

En Panamá, en la capital se registra la variante totalmente fricativa (Vaquero, 1996). Los estudios concluyen que los hombres adoptan, en mayor medida que las mujeres, la articulación africana. Es decir: son las mujeres quienes tienden más a fricativizar el fonema (39.5%) que los hombres (13.2%) (Cardona, 2010). Asimismo, la articulación fricativa es más común en la clase media que en otras clases sociales (Aleza & Enguita, 2002).

En Puerto Rico, en el área metropolitana se presenta la variante adherente (con predominio del elemento fricativo) (López, 1983; Navarro, 1948). Esta variante es más frecuente en el sexo femenino, como ocurre en Panamá; asimismo, el fenómeno no es propio de las generaciones jóvenes (López, 1983).

Una vez descrita la información fonética del fenómeno, ésta se acomoda en una tabla como la que se muestra en el Cuadro 1. A partir de la información contenida en la tabla se definen los atributos de identidad en la base de datos, así como la asociación que existirá entre las diferentes entidades.

## 4 Diseño de la base de datos

El diseño de la base de datos debe responder a la necesidad de clasificar las grabaciones del banco de voces de manera que éstas puedan ser recuperadas de acuerdo a una consulta que incluya los descriptores mencionados anteriormente junto con su información diatópica y diastrática.

El análisis lingüístico previo de la voz dubitada debe servir para reducir el conjunto de referencia a un grupo de muestras de voz que comparten características lingüísticas. Por lo tanto, debe ser posible hacer una consulta que permita especificar los descriptores lingüísticos presentes en la grabación dubitada, para así restringir el conjunto de referencia al hacer la comparación automatizada.

Como cada descriptor lingüístico en la base de datos estará acompañado de su información diatópica y diastrática, las grabaciones que finalmente serán mostradas al usuario serán aquellas que contengan todos los descriptores solicitados y además dichos descriptores presenten similitud en sus atributos tales como país, región, género y nivel social.

### 4.1 Modelo entidad-relación

En la figura 1 se muestra el modelo entidad relación propuesto para el diseño de la base de

datos que cumple con las características anteriores.

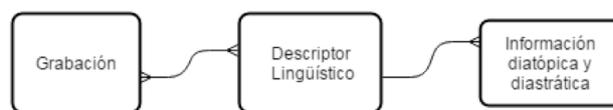


Figura 1: Diagrama Entidad-Relación.

En este modelo se tiene que una grabación puede estar asociada a uno o más descriptores lingüísticos, mientras que a su vez un descriptor lingüístico puede estar asociado a varias grabaciones (relación muchos a muchos). De esta manera se puede tener que, por ejemplo, la grabación A presente “nasalización vocálica” y “articulación fricativa de  $/tʃ/$ ”, mientras que la grabación B presente “articulación fricativa de  $/tʃ/$ ” y “vocalización de líquidas”.

Por su parte, cada descriptor lingüístico deberá tener asociada su respectiva información diatópica y diastrática. Esta información está representada por la entidad identificada en la figura como “información diatópica y diastrática”, la cual consiste en una combinación de valores para atributos previamente definidos (ej. país, región, género y nivel social).

Siguiendo las buenas prácticas de un modelo relacional, es recomendable que el conjunto de países, regiones, géneros y niveles sociales sea también representado cada uno por una entidad, de manera que la combinación de valores de estos atributos sean referencias a un catálogo.

### 4.2 Tablas de la base de datos

A continuación se presenta la propuesta de la estructura de las tablas de la base de datos a nivel más concreto que el diagrama entidad-relación.

Primeramente, en el Cuadro 2 se muestra la información que funge como catálogo: país, región, género y nivel social. Las tablas país y región están ligadas puesto que una región debe pertenecer necesariamente a un país. Esto hace posible que en la tabla de la información diatópica y diastrática, baste con hacer referencia a la región. Todas las entradas de esta tabla son asignadas con un ID único de manera que pueden ser identificadas unívocamente en las demás tablas.

En el Cuadro 3 se muestra la tabla que representa a la entidad “descriptor lingüístico”. En ella se asigna un ID para hacer referencia a cada descriptor. También es posible, de considerarse conveniente, agregar un tercer campo con una descripción del fenómeno, para que esta información pueda ser mostrada al usuario del sistema al

género		nivel	
<i>id</i>	<i>nombre</i>	<i>id</i>	<i>nombre</i>
1	H	1	Bajo
2	M	2	Medio
		3	Alto

país		región		
<i>id</i>	<i>nombre</i>	<i>id</i>	<i>pais_id</i>	<i>nombre</i>
1	México	1	1	Noroeste
2	Panamá	2	1	Region occidental del centro
3	Chile	3	5	Andalucía
4	Puerto Rico	4	2	Capital
5	España	5	4	Área metropolitana

Cuadro 2: Tablas catálogo de la base de datos.

descriptor	
<i>id</i>	<i>nombre</i>
1	Articulación fricativa de /tʃ/
2	Nasalización vocálica
3	Debilitamiento y pérdida de /d/ intervocálica y final de sílaba
4	Confusión de líquidas
5	Aspiración y velarización de /f/
6	Bilabialización de /f/

Cuadro 3: Tabla de los descriptores lingüísticos.

información_día				
<i>id</i>	<i>descriptor_id</i>	<i>region_id</i>	<i>genero_id</i>	<i>nivel_id</i>
1	1	1	2	1
2	1	1	2	2
3	1	1	2	3
4	1	2	2	1
5	1	2	2	2
6	1	2	2	3
7	1	3	n/a	1
8	1	4	n/a	2
9	1	5	2	n/a

Cuadro 4: Tabla de la información diatópica y diastrática.

momento de ser utilizado.

Finalmente, en el Cuadro 4 se muestra la tabla referente a la entidad “información diatópica y diastrática”. Esta tabla contiene un registro por cada combinación de atributos que pueda presentar un descriptor lingüístico. El Cuadro 4 presenta cómo quedaría la información contenida en el Cuadro 1 almacenada dentro de la base de datos. Cada una de sus filas representa una combinación de atributos para un fenómeno. Por ejemplo, la primera fila tiene los valores (1, 1, 2, 1), los cuales representan el fenómeno lingüístico, la región, el género y el nivel social. Esta fila representa la información sobre el descriptor 1 (correspondiente a la “articulación fricativa de /tʃ/” según el ID que aparece en la tabla de descrip-

tores lingüísticos), que está presente en la región 1 (correspondiente a la región Noroeste de México), en el género 2 (mujeres), y en el nivel social 1 (bajo). Aparecen ocho filas más, cada una de las cuales representa otra combinación de atributos que existe para ese descriptor lingüístico.

### 4.3 Búsquedas en la base de datos

Bajo este esquema se pueden hacer búsquedas de la siguiente naturaleza. Supóngase que se determina que una cierta grabación dubitada posee los descriptores lingüísticos 1 y 2 (“articulación fricativa de /tʃ/” y “nasalización vocálica”). El sistema debe buscar en la base de datos, en la tabla de “información diatópica y diastrática”, todos los registros o filas cuya columna *descriptor\_id* tenga el valor 1 o 2. Como resultado se obtienen varias tuplas de atributos correspondientes a las filas de las tablas: (1, 1, 2, 1), (1, 1, 2, 2) y (1, 1, 2, 3) serán algunas de las tuplas que se extraen de la tabla presentada en el Cuadro 4. El sistema deberá buscar que las tuplas que corresponden al descriptor 2 coincidan lo más posible con las tuplas del descriptor 1, en cuanto a región, género y nivel social. La intersección entre los dos conjuntos de tuplas describe las características que han de contener el grupo de las grabaciones del conjunto de referencia contra las cuales resultará lógico hacer la comparación automatizada, en lugar de realizarla contra todo el banco de voces.

## 5 Conclusiones

El proyecto que aquí se presenta pretende ser la base teórica para generar una herramienta que contribuya a mejorar los sistemas automáticos de comparación forense de voz. Actualmente dichos sistemas únicamente analizan las características

sonoras de la muestra de voz. No obstante, identificar los fenómenos lingüísticos que describen una muestra de voz desde diferentes niveles de la lengua permitiría generar grupos de grabaciones dentro del conjunto de referencia, y con ello aumentar la asertividad de las comparaciones automatizadas.

Asimismo, la información generada en este proyecto será la base teórica para desarrollar herramientas que permitan automatizar la clasificación del conjunto de referencia. Por ejemplo, en el caso de los 17 fenómenos fonéticos propuestos es posible entrenar a un sistema automático de reconocimiento de voz que permita identificar aquellos que estén presentes en una nueva muestra de voz sin necesidad de que el experto realice dicha tarea.

Se ha descrito el problema que se pretende abordar y solucionar a través del proyecto “Propuesta de clasificación de un banco de voces con fines de identificación forense”. Asimismo, se ha hecho un breve recorrido por algunos de los estudios más reconocidos que se han realizado en el ámbito de la dialectología y la sociolingüística.

Con el fin de comprender la manera en la que se propone la clasificación de un banco de voces a partir de descriptores lingüísticos, se han expuesto los 17 fenómenos que, en el nivel fonético, se considera pueden ser fácilmente identificables auditivamente por una persona que no sea experta en fonética y que podrían servir para entrenar a un sistema automático de voz. A pesar de que la descripción se hace únicamente a partir de estos 17 fenómenos, el proyecto considera la definición de fenómenos en todos los niveles de la lengua.

Finalmente, se describe la manera en la que la información lingüística es utilizada por una base de datos, no sólo para clasificar y generar grupos de voces con los cuales hacer una comparación automatizada; sino, para realizar descripciones lingüísticas o perfiles lingüísticos de las voces dubitadas, aún y cuando el formato de las grabaciones no lo permite.

Con este proyecto se espera contribuir en la generación e implementación de metodologías, protocolos y técnicas que doten a los laboratorios y personal de las instituciones de procuración de justicia en México de herramientas automáticas que coadyuven en la impartición de justicia.

## Referencias

- Aguilar, Miguel Oscar. 2011. Bases de datos criminalísticos en la procuraduría general de la república. En Sergio García & Olga Islas de González (eds.), *La situación actual del sistema penal en México. XI Jornadas sobre Justicia Penal*, 445–450.
- Aleza, Milagros & José M. Enguita. 2002. *El español de américa: aproximación sincrónica*. Valencia: Tirant lo Blanch.
- Boyd-Bowman, Peter. 1960. *El habla de guajuato*. México: UNAM.
- Canfield, Delos Lincoln. 1962. *La pronunciación del español de américa. ensayo histórico-descriptivo*. Bogotá: Instituto Caro y Cuervo.
- Cardona, Mauricio. 2010. Fonética del español de panamá. En Miguel A. Quesada (ed.), *El español hablado en América Central. Nivel fonético*, Frankfurt: Vervuert.
- Cortichs de Mora, Estrella. 1951. *El habla de tepotztlán*. México: UNAM.
- Cárdenas, Daniel. 1967. *El español de jalisco. contribución a la geografía lingüística hispanoamericana*. Madrid: Consejo Superior de Investigaciones Científicas.
- Espejo, María Bernarda. 1998. Reseña a atlas lingüístico de México. En *THESAURUS*, vol. Tomo III 1, Centro Virtual Cervantes.
- Jiménez, Rafael. 1999. *El andaluz*. Madrid: Arco/libros.
- Lope Blanch, Juan M. 1970. Las zonas dialectales de México. proyecto de delimitación. *Nueva Revista de Filología Hispánica* 19(1). 1–11.
- Lope Blanch, Juan M. 1990. *Atlas lingüístico de México*. México: El Colegio de México – UNAM – FCE.
- López, Humberto. 1983. *Estratificación social del español de san juan de puerto rico*. México: UNAM.
- Matluck, José. 1951. *La pronunciación en el español del valle de México*. México: Edición de autor.
- Mendoza, José E. 2004. *Notas sobre el español del noroeste*. México: El Colegio de Sinaloa.
- Moreno de Alba, José G. 1994. *La pronunciación del español en México*. México: El Colegio de México.
- Moreno de Alba, José G. 1998. *El español en américa*. México: Fondo de Cultura Económica.
- Moreno Fernández, Francisco. 1993. Las áreas dialectales del español americano: historia de un problema. *La división dialectal del español de América* 11–38.

- Morrison, Geoffrey Stewart, Farhan Hyder Sahito, Gaëlle Jardine, Djordje Djokic, Sophie Clavet, Sabine Berghs & Caroline Goemans Dorny. 2016. INTERPOL survey of the use of speaker identification by law enforcement agencies. *Forensic Science International* 263. 92–100.
- Moya, Juan A. & Emilio Weiderman. 1995. La ‘ch’ fricativa en granada: un sonido del habla masculina. En *Actas del XII Congreso de la Asociación Internacional de Hispanistas*, vol. I, Centro Virtual Cervantes.
- Navarro, Tomás. 1918. *Manual de pronunciación española*. Madrid: Consejo Superior de Investigaciones Científicas.
- Navarro, Tomás. 1948. *El español en puerto rico. contribución a la geografía lingüística hispanoamericana*. Río Piedras: Editorial de la Universidad de Puerto Rico.
- Perissinotto, Giorgio. 1975. *Fonología del español hablado en la ciudad de méxico. ensayo de un método sociolingüístico*. México: El Colegio de México.
- Quilis, Antonio. 1999. *Tratado de fonología y fonética españolas*. Madrid: Gredos.
- Ramos, Daniel. 2007. *Forensic evaluation of the evidence using automatic speaker recognition systems*: Universidad Autónoma de Madrid. Tesis Doctoral.
- Rona, José Pedro. 1964. El problema de la división del español americano en zonas dialectales. En *Presente y futuro de la lengua española I (Actas de la Asamblea de Filología del I Congreso de Instituciones Hispánicas)*, vol. I, 215–226. Madrid: Ediciones Cultura Hispánica.
- Rose, Philip. 2002. *Forensic speaker identification*. Taylor & Francis.
- Vaquero, María. 1996. *El español de américa i: Pronunciación*. Madrid: Arco/libros.
- Vivanco, Hiram. 1999. Análisis fonético acústico de una pronunciación de ‘ch’ en jóvenes del estrato social medio-alto y alto de santiago de chile. *Boletín de Filología de la Universidad de Chile* XXXVII. 1257–1269.
- Zamora, Munné Juan C. & Jorge M. Guitart. 1982. *Dialectología hispanoamericana. teoría, descripción, historia*. Salamanca: Colegio de España.