



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

**FACULTAD DE FILOSOFÍA Y LETRAS
COLEGIO DE LETRAS HISPÁNICAS**

**ANÁLISIS DE LOS CONTORNOS ENTONATIVOS DE LAS
ORACIONES DECLARATIVAS NEUTRAS EN EL ESPAÑOL DE LA
CIUDAD DE MÉXICO COMO POSIBLE RASGO PARA LA
COMPARACIÓN FORENSE DE VOZ**

TESIS

Que para obtener el título de

**LICENCIADA EN LENGUA Y LITERATURAS
HISPÁNICAS**

P R E S E N T A

María Alicia Martín del Campo Rodríguez

DIRECTORA DE TESIS

Dra. Fernanda López Escobedo

Ciudad de México, Septiembre 2022.





Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Agradecimientos

Agradezco profundamente a mi asesora, la Dra. Fernanda López Escobedo por su paciencia y apoyo durante este proceso de titulación.

Agradezco también a mis sinodales por sus comentarios, ya que estos enriquecieron este trabajo y me ayudaron poner atención a los detalles, buscar áreas de mejora y pensar en investigaciones futuras.

Gracias a Esteban por haberme apoyado a lo largo de estos años.

Y gracias a mí mamá, por haberme asegurado la oportunidad de estudiar y conseguir este grado.

Índice

1. Introducción	6
2. Objetivo	9
2.1 Objetivo general.....	9
2.2 Objetivos específicos	9
3. Hipótesis	10
4. Estado de la cuestión	10
4.1 La entonación	10
4.2 Formas de medir la entonación	10
4.3 El contorno entonativo en la ciudad de México	11
4.4 Teoría variacionista	18
4.5 Lingüística forense: fonética forense	21
4.6 Uso del contorno entonativo con fines de identificación forense	27
5. Metodología de análisis	31
5.1 Corpus CLOE México	33
5.2 Análisis de datos.....	35
6. Resultados	48
6.1 Oraciones finales	49
6.2 Oraciones no finales	56
7. Análisis de resultados	63
8. Conclusiones	70

9. Bibliografía	72
------------------------------	-----------

Índice de figuras

Figura 4.3.1	13
Figura 4.3.2	16
Figura 4.3.3	17
Figura 4.4. 1	19
Figura 4.4. 2	20
Figura 5.1	31
Figura 5.1.1	34
Figura 5.1.2	35
Figura 5.2.1	35
Figura 5.2.2	38
Figura 5.2.3	39
Figura 5.2.4	40
Figura 5.2.5	41
Figura 5.2.6	41
Figura 5.2.7	42
Figura 5.2.8	43
Figura 5.2.9	44

Figura 5.2.10	45
Figura 6.1.1	50
Figura 6.1.2	52
Figura 6.1.3	55
Figura 6.2.1	57
Figura 6.2.2	60
Figura 6.2.3	62

Índice de tablas

Tabla 5.2.1	36
Tabla 5.2.2	46
Tabla 6.1.1	51
Tabla 6.1.2	53
Tabla 6.1.3	56
Tabla 6.2.1	59
Tabla 6.2.2	61
Tabla 6.2.3	63
Tabla 7.1	65
Tabla 7.2	66
Tabla 7.3	67

1. Introducción

Este proyecto de investigación se enfoca en el análisis de los contornos entonativos de oraciones declarativas neutras que terminan en palabra grave en el español de la Ciudad de México con el propósito de demostrar si las diferencias que presenta el contorno entonativo de una persona al ser comparado con contornos entonativos pertenecientes a otros hablantes son suficientemente significativas como para poder identificarlo.

La finalidad de este análisis recae en el uso que podría tener dentro del ámbito de la lingüística forense al partir de la hipótesis de que el contorno entonativo de cada individuo sirve como un rasgo para identificar hablantes en comparaciones forenses de voz debido a que presenta mayor variación interhablante y menor variación intrahablante.

Las investigaciones realizadas sobre el contorno entonativo del español de la Ciudad de México suelen centrarse en describir dicho contorno de acuerdo con su curva, su aparición en diferentes estratos sociales o en diferentes puntos geográficos y las modificaciones que puede presentar. El modelo de análisis más común hoy en día es el Sp-ToBI, el cual es un modelo de análisis métrico-autosegmental. Este tipo de análisis depende en gran medida del investigador a cargo, ya que está sujeto a la interpretación que se les dé a los datos en cuestión.

Lo anterior presenta un problema al intentar utilizar contornos entonativos en investigaciones del ámbito de la fonética forense, dado que esta tiene como principio que los modelos utilizados sean fáciles de replicar e independientes de interpretación por parte de un investigador, en la medida de lo posible.

Tomando en cuenta lo ya expuesto, se decidió presentar un modelo de análisis de contornos entonativos que sea capaz de ser reproducido de manera sencilla sin mucha intervención de un especialista para que pueda así ser utilizado en el ámbito forense. Dentro de la revisión

bibliográfica se encontró que de forma general la entonación ha sido poco estudiada como un posible rasgo para la identificación de hablantes en comparaciones de voz tanto en países de habla hispana como anglosajones.

Para llevar a cabo la investigación se tomaron muestras de voz del Corpus de Lengua Oral del Español de México¹, CLOE México de aquí en adelante, pertenecientes a 15 hablantes hombres nacidos en la Ciudad de México de entre 20 y 34 años de edad que tuvieran un nivel de instrucción alto. Se utilizó el habla semi-espontánea de estos hablantes y de ahí se seleccionaron 10 oraciones declarativas neutras por hablante que terminaran en palabra grave. Estas oraciones fueron divididas en finales y no finales para el análisis.

Una vez divididas las oraciones, se generó una oración prototípica final y una no final para cada uno de los 15 hablantes y con ellas se crearon gráficas de líneas con la finalidad de comparar los contornos y así formar grupos que compartieran características similares. De esta comparación se obtuvieron tres grupos de oraciones finales y tres grupos de oraciones no finales. Estos grupos representarían a la población relevante en una investigación forense.

En el siguiente paso, ya con los grupos establecidos, se realizó un cálculo de distancias entre cada una de las oraciones de cada uno de los hablantes de un mismo grupo con la finalidad de buscar si los hablantes presentaban menor variación intrahablante que interhablante. El cálculo de distancias se busca para dar un valor numérico a las diferencias entre oraciones, siendo así que entre más alto sea el valor, más distancia habrá entre oraciones y, por lo tanto, más diferentes serán.

Como base teórica para la metodología se tomó la teoría de variación, la cual establece que el lenguaje de un individuo varía dependiendo de las condiciones bajo las cuales se encuentra al momento de la producción de habla. Las variaciones que se producen al momento de hablar son

¹ Acceso a través de <https://selifo.com.mx/corpus-de-lengua-oral-del-espanol-de-mexico/>

inevitables, dos producciones nunca van a ser iguales, aunque sean producidas por la misma persona y bajo las mismas circunstancias. Debido a esto, es necesario que las diferencias entre las muestras de habla de una misma persona sean menores que las diferencias entre muestras de habla de una persona contra las muestras pertenecientes a otras personas, en otras palabras, para poder identificar a un individuo es necesario que presente menor variación intrahablante que interhablante.

A lo largo del análisis surgió un nuevo problema: ¿qué distancia se considera demasiada como para ser parte de la variación intrahablante estándar? Dado que en la literatura no se hace mención al respecto, nuestra propuesta es calcular la variación intrahablante e interhablante utilizando la media del valor de distancias entre las oraciones de un mismo hablante para calcular la variación intrahablante y la media de las distancias del resto de los hablantes del grupo para la variación interhablante. Esto con el propósito de tener un valor numérico y así hacer la interpretación menos subjetiva.

Dentro de los pasos de la metodología únicamente dos requirieron de la interpretación del investigador (1) la selección de oraciones para el análisis y (2) la clasificación de dichas oraciones en finales o no finales. El resto de los pasos se realizaron utilizando programas de análisis fonético y estadístico, por lo que se requiere de la supervisión de un investigador, mas no de su interpretación.

En caso de que la metodología anterior resultara exitosa, se podría comparar la entonación de un hablante sin la necesidad de basarse en su valor en Hertz, el cual puede verse alterado dependiendo el estado de ánimo de la persona, el contexto y la situación de habla, entre otros factores. Asimismo, al no depender del valor en Hertz, se puede realizar un análisis con muestras de voz obtenidas a través de dispositivos que reducen el ancho de banda, como el celular. El

análisis del comportamiento del contorno entonativo como una curva y no como un valor en Hertz permite comparar muestras de voz sin verse atados a las limitantes de variación sociolingüística.

2. Objetivo

2.1 Objetivo general

Explorar el comportamiento del contorno entonativo de las oraciones declarativas neutras en el español de la Ciudad de México y probar su uso como posible método de identificación forense que sea capaz de caracterizar individualmente a un hablante bajo el presupuesto de que existe menor variación intrahablante y mayor variación interhablante, a pesar de pertenecer al mismo dialecto.

2.2 Objetivos específicos

Dentro de los objetivos específicos desarrollados durante la investigación se encuentran:

- Identificar las formas existentes para medir la entonación.
- Seleccionar y analizar muestras de voz.
- Buscar y proponer una forma de determinar la variación intrahablante e interhablante.
- Proponer un modelo de análisis objetivo, fácil de replicar y que no dependa en gran medida de la interpretación del investigador que lo utilice.

3. Hipótesis

El contorno entonativo de cada individuo sirve como un rasgo para identificar hablantes en el ámbito forense debido a que presenta mayor variación interhablante y menor variación intrahablante.

4. Estado de la cuestión

4.1 La entonación

Cantera (2002) define la entonación como las variaciones de la frecuencia fundamental que cumplen una función lingüística a lo largo de la emisión de la voz. Por su parte, Quilis (1981:376) afirma que la entonación actúa en tres niveles: en el nivel lingüístico cumple una función distintiva, en el nivel sociolingüístico informa e identifica, y en el nivel expresivo transmite el estado emocional del hablante. Más tarde, en su *Tratado de Fonología y Fonética Españolas* (1999:411) habla de la entonación como “la función lingüísticamente significativa, socialmente representativa e individualmente expresiva de la frecuencia fundamental en el nivel de la oración.”

Para Cantera (2002), la entonación cumple tres funciones: la función expresiva permite ver la perspectiva del interlocutor, la focalizadora sirve para que el hablante dé relevancia a la parte central del mensaje, y la demarcativa, en la cual el emisor divide el discurso en unidades tonales para que el oyente pueda segmentarlo e interpretarlo. Asimismo, la entonación cuenta con tres parámetros: frecuencia fundamental, duración e intensidad.

4.2 Formas de medir la entonación

El elemento más utilizado al momento de medir la entonación es la frecuencia fundamental. Künzel (1995:292), la define como “el correlato acústico del tono percibido”, el cual puede verse

afectado por factores psicológicos, fisiológicos o situacionales presentes al momento en el que el hablante produce la oración. La frecuencia fundamental también puede ser definida como un parámetro acústico producido por las vibraciones de las cuerdas vocales (Cantera, 2002), o como la nota base de la voz, el ritmo al cual vibran las cuerdas. (Olsson, 2008:162).

Lo que se conoce como *contorno entonativo* es un reflejo de la frecuencia fundamental, el cual cuenta con un significado fonológico determinado y constante, y está compuesto por el *significante entonativo*, que es el nombre otorgado a la melodía contenida en el grupo fónico. Por su parte, el grupo fónico está conformado por una serie de palabras fónicas organizadas alrededor de un acento sintagmático, estas palabras son sonidos consonánticos que se agrupan alrededor de sonidos vocálicos formando sílabas, y éstas, a su vez, se concentran alrededor de vocales más tónicas. El acento sintagmático u oracional es un acento puesto en relieve sobre los demás y generalmente es el último acento del grupo fónico. (Cantera, 2002).

4.3 El contorno entonativo en la ciudad de México

La mayoría de los estudios de análisis entonativo utilizan el modelo métrico-autosegmental de análisis entonativo, Sp_ToBI (*Spanish ToBI*). De acuerdo con Vilaplana y Prieto (2008), el sistema ToBI distingue cuatro niveles de análisis: 1) el nivel ortográfico, 2) el nivel tonal, 3) el nivel de separación prosódica y 4) el nivel misceláneo. Asimismo, mencionan que el sistema ToBI describe las curvas melódicas en relación con la estructura métrica de los enunciados y propone la existencia de dos unidades fonológicas: 1) acentos tonales, que se asocian a sílabas con acento sintagmático, es decir, la última sílaba acentuada en una oración, y 2) tonos de frontera, que se asocian a las fronteras de los dominios prosódicos. Los tonos se representan mediante sus iniciales en inglés: L (tono bajo, «low tone») y H (tono alto, «high tone»). (Vilaplana y Prieto, 2008:26).

De la Mota (et. al., 2010) proponen un inventario de acentos y tonos de frontera para el español de México, el cual se describe a continuación.

En el caso de los acentos tonales marcan las siguientes posibilidades:

- L*: describe un acento bajo y plano, realizado en el límite mínimo de F0 del hablante.
- H*: describe un acento plano y alto sin valle de F0 que lo preceda.
- L+H*: describe un movimiento ascendente durante la sílaba acentuada con el pico de F0 localizado al final de la sílaba.
- L+>H*: describe un movimiento ascendente durante la sílaba acentuada con el pico de F0 alineado con la sílaba postónica.
- L*H: describe un valle de F0 en la sílaba acentuada con un ascenso subsecuente en la sílaba posterior.
- H+L*: describe una caída de F0 dentro de la sílaba acentuada.

El inventario de tonos de frontera es el que sigue:

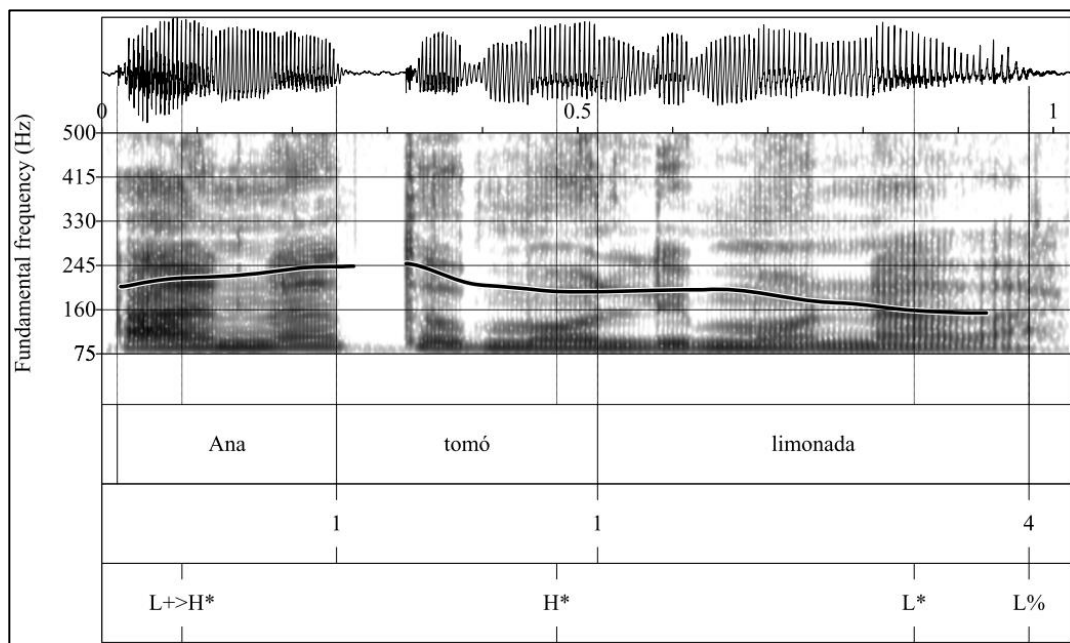
- L%: describe un tono bajo y mantenido en la base fonética del hablante.
- M%: describe un movimiento de ascenso o descenso hacia un punto medio.
- H%: describe un movimiento de ascenso procedente de un acento tonal alto o bajo.
- LH%: describe un valle de F0 seguido por una subida.
- HL%: describe un pico de F0 seguido por una caída.
- LM%: describe un valle de F0 seguido por un movimiento de ascenso a un punto medio.

Adicionalmente, se utilizan los signos ¡ y ! para marcar escalonamiento en los acentos tonales. El escalonamiento significa que cada pico tonal es más bajo o más alto que el anterior. Para marcar escalonamientos descendentes se utiliza el símbolo ! y para marcar escalonamientos ascendentes se utiliza ¡. (Beckman et. al., 2002:16).

Lo mencionado anteriormente puede verse ejemplificado en la figura 4.3.1, la cual muestra un espectrograma con el contorno entonativo marcado por la F0 de la oración “Ana tomó una limonada” la cual ha sido etiquetada utilizando el modelo Sp-ToBI, y tiene una configuración nuclear de L* L%.

Figura 4.3.1

Ejemplo configuración L L%. (adaptado de P. Prieto y P. Roseano (eds.), 2010: 325.)*



Nota. Curva, espectogramas y F0 de la oración ‘Ana tomó una limonada’, producido con una configuración L* L%.

De la Mota (et. al., 2010) establecen que en la Ciudad de México se presenta un contorno entonativo muy peculiar conocido como “circunflejo”, el cual se caracteriza por un movimiento de ascenso y descenso particular. De acuerdo con el modelo entonativo Sp- ToBI, se considera que el acento prototípico de este tipo de contornos es L+;H*, el cual suele estar seguido por un descenso en el tono de frontera (L-H%). Martín Butragueño (2008), describe esta configuración

como un acento que presenta un ascenso tonal pronunciado, el cual se extiende a lo largo de la sílaba nuclear, y cuyo pico suele estar alineado con el final de la sílaba. Dicho pico declina rápidamente en la sílaba postónica en forma de tono de juntura. Este tipo de configuración puede ser menos común entre hablantes de estrato social alto, mujeres o en habla formal.

La configuración circunfleja puede presentarse en declaraciones de enfoque estrecho, con un acento L+H* y un tono de frontera L%, o de enfoque amplio, conservando el mismo acento, pero con la posibilidad de tonos de frontera L%, HL%, M% y LM%. Además de las posibles configuraciones anteriores, existe otra configuración circunfleja que difiere notablemente de la prototípica, la cual tiene un acento L* y un tono de frontera L%. (De la Mota et. al., 2010). En el estudio llevado a cabo por los autores, se presentaron diferencias en las configuraciones circunflejas entre los tres hablantes entrevistados, sin embargo, debido al reducido corpus de análisis, no afirman que sus resultados sean definitivos.

Otros autores han hablado previamente sobre la entonación circunfleja, Henríquez Ureña (1938), por ejemplo, marca que:

es característica la cadencia final de la frase enunciativa, muy distinta de la cadencia usual en Castilla. [...] en el habla popular de Méjico (sic), de la antepenúltima a la penúltima sílaba se asciende aproximadamente una tercera y de la penúltima a la última se descende aproximadamente una sexta, siendo la penúltima larga y la final muy breve. (p. 335).

El término *cadencia* refiere a un tono descendente al final de una oración. Por su parte *anticadencia* significa que el tono asciende y *suspensión* se entiende como un tono que se mantiene. (Estebas-Vilaplana, E y Prieto, P., 2010: 17).

Matluck (1965) toma un acercamiento diferente y le otorga un valor exclusivamente fonético al contorno circunflejo, de acuerdo con sus anotaciones, la curva circunfleja consistiría en un nivel de tono /2/ en vez de /1/ para la sílaba acentuada y una terminación suspensiva en vez de descendente, en donde la sílaba final puede descender o no y suele alargarse notablemente. Para él, la impresión circunfleja se intensifica por la fuerza de las sílabas fuertes y la debilidad de las átonas. Es necesario señalar aquí que el nivel tonal /2/ refiere a un tono medio, mientras que el nivel tonal /1/ refiere a un tono bajo. Junto con los dos niveles tonales anteriores existe un nivel tonal /3/ que se utiliza para marcar tonos altos. (Quilis, A. 1999: 422).

Por su parte, Kvavik (1974) habla al respecto de la configuración circunfleja y distingue dos tipos: las cadencias simples, que son unidireccionales y se desarrollan a partir de la postónica final; y las complejas, que adoptan diferentes formas; en estas últimas, el movimiento cambiante inicia en la sílaba tónica final. Las cadencias complejas circunflejas, en específico, pueden ser suspensivas, descendentes o ascendentes. Para Quilis (1985:160), la entonación circunfleja consiste en un movimiento de ascenso-descenso al final del enunciado. La frecuencia fundamental termina de forma descendente en un nivel igual o mayor que el del cuerpo del enunciado, mientras que la cima, o acento sintagmático, coincide con la última vocal tónica. En el caso de las mujeres, a lo largo del enunciado, el F0 asciende 33% desde el inicio hasta la cima y la gama de frecuencias en este tramo aumenta en promedio 59 Hz. A partir de la cima, el F0 disminuye dicho 33% y la gama de frecuencias desciende 59 Hz, es decir, ambos rubros bajan lo mismo que subieron. El caso de los hombres es distinto, del inicio del enunciado a la cima, el F0 asciende un 41.6% y la gama de frecuencias sube una media de 29.5 Hz, posteriormente, el F0 desciende 33% y la gama de frecuencias baja una media de 23.6 Hz.

Finalmente, Martín Butragueño (2008) marca un número de variantes para el acento tonal en la entonación circunfleja: H^* , $L+H^*$, $L+;H^*$, $L+(;)H^*L-$, L^* , $H+L^*$, $H+!L^*$. Siendo $L+;H^*$ el acento más típicamente asociado a la circunflexión, el cual presenta un ascenso pronunciado que se extiende a lo largo de la sílaba nuclear, o desde antes, y cuyo pico suele estar alineado al final de la sílaba tónica, para posteriormente descender en la postónica con un tono de juntura $L\%$.

Con base en los párrafos anteriores, se puede inferir que la parte más relevante del contorno entonativo circunflejo se encuentra al final del enunciado, el cual está compuesto por la sílaba tónica o núcleo oracional y el tono de frontera o cola, siendo estos apartados los que más variación presentan. La figura 4.3.2 muestra un contorno circunflejo, con el núcleo oracional o sílaba tónica en “ca” y la cola o tono de frontera en “sas”. Por su parte, en la figura 4.3.3 podemos ver un contorno entonativo descendente, el más común en la Ciudad de México, con el núcleo oracional o sílaba tónica en “na” y la cola o tono de frontera en “n(d)o”.

Figura 4.3.2

Ejemplo de contorno circunflejo.

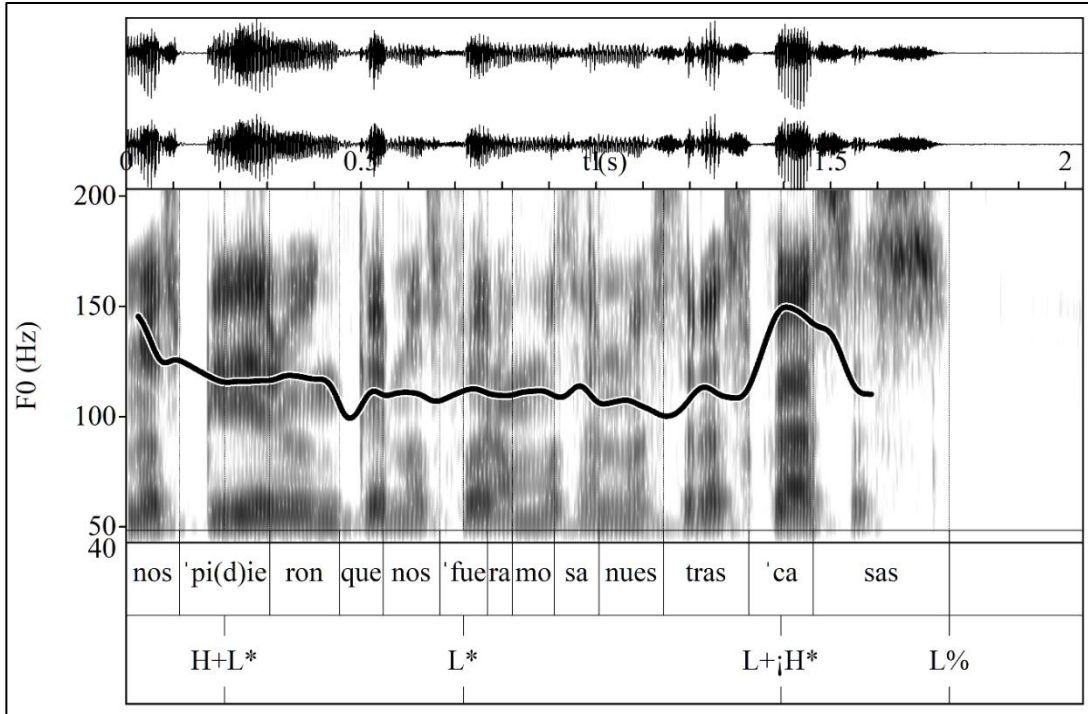
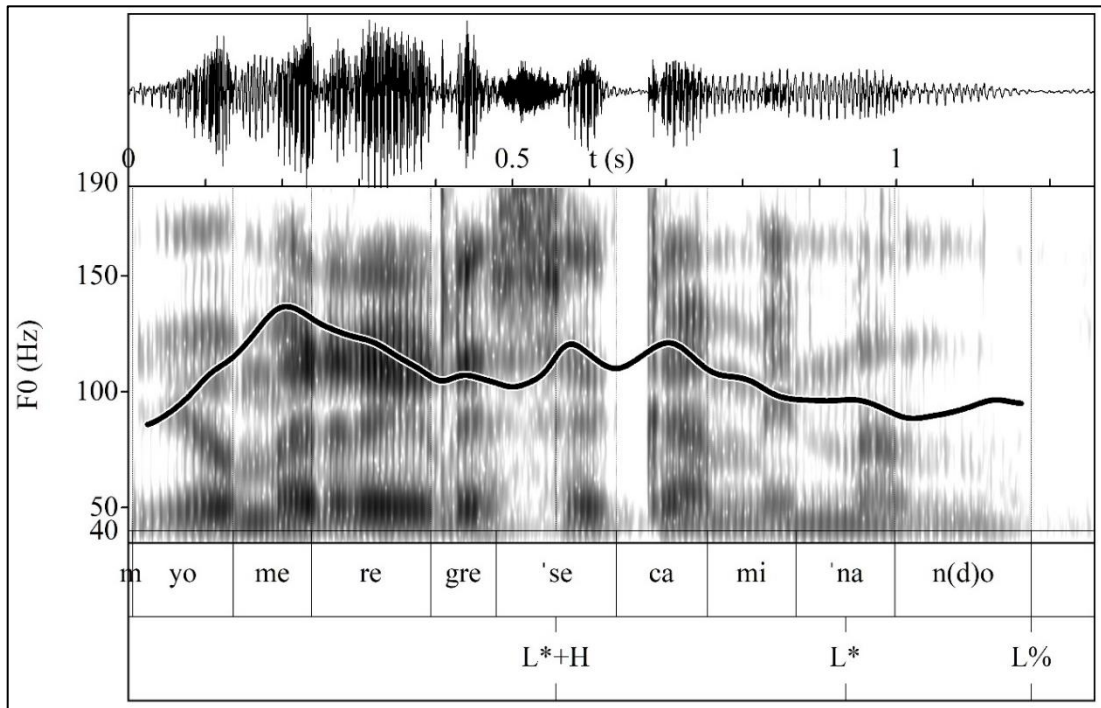


Figura 4.3.3

Ejemplo de contorno descendente.



Si comparamos la oración de la figura 4.3.2 con la mostrada en la figura 4.3.3, podemos apreciar las diferencias entre ambos contornos, siendo la más importante la subida que presenta la figura 4.3.2 en la sílaba tónica y su consecuente caída a un tono bajo (L%) en el tono de frontera.

Tomando en cuenta lo mencionado, se puede interpretar que en la Ciudad de México existen dos contornos entonativos predominantes, estando en primer lugar el contorno descendente al ser el más utilizado. El otro contorno que se puede encontrar es el circunflejo el que, aunque menos utilizado, es común en el español de la Ciudad de México.

4.4 Teoría variacionista

La sociolingüística es la disciplina encargada de estudiar la relación entre lenguaje y sociedad, dentro de ella se han desarrollado diversas teorías que apoyan al análisis lingüístico. Una de ellas es la teoría variacionista, la cual tiene como base el análisis de las características lingüísticas que presentan variación como respuesta a diferencias sociológicas, contextuales, temporales y geográficas de los individuos. Dicho en otras palabras, cuando las diferencias sociológicas, contextuales, temporales y geográficas de un individuo varían, su lenguaje también varía. (Caicedo, 1991).

Para comprender la teoría variacionista es importante entender los conceptos de variación lingüística y lengua estándar. Se entiende por variación lingüística los cambios que sufre una lengua que están motivados por fenómenos regionales o sociales, circunstancias comunicativas ocasionales u otros fenómenos extralingüísticos. Por otra parte, lengua estándar es aquella que tiene reconocido prestigio social en la comunidad y que sirve como modelo de comunicación. (Caicedo, 1991).

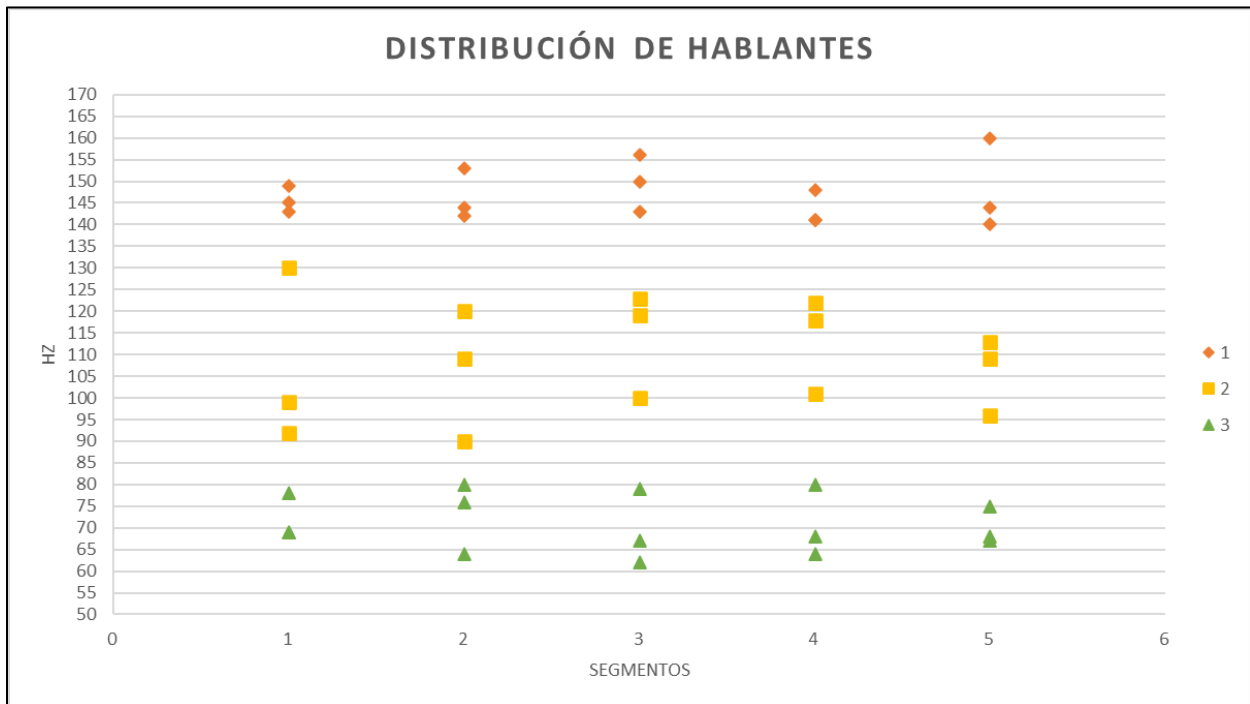
La teoría variacionista se vuelve relevante en el ámbito de la lingüística forense debido a que, para diferenciar entre voces, se busca encontrar variaciones en las características lingüísticas

antes mencionadas para utilizarlas en la comparación de voz. Como menciona Rose (2002), “si se planea discriminar entidades con base en uno de sus atributos, dichas entidades deben diferir en ese atributo [...]. Por lo tanto, si se planea discriminar individuos con base en sus voces, los individuos deben diferir en sus voces.” (p.10)

La figura 4.4.1 muestra la distribución de hablantes de acuerdo con la media en Hz de las cinco sílabas más prominentes en una oración. Se puede observar que cada hablante ocupa un espacio específico de la gráfica. En este caso, la variación existente entre estos hablantes es suficiente para decir que se trata de diferentes individuos.

Figura 4.4. 1

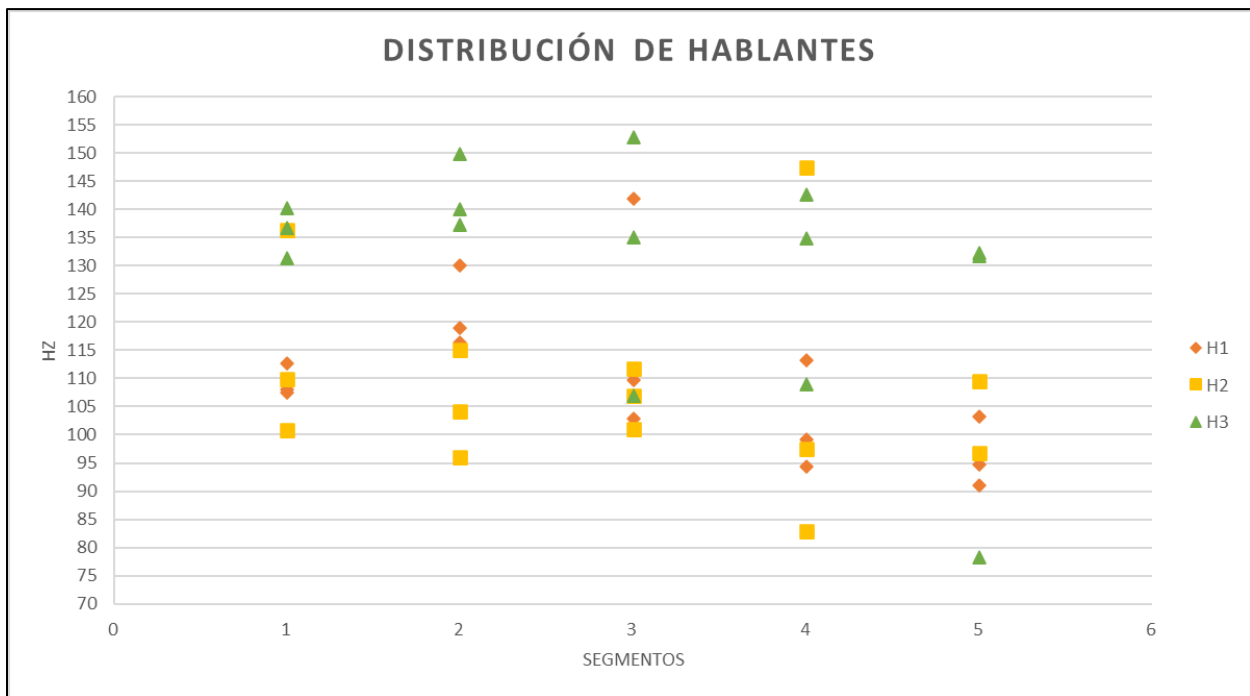
Distribución ideal de hablantes.



Es necesario mencionar que la figura 4.4.1 fue hecha con datos falsos con la finalidad de ilustrar la variación entre hablantes. Sin embargo, en la realidad los límites entre las muestras de un mismo hablante suelen ser mucho más difusos. Esto puede apreciarse en la figura 4.4.2, la cual contiene los datos de la media de las sílabas en Hertz de tres hablantes, tomadas del CLOE México.

Figura 4.4. 2

Distribución real de hablantes.



En la gráfica anterior se puede observar la distribución real que suele encontrarse al realizar una comparación forense de voz. A diferencia de la figura 4.4.1 donde las diferentes muestras de un mismo hablante se graficaban en valores cercanos, de tal manera que eran fáciles de diferenciar entre ellos, en la figura 4.4.2 podemos ver que la dispersión de las muestras de los hablantes suele ser mayor. En el caso de estos tres hablantes, es imposible diferenciar entre H1 y H2, ya que la mayoría de sus valores están sobrepuestos dentro del mismo espacio. Finalmente, H3 resulta más

fácil de diferenciar ya que su variación intrahablante es más uniforme en contraste con la variación interhablante que presenta.

Es así como, con base en los fundamentos teóricos planteados por la teoría variacionista, uno de los principios más relevantes en la lingüística forense es que para identificar a un sujeto, su variación intrahablante debe ser menor a su variación interhablante. El cumplimiento de ello dependerá de una diversidad de factores, no todos ellos posibles de controlar, por lo que se requiere realizar una búsqueda de características que cumplan con este principio. Entre más características lingüísticas se puedan medir en la voz, mayor será la probabilidad de encontrar el conjunto de aquellas que en un individuo cumplan con dicho principio.

Es necesario mencionar que la variabilidad en las características lingüísticas de un individuo no será universal, es decir, cada individuo presentará mayor o menor variabilidad intrahablante en determinadas características, y esto no necesariamente será igual para todos. Ello nos lleva a sugerir la medición de distintas características lingüísticas que permitan que el principio de menor variación intrahablante y mayor variación interhablante se cumpla en la voz de un individuo. Lo que se busca en este trabajo de investigación es explorar si la medición del contorno entonativo de un hablante sirve como un rasgo para su identificación en el ámbito forense debido a que presenta mayor variación interhablante y menor variación intrahablante.

4.5 Lingüística forense: fonética forense

Se entiende por lingüística forense a la aplicación del conocimiento científico del lenguaje dentro del contexto legal. El término fue usado por primera vez por el profesor Jan Svartvik en 1968 en su libro *“The Evans statements. A case of forensic linguistics”* (Ariani et. al. 2014).

Para Olsson (2008) la lingüística forense es el punto de contacto entre el lenguaje, el crimen y la ley, por lo cual, el lingüista forense debe aplicar el conocimiento y las técnicas lingüísticas que domina al lenguaje implicado en los casos o procedimientos legales, o en disputas privadas que puedan resultar en la toma de acciones legales. Por su parte, Coulthard et. al. (2010: 529) sugieren que es la aplicación de la lingüística a los textos legales escritos, las prácticas legales orales y la aportación de evidencia para las investigaciones criminales y civiles.

Vista por Perkins y Grant (2013), la lingüística forense puede entenderse como una rama de la lingüística aplicada que se relaciona con el campo de la ley y los procesos legales. Como se puede ver, la mayoría de las definiciones tratan a la lingüística forense como una herramienta de apoyo para los procesos legales que involucran lengua hablada o escrita.

La lingüística investigativa y el suministro de evidencia forman parte de la lingüística forense y consisten en un análisis comparativo de autoría y perfilado sociolingüístico, entre otros aspectos. El perfilado sociolingüístico refiere a que la producción lingüística de cada individuo tiene influencia de varios factores sociales, tales como: edad, sexo y estatus educativo; con él se pretende determinar información sobre un autor anónimo o los orígenes de un texto (Grant y Perkins, 2013: 176).

Existe una rama de la lingüística forense que consiste en el uso de las técnicas fonéticas en el análisis de voz aplicado a la investigación criminal denominada fonética forense, la cual abarca comparaciones técnicas de voz, reconocimiento de voz, transcripción del lenguaje hablado (incluyendo emisiones en disputa), la mejora de la señal del habla y la autenticación de grabaciones. (Olsson, 2008: 156)

De acuerdo con Hollien (2012), la fonética forense se especializa en buscar las partes de los procesos comunicativos del ser humano que pueden aplicarse para cubrir las necesidades de

los sistemas de justicia penal, judicial y las agencias de inteligencia a través del uso de técnicas especializadas. Está conformada por dos elementos: el primero involucra el análisis electroacústico de las señales de habla que han sido transmitidas y almacenadas. Trata los temas de transmisión y almacenamiento adecuado de habla, la autenticación de grabaciones, el mejoramiento de inteligibilidad del habla, el decodificado del habla y aspectos relacionados. El segundo elemento refiere al análisis de los actos comunicativos mismos y consiste en el reconocimiento de hablantes a partir de un análisis de voz, la identificación del estado de salud, emocional y psicológico del hablante, y el análisis del habla como evidencia de fingimiento de voz.

Para San Segundo (2013: 587-588), la fonética forense es la aplicación del conocimiento fonético general a los problemas legales, por ejemplo, el contribuir a la identificación de un hablante. Por lo tanto, el fonetista forense debe ser capaz de diseñar alineaciones de voz y crear el perfil de voz del hablante, entre otras cosas. En el caso de las comparaciones forenses de voz, es imprescindible que el experto compare una o varias muestras de habla de una voz desconocida con una o varias muestras de voz conocida. Es necesario poder discriminar entre una muestra de voz del mismo hablante y muestras de diferentes hablantes para poder afirmar que una misma voz está presente en dos o más muestras.

Es bien sabido que personas distintas tienen voces distintas, pero incluso la voz de una misma persona presenta variaciones ya que nadie dice una frase exactamente de la misma forma, esto es conocido como variación intrahablante. A la variación que se presenta entre dos o más hablantes de la misma lengua se le conoce como variación interhablante. Por lo tanto, siempre van a existir diferencias entre muestras de voz, aunque provengan de un mismo individuo, pero la variación interhablante debe ser mayor que la variación intrahablante para poder identificarlos. Entre más grande sea la diferencia entre ellas, más fácil será la identificación del hablante. La

identificación forense de voz involucra la capacidad de definir si las diferencias inevitables que pueden presentarse entre dos muestras es más probable que se deban a diferencias intrahablante o interhablante. (Rose, 2002).

Siguiendo a Morrison (2009), existen cuatro tipos de acercamiento a la comparación forense de voz: auditivo, espectrográfico, acústico-fonético y automático. El acercamiento auditivo lo llevan a cabo fonetistas con años de entrenamiento y experiencia, estos fonetistas escuchan las muestras de voz y hacen comentarios referentes a las propiedades de la voz que puedan considerarse como inusuales o distintivas.

El acercamiento espectrográfico se basa en el uso de espectrogramas, los cuales permiten visualizar los cambios en la amplitud de las propiedades acústicas de una frecuencia. Es considerado un acercamiento no científico ni fiable debido a que el análisis consiste en la comparación visual de espectrogramas con el fin de llegar a una expresión cualitativa de la probabilidad de que las muestras tengan un mismo origen o provengan de diferentes fuentes.

El acercamiento acústico-fonético involucra llevar a cabo medidas cuantitativas de las propiedades acústicas de los sonidos del habla. Las unidades fonéticas comparables son identificadas tanto en la muestra de habla dubitada (de origen desconocido) como en la no dubitada (de origen conocido) para posteriormente medir las propiedades acústicas de las unidades. Este tipo de análisis se basa en el uso de algoritmos que procesan señales bajo la supervisión de un experto. El tiempo y el presupuesto necesarios para este acercamiento representan uno de los obstáculos para su aplicación.

El acercamiento automático fue diseñado por ingenieros especializados en el procesamiento de señales y también se basa en medidas cuantitativas de las propiedades acústicas del habla, pero no se han hecho intentos para explotar la información relacionada con unidades

fonéticas. Este acercamiento procesa rápida y económicamente grandes cantidades de información, la desventaja es que incluye sonidos sin importancia, como el ruido de fondo, dentro de su análisis.

En los últimos años se presentó un cambio en el paradigma de la investigación forense. Saks y Koehler (2005:892) describen el nuevo paradigma como una “ciencia empíricamente fundamentada, lo cual se puede ver en la valoración probabilística de información.” Los autores recomiendan que otro tipo de comparaciones forenses imiten las comparaciones de ADN, incluyendo la construcción de bases de datos que presenten las características encontradas en las muestras a analizar, para posteriormente hacer uso de dichas bases en un acercamiento probabilístico. Otro aspecto al que los autores dan importancia en el nuevo paradigma es la cuantificación y el reporte de las limitaciones de la comparación forense a partir de los rangos de error. (p.893).

Morrison (2014:246) establece que la evaluación de evidencia forense debe basarse en tres elementos:

1. El uso obligatorio del marco teórico del *likelihood ratio* (razón de verosimilitud).
2. El uso de acercamientos basados en medidas cuantitativas, bases de datos representativas de la población relevante y modelos estadísticos.
3. Una prueba obligatoria de validez y fiabilidad bajo condiciones parecidas al caso en investigación utilizando información proveniente de la población relevante.

Morrison (2009:299), explica que el primer elemento, la razón de verosimilitud, se basa en la pregunta “¿qué tanto más probable es que se presenten las similitudes o diferencias observadas entre la muestra conocida y desconocida bajo la hipótesis de que ambas tienen el mismo origen que bajo la hipótesis de que tienen orígenes diferentes?”. Por su parte, Hughes y Foulkes (2015:

218) explican el likelihood ratio como el rango de probabilidad (p) de la evidencia (E) asumiendo la proposición de la acusación (Hp) y la probabilidad de la evidencia asumiendo la proposición de la defensa (Hd). Esto puede expresarse como:

$$\frac{p(E|H_p)}{p(E|H_d)}$$

Donde:

- Hp se entiende como la fuente de la grabación del delincuente es el sospechoso.
- Hd puede expresarse como la fuente de la grabación del delincuente no es el sospechoso, sino otro hablante de la población relevante.
- El numerador es equivalente a la similitud entre las muestras del delincuente y del sospechoso, y
- El denominador es equivalente a la tipicidad de la muestra del delincuente con respecto a la población relevante.

Para determinar la población relevante, se deben tomar en cuenta las características del delincuente, por lo que el analista deberá realizar un perfil lingüístico. Usualmente se limita al sexo del delincuente, su estrato geográfico, clase socioeconómica, edad, etnia y la comunidad a la que pertenece (Hughes y Foulkes, 2015: 219).

Respecto al segundo elemento para la evaluación de evidencia forense, Morrison (2014:246) plantea los acercamientos basados en medidas cuantitativas, las bases de datos representativas de la población relevante y los modelos estadísticos. Estos son preferidos sobre los acercamientos basados en la experiencia humana, debido a que este tipo de acercamientos y modelos son más transparentes, más fáciles de replicar y más sencillos de interpretar al ser puestos a prueba bajo criterios de validez y fiabilidad. Asimismo, cuentan con la posibilidad de describir la información utilizada, las medidas llevadas a cabo y los modelos estadísticos aplicados con el

suficiente detalle como para que otro analista pueda copiar lo que se ha hecho.

Con respecto al tercer elemento, el término validez refiere al alcance con el cual el conjunto de medidas promedio se aproximan al verdadero valor de la propiedad que está siendo medida. Fiabilidad, por otro lado, trata con el margen del conjunto de medidas alrededor del valor promedio de las mismas. Ambos aspectos pueden calcularse utilizando una base de datos de pares de muestras de prueba procedentes de un mismo origen o de orígenes distintos. La persona que lleve a cabo la prueba debe saber cuáles pares son de un mismo origen y cuáles son de diferente procedencia, pero el sistema que se está poniendo a prueba no debe tener acceso a esta información. El sistema deberá proporcionar un resultado para cada par y la persona a cargo compara las respuestas del sistema con la realidad, asigna una penalización de acuerdo a la exactitud de la respuesta y saca un promedio de penalizaciones como indicador de la validez del sistema (Morrison, 2014: 247).

Tomando en cuenta los elementos de análisis forense de Morrison (2014: 246), podemos decir que la frecuencia fundamental es un parámetro del habla capaz de ajustarse a modelos cuantitativos de análisis lo que, a su vez, significa que puede proporcionar información significativa sobre los hablantes en una comparación forense de voz.

4.6 Uso del contorno entonativo con fines de identificación forense

Considerando lo descrito en el apartado anterior, es necesario establecer los parámetros acústicos suprasegmentales que son relevantes en la identificación de hablantes, al igual que analizar la variación intrahablante e interhablante para poder escribir de forma detallada los rasgos que caracterizan a un individuo (Battaner et. al., 2003). Es importante remarcar que los estudios presentados a continuación se centran en dos aspectos diferentes, el primer aspecto refiere al valor

numérico de la frecuencia fundamental presentado usualmente en Hertz, el cual puede verse afectado por diversos aspectos. El segundo aspecto trata el estudio del contorno entonativo, este campo no ha sido poco explorado y puede contener aspectos útiles para la identificación de hablantes.

Con respecto a la frecuencia fundamental, algunas investigaciones han intentado relacionar este parámetro con la altura del hablante, tal es el caso de Lass (1980) y Lass et. al. (1978 y 1980) cuyos estudios afirmaban que la estatura y el peso, tanto de hablantes masculinos como femeninos, podía ser estimada con un rango de error de menos de 1.35 pulgadas (3.4cm) y 4 libras (1.8 kg) por oyentes sin preparación. Más tarde, Künzel (1989) llevó a cabo un estudio en el que utilizó muestras que semejaban el mensaje típico por parte de un secuestrador a los padres de la víctima en una prueba simulada, dicha prueba tenía una duración de 60 segundos y fue grabada 15 veces por cada hablante con intervalos de seis semanas entre cada grabación. Los sujetos grabados fueron 105 hombres y 78 mujeres en un rango de edad de 19 a 61 años, lo que concluyó en 915 muestras. Los resultados de ambos estudios demostraron que no existe una relación entre la frecuencia fundamental y la altura o el peso del hablante, sea este hombre o mujer, por lo que se puede asumir que esta información se localiza no sólo en un parámetro acústico.

Por su parte, Akagi e Ienaga (1997) proponen que algunas características relacionadas con la individualidad del hablante se encuentran insertadas en los contornos entonativos de palabras, por lo que investigaron la relevancia de los parámetros que describen la frecuencia fundamental en los modelos utilizados para extraer y manipular características físicas del contorno. Para lo anterior, realizaron tres experimentos, el primero demostró que la individualidad del hablante existe en los contornos de la frecuencia fundamental; el segundo puso en evidencia que los parámetros relativos a las dinámicas de los contornos de la frecuencia fundamental son importantes

en la identificación de hablantes; el tercero manifestó que cambiar el promedio de tiempo de un contorno de un hablante a otro no ofrece un mayor rango de identificación.

Los estudios llevados a cabo por Pruzansky (1963), Hargreaves & Starkweather (1963), Li, Dammann & Chapman (1966), Luck (1969), Das & Mohn (1969), Bricker et.al (1971) y Doddington (1971) demostraron que el reconocimiento automático de hablantes es factible tomando una población pequeña. Tomando estas investigaciones como base, Atal (1972) se dio a la tarea de analizar la posibilidad de utilizar los contornos entonativos para identificar hablantes debido a las ventajas que presenta, como el no verse afectada por variaciones causadas por el sistema de transmisión y ser independiente tanto del volumen de voz del hablante como de la distancia entre el hablante y el micrófono. Para su estudio utilizó el contorno entonativo en su totalidad, puesto que es poco probable que un impostor sea capaz de imitar la variación presente a lo largo del contorno como una función de tiempo, en otras palabras, el autor se concentra en la duración de los contornos entonativos, tomando en cuenta el tiempo que le toma al hablante producir un enunciado y lo compara con otro. Lamentablemente, la duración de los contornos entonativos demostró tener una gran variación intrahablante, por lo que no resulta útil para el reconocimiento de hablantes.

Para el español en específico, se han llevado a cabo investigaciones de carácter geoprosódico y, de acuerdo con Congosto (2015), estas han demostrado que la entonación tiene mucho que ofrecer a la fonética forense puesto que puede proporcionar información de tipo dialectal y sociolingüístico de especial relevancia sobre el individuo, tal como el grupo social al que pertenece, su origen geográfico, el medio social en el que se inserta, entre otros. Asimismo, aporta información sobre la edad, sexo e incluso el temperamento, carácter y estado de ánimo del hablante.

Los estudios que se han llevado a cabo sobre la frecuencia fundamental como base para la identificación de hablantes en español en su mayoría han tomado como metodología la identificación por parte de personas, es decir, hablantes que sean capaces de relacionar o diferenciar un sonido de otro de entre un grupo de varias muestras utilizando como punto de referencia únicamente lo que escuchan.

Como ejemplo de lo anterior tenemos el estudio de Congosto (2015), en el cual la autora quería observar si se podía diferenciar entre el español hablado en la Ciudad de México y el español de Los Ángeles, EE.UU., para cuyo fin realizó tres pruebas: (1) en la cual se debía diferenciar entre aseveración y pregunta, (2) en la que se tenía que discriminar entre sonidos diferentes o iguales y (3) donde se reproducían tres sonidos, siendo el primero y el último diferentes, para al final decir qué sonido era semejante al segundo. Congosto seleccionó jueces cuya tarea era identificar o discriminar entre las dos variedades de español. El grupo de jueces se compuso por 9 personas de la Ciudad de México y 7 procedentes de otros países de habla hispana. Los resultados demostraron que la prueba 1 y 3 tuvieron un alto porcentaje de reconocimiento por parte de los informantes, mientras que la prueba 2 no resultó tan eficaz.

Lo anterior demuestra que a pesar de que ha habido estudios que utilizan la frecuencia fundamental como base para la identificación de hablantes, sólo el trabajo de Akagi y Ienaga (1997) se centra en las peculiaridades que presenta el contorno entonativo por sí mismo, mientras que el resto de las investigaciones (Lass (1980), Lass et. al. (1978 y 1980), Künzel (1989), Atal (1972) y Congosto (2015)) buscan parámetros externos a la frecuencia fundamental, la mayoría con la finalidad de relacionarlos a alguna cualidad física o sociocultural del hablante. Cabe agregar que la metodología empleada por una gran cantidad de autores, la cual consiste en utilizar oyentes de la misma lengua dentro de sus experimentos para identificar las variaciones en la entonación,

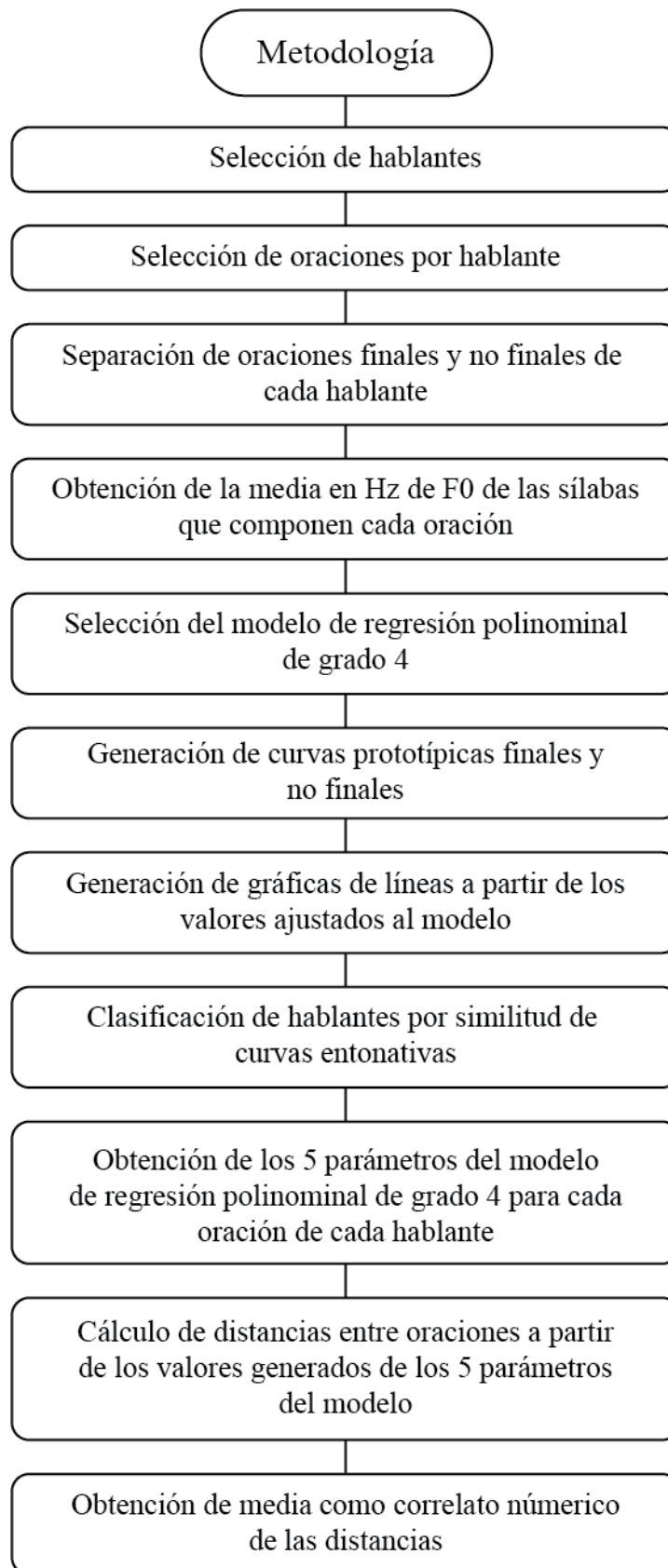
resulta extremadamente subjetiva y difícil de llevar a cabo en el ámbito forense puesto que es complicado establecer la serie de parámetros que utilizan los oyentes para discernir entre una muestra de voz y otra, por lo cual se sugeriría apegarse a modelos más estadísticos y fáciles de replicar.

5. Metodología de análisis

El siguiente diagrama contiene los pasos llevados a cabo a lo largo de esta tesis a forma de resumen.

Figura 5.1

Pasos de la metodología.



5.1 Corpus CLOE México

El Corpus de Lengua Oral del Español de México (CLOE México) está conformado por grabaciones de voz llevadas a cabo en la cabina de grabación de la Licenciatura en Ciencia Forense de la Facultad de Medicina de la UNAM. El corpus cuenta con divisiones de acuerdo al nivel de instrucción, la edad y el sexo del hablante, asimismo, se separa a los hablantes por su lugar de residencia. Cuenta con grabaciones de varios estados de la República Mexicana, incluyendo el Estado de México, Guanajuato, Puebla y Yucatán. Para que un hablante sea considerado como parte de un estado es necesario que lleve más de 20 años en la localidad, ya sea que sea que haya nacido ahí o se haya mudado en algún momento.

El protocolo de grabación del corpus consiste en cinco partes: fonación, lectura de un texto fonéticamente balanceado, habla semi-espontánea, entonación y map task.

Para esta investigación se utilizaron grabaciones pertenecientes a hombres de 20 a 34 años con nivel de instrucción alto cuyo lugar de nacimiento y residencia fuera la Ciudad de México, dicha selección de hablantes se debió a que estadísticamente es más probable que una muestra de voz utilizada en el ámbito forense pertenezca a un hombre. Los parámetros de edad, nivel de instrucción y lugar de residencia fueron seleccionados así por practicidad, al ser los rubros del corpus que contaban con el mayor número de hablantes.

Se tomaron muestras de habla de 15 hablantes del apartado de habla semi-espontánea. Por cada hablante se seleccionaron diez oraciones declarativas neutras que terminaran en palabra grave, tanto oraciones finales como no finales.

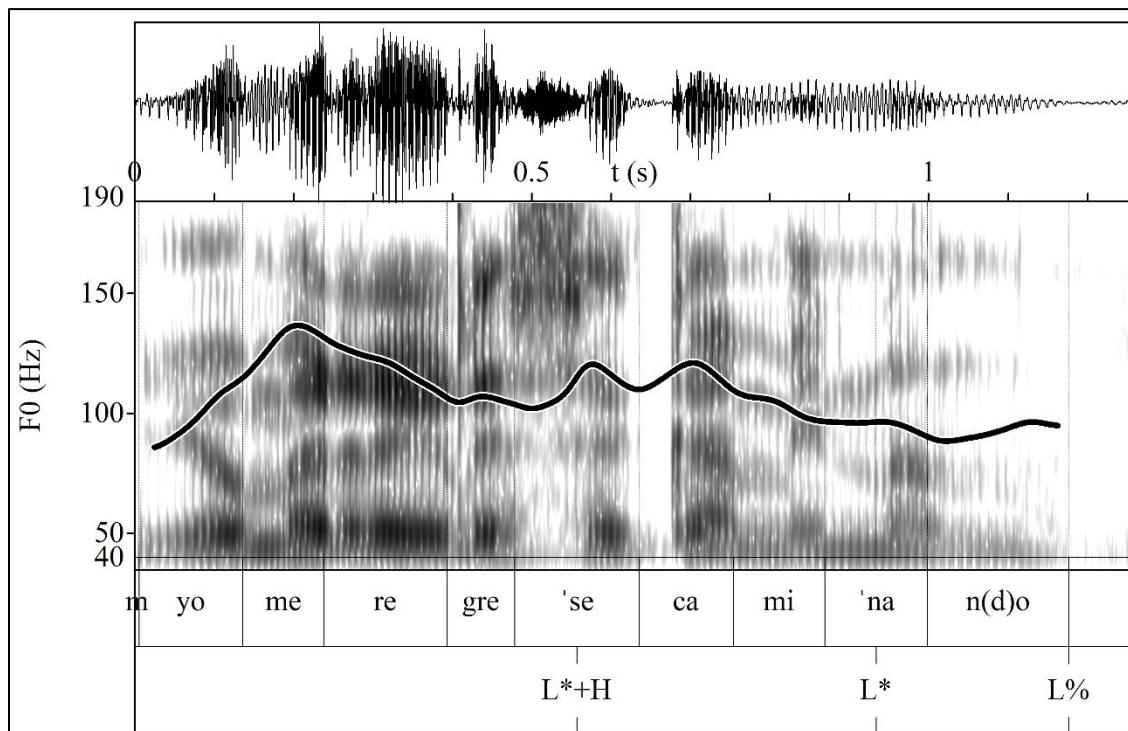
Orozco (2016) define las oraciones declarativas neutras o declarativas de foco amplio como aquellas que presentan información que es nueva para el oyente. (p.19). Se decidió usar esta terminología al ser la propuesta por De la Mota (et. al., 2010) en su estudio *Mexican Spanish*

intonation.

Por otro lado, las oraciones finales (Figura 5.1.1) refieren a aquellas que produce el hablante con una curva entonativa descendente que permite al oyente entender que la persona ha terminado su turno de habla.

Figura 5.1.1

Ejemplo de oración final.

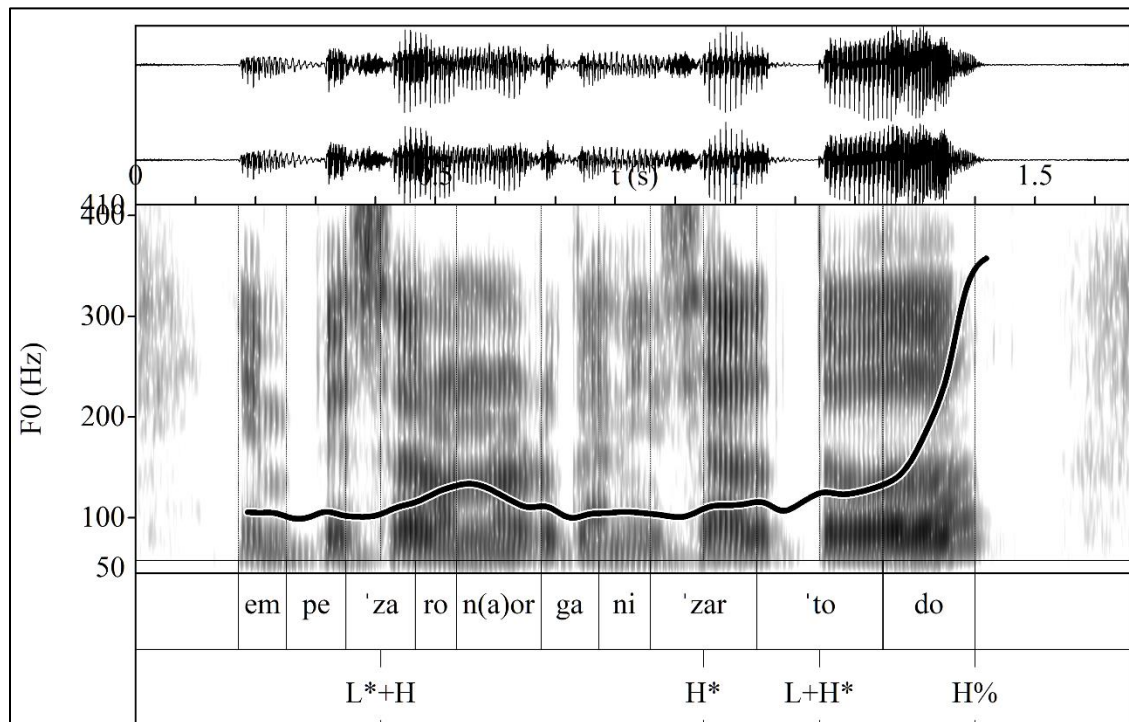


En contraste, las oraciones no finales (Figura 5.1.2) tienen un contorno entonativo ascendente que deja saber al oyente que la persona va a continuar hablando y se producen al inicio o a la mitad del turno de habla.

Se buscó que ambos tipos de oraciones terminaran en palabra grave ya que esto facilita percibir las diferencias presentes en el acento sintagmático y la cola.

Figura 5.1.2

Ejemplo de oración no final.



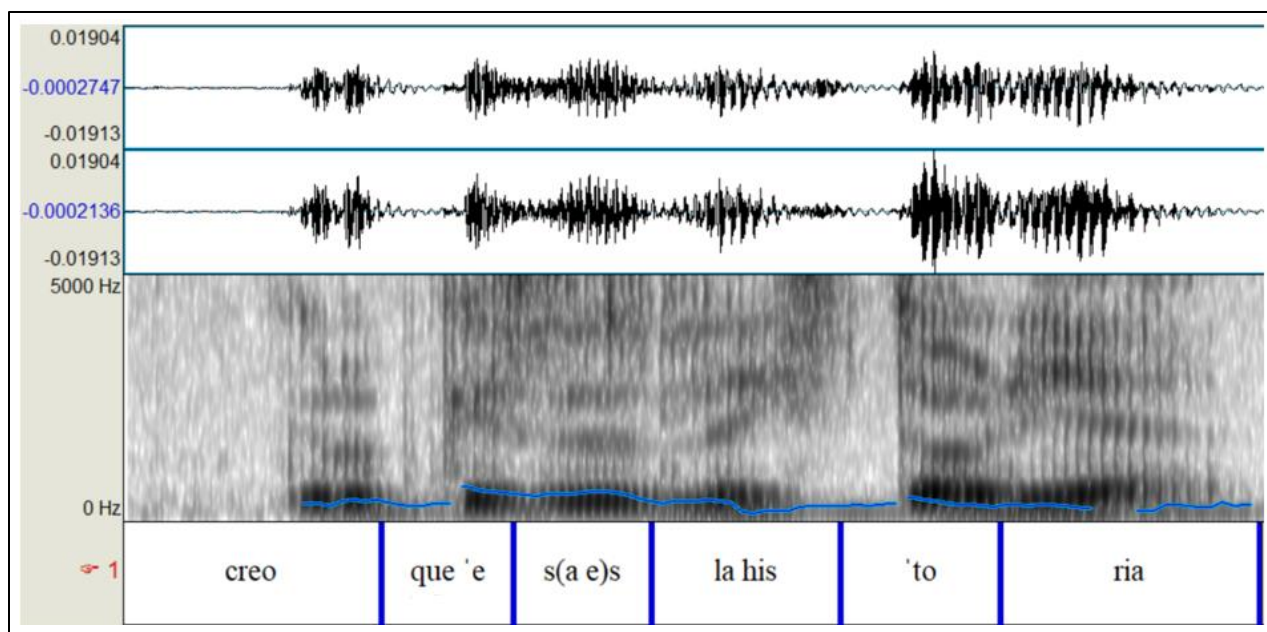
5.2 Análisis de datos

Las diez oraciones seleccionadas para cada hablante fueron introducidas al programa de análisis fonético Praat², donde fueron etiquetadas silábicamente de forma manual. La figura 5.2.1 muestra un ejemplo de dicho etiquetado.

Figura 5.2.1

Ejemplo de etiquetado silábico de una oración en Praat.

² Praat para Windows (versión 6.1.09 del 20 de enero del 2020).



Posteriormente, en el mismo programa, se utilizó un script desarrollado por Henning Reetz ³ del Instituto de Fonética de la Universidad de Frankfurt con la finalidad de obtener la media de la frecuencia fundamental de cada sílaba en Hertz. Con estos datos, se seleccionaron cinco elementos por cada oración: el inicio, los dos acentos intermedios más prominentes, el núcleo oracional o acento sintagmático y la cola. En el caso de que se tratara de una oración corta y faltaran acentos, se seleccionó la sílaba átona dentro de la oración que tuviera el valor más elevado de F0. Una vez hecho esto se generaron archivos con datos como el que se muestra en la tabla 5.2.1. Las celdas sombreadas corresponden a los cinco elementos seleccionados para análisis de la oración en cuestión.

Tabla 5.2.1

Ejemplo de datos seleccionados.

³ Mean_pitch.praat (versión 1.1 del 22 de junio del 2009).

Oración	Etiqueta	Media F0 [Hz]	DevEs F0 [Hz]
H14-10	creo	108.1	.
	que 'e	114.4	1
	s(a e)s	122.3	3
	la his	103.2	7.9
	'to	106.1	4.8
	ria	101.2	25.3

Estos cinco datos fueron introducidos al programa Curve Expert⁴ para generar un análisis de regresión. El propósito del análisis de regresión es ajustar un modelo matemático a un conjunto de datos y determinar si existe relación entre una variable dependiente y una o más variables independientes. Para esta investigación, se tomó como variable dependiente al hablante y como variables independientes los cinco elementos de la oración.

La ventaja del Curve Expert es que, además de ser una herramienta gratuita, permite ajustar diferentes modelos a los datos. Después de probar varios, se eligió una regresión polinomial de grado cuatro al ser la que mejor se ajustó a la curva de las oraciones individuales. La ecuación del modelo es $y = a + b*x + c*x^2 + d*x^3 + e*x^4$.

En una regresión no lineal de este tipo, la dependencia entre la variable y las independientes ($x_1 \dots x_k$) no es lineal, por lo que las variables independientes originales se elevan a una potencia para añadir curvatura al modelo. En el modelo polinomial de grado cuatro se elevan hasta la cuarta potencia (x, x^2, x^3, x^4) para lograr un mejor ajuste.

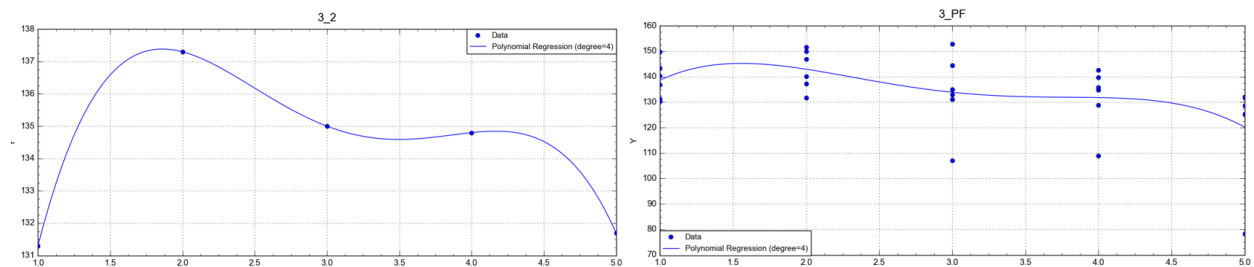
⁴ Curve Expert para Windows (versión 2.7.3 del 17 de octubre del 2020)

Asimismo, con Curve Expert se generaron curvas prototípicas y con ellas se hicieron gráficas de líneas con la finalidad de comparar y formar grupos con hablantes que se comportaran de forma similar. Para obtener las curvas prototípicas de los hablantes, se introdujeron los valores de todas las oraciones de un hablante y se ajustaron a la regresión polinomial de grado cuatro. Se generó una oración prototípica final y una no final para cada hablante.

La figura 5.2.2 muestra, del lado izquierdo, un ejemplo de la gráfica arrojada por el programa generada a partir de una oración individual, correspondiente al hablante número 3, en la que dice “todo estaba súper mal organizado”. Aquí podemos ver el ajuste que presenta el modelo con oraciones individuales. Del lado derecho se ve el ajuste cuando ya no es una oración, sino 6 oraciones de un mismo hablante. Esta es la curva prototípica final que se genera para las oraciones del hablante 3. Como es de esperarse, el ajuste de las oraciones prototípicas es menor que aquel de las oraciones individuales.

Figura 5.2.2

Ejemplo del ajuste del modelo con una oración y con más de una.



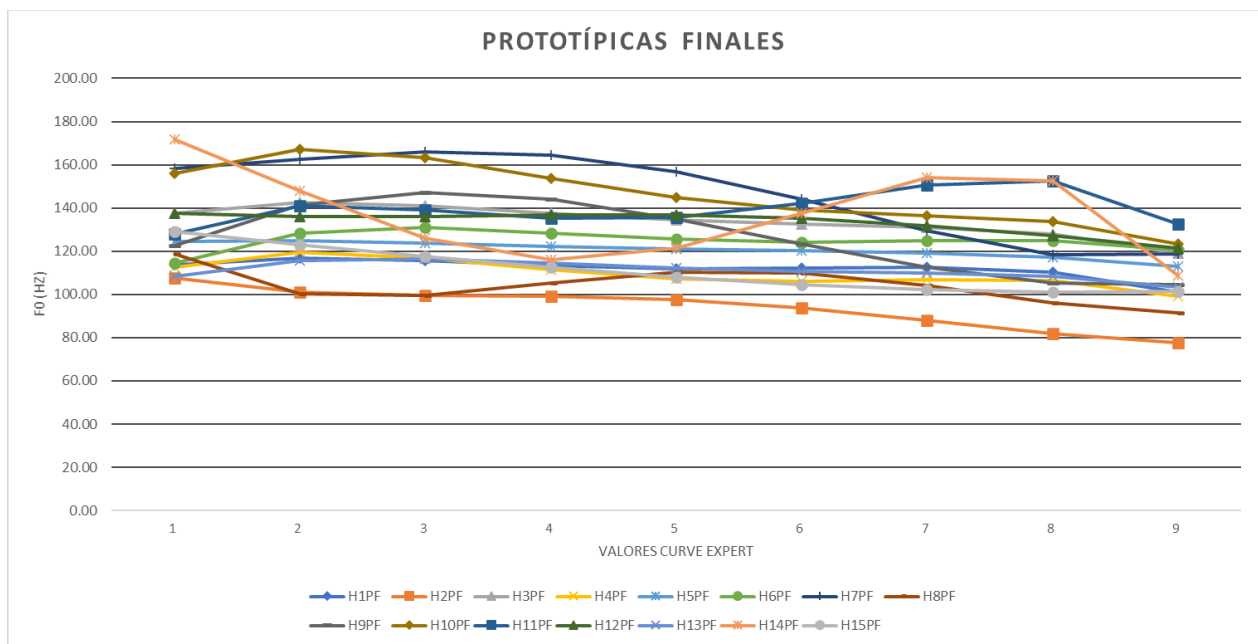
El siguiente paso consistió en tomar la curva entonativa prototípica de cada hablante para generar una gráfica de líneas que representara el comportamiento en la entonación de las oraciones de los diferentes hablantes y con ella comparar los contornos entonativos de la población total utilizada para este estudio. Esto se realizó haciendo uso de los valores ajustados a la regresión polinomial

de grado cuatro arrojados por Curve Expert al introducir todas las oraciones de un mismo hablante. Las gráficas fueron divididas, nuevamente, en contornos finales y no finales.

La figura 5.2.3 muestra la gráfica con las curvas prototípicas finales de los 15 hablantes. Cada color representa la curva prototípica (P) de un hablante (H) en sus oraciones finales (F). De ahí que las etiquetas tengan el código: H corresponde a “hablante”, seguido por el número que le fue asignado a dicho hablante, una P que significa “prototípica” y por último podemos tener F correspondiente a “final” o NF “no final”.

Figura 5.2.3

Gráfica de líneas de oraciones prototípicas finales.



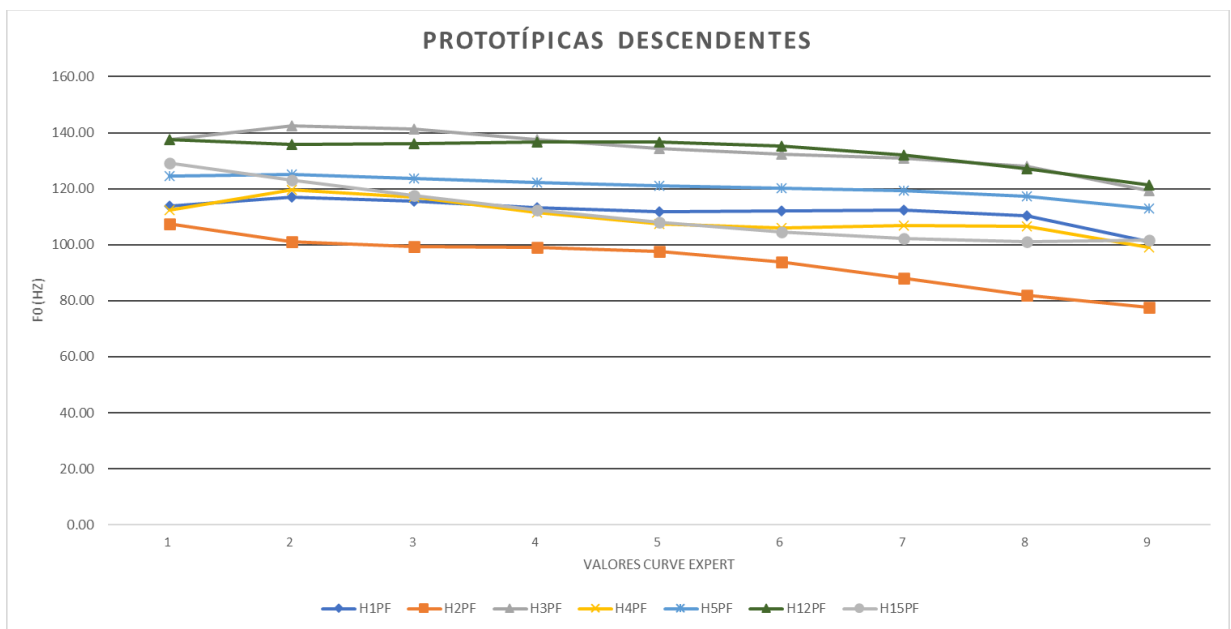
Las gráficas de líneas fueron comparadas visualmente buscando patrones de contornos entonativos. De esta confrontación se pudieron identificar tres grupos de hablantes que compartían contornos entonativos similares. En el campo de la lingüística forense, estos grupos corresponderían a la población relevante contra la cual se debe comparar la grabación dubitada

debido a que todos los miembros del grupo son del mismo sexo, estrato geográfico, nivel de instrucción y edad.

En el primero de estos grupos podemos observar el contorno descendente, uno de los más comunes en la Ciudad de México (ver figura 5.2.4). El grupo estaba compuesto por 7 hablantes, lo que representa el 46.6% de la población total utilizada en el estudio. Este contorno se caracteriza por tener un constante descenso a partir de su inicio. La figura 5.2.4 muestra la gráfica que contiene los contornos entonativos de los hablantes que conforman este grupo.

Figura 5.2.4

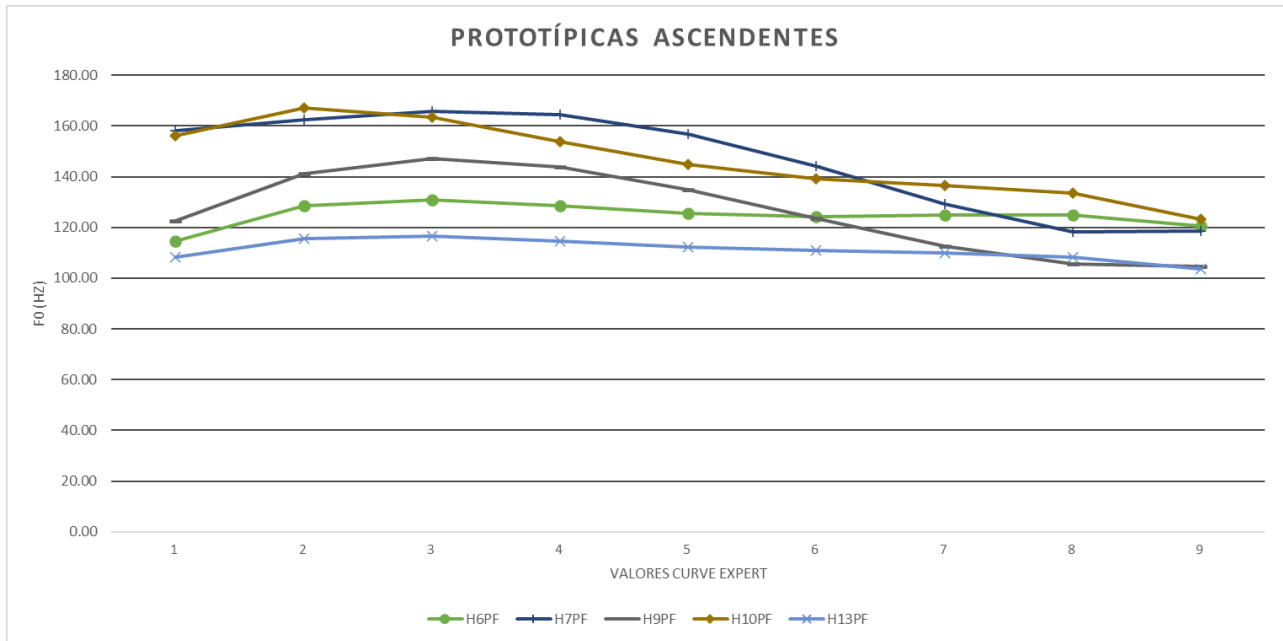
Gráfica de líneas de las oraciones prototípicas del grupo descendente.



El segundo grupo que se obtuvo se le denominó ascendente debido a que los hablantes presentaban un ascenso ligeramente pronunciado del inicio de la oración, visible a partir del segundo segmento para posteriormente descender. Este grupo estaba compuesto por 5 hablantes, lo que representa al 33.3% del total de la población utilizada en el estudio. Las curvas de los hablantes de este grupo se pueden ver en la figura 5.2.5.

Figura 5.2.5

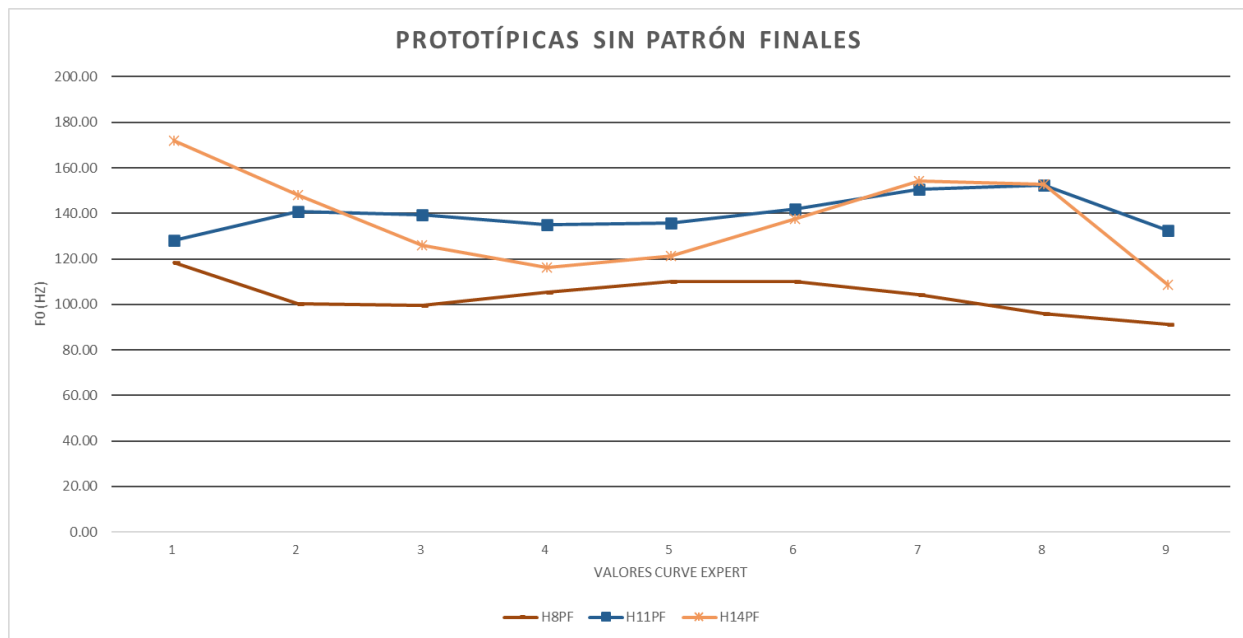
Gráfica de líneas de las oraciones prototípicas del grupo ascendente.



El último grupo dentro de las oraciones finales está compuesto por únicamente 3 hablantes, a este grupo se le denominó sin patrón debido a que las curvas de los que lo componen no se parecen a ninguno de los patrones vistos con anterioridad ni se parecen entre sí. Las oraciones finales prototípicas de estos hablantes presentan mayor variación al compararlas con el resto de la población utilizada en el estudio. Este grupo representa al 20% de la población. La figura 5.2.6 muestra la gráfica con las curvas entonativas de los hablantes de este grupo.

Figura 5.2.6

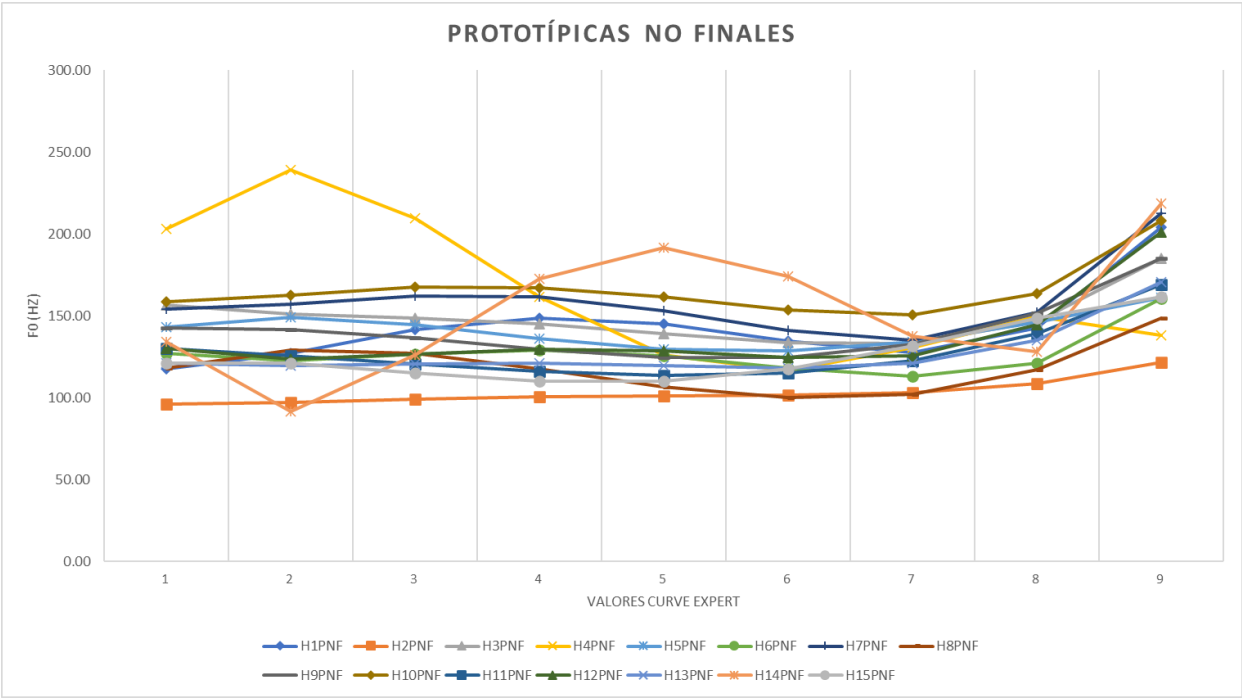
Gráfica de líneas de las oraciones prototípicas del grupo sin patrón finales.



Las oraciones no finales fueron comparadas, al igual que las finales, en una gráfica de líneas. La figura 5.2.7 muestra los contornos entonativos no finales prototípicos de los 15 hablantes que componen la población utilizada en el estudio.

Figura 5.2.7

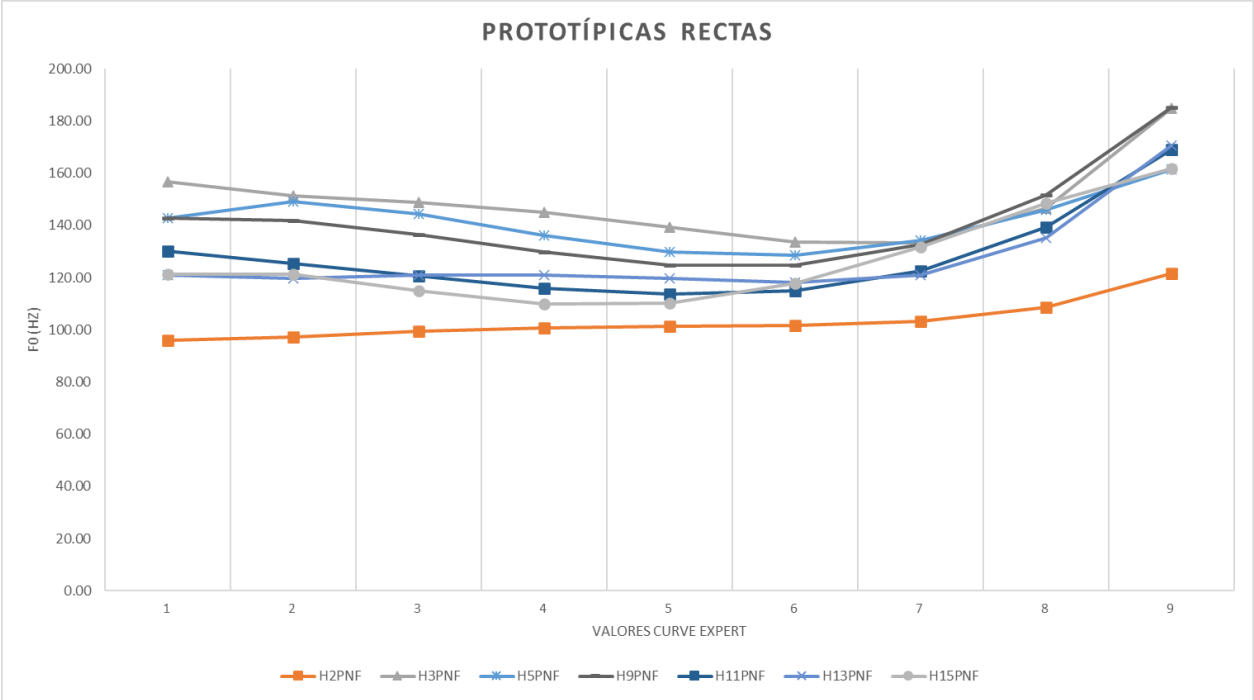
Gráfica de líneas de las oraciones prototípicas del grupo no finales.



Al primer grupo que se pudo observar de este tipo de oraciones se le denominó no finales rectas. El grupo está compuesto por 7 hablantes, lo que representa el 46.6% de la población utilizada en el estudio. Este contorno se caracteriza por mantenerse recto o presentar un ligero descenso desde el inicio de la oración hasta el acento oracional, a partir del cual comienza a subir. Este tipo de contorno es el más común en la Ciudad de México para las oraciones no finales. Este grupo de hablantes puede observarse en la figura 5.2.8.

Figura 5.2.8

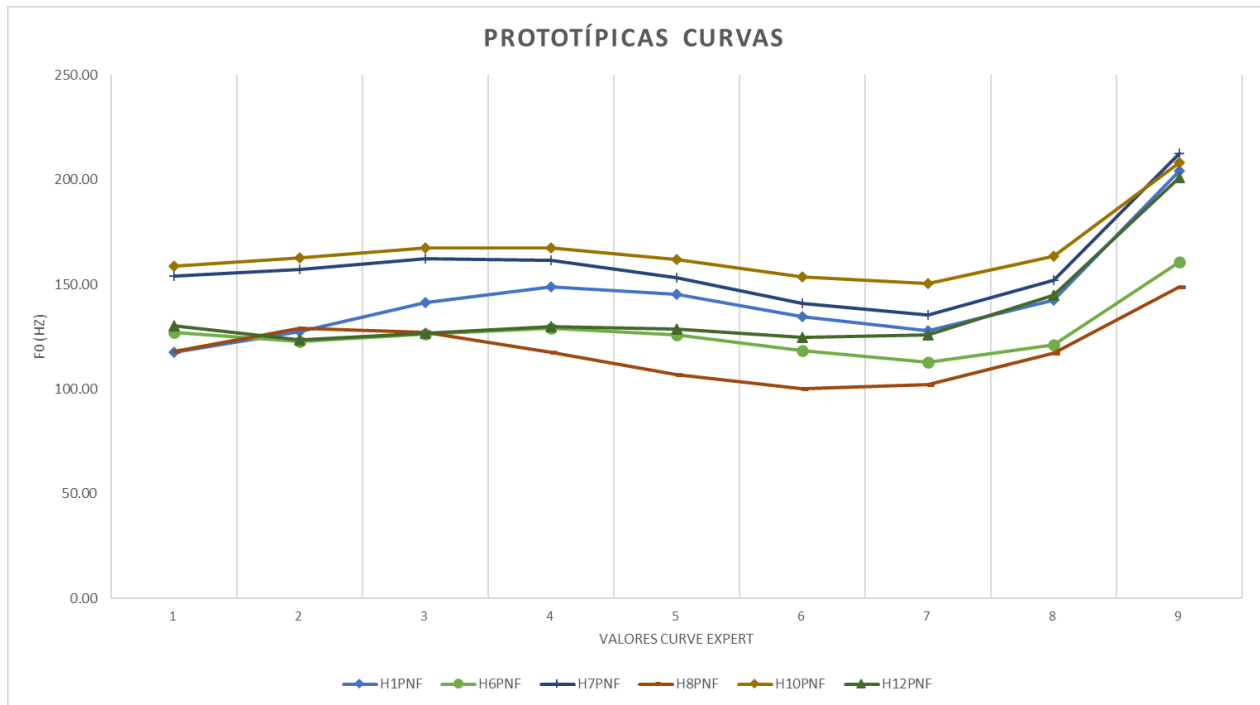
Gráfica de líneas de las oraciones prototípicas del grupo no finales rectas.



Al segundo grupo que se obtuvo se le denominó no finales curvas debido a que presentan un pequeño ascenso al inicio de oración para después descender y volver a subir a partir del acento oracional. Este grupo está compuesto por 6 hablantes, lo que representa el 40% de la población total utilizada en el estudio. Los contornos entonativos de este grupo pueden observarse en la figura 5.2.9.

Figura 5.2.9

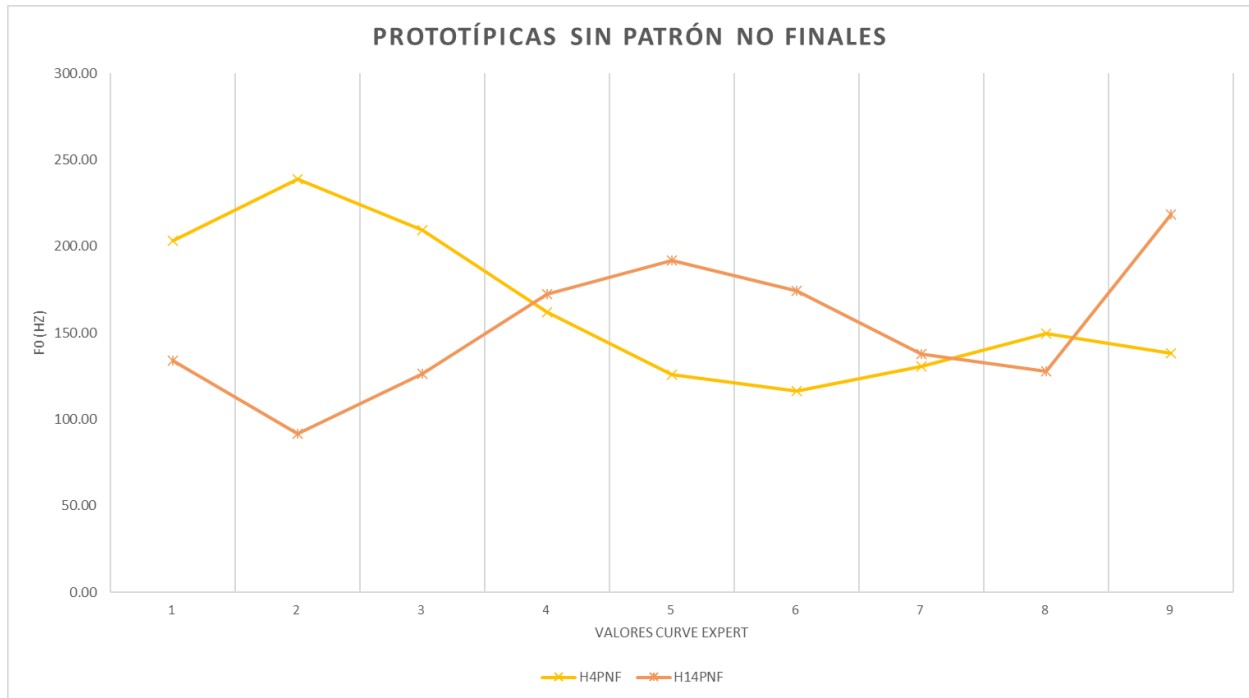
Gráfica de líneas de las oraciones prototípicas del grupo no finales curvas.



El último grupo de las oraciones no finales está compuesto por los hablantes cuyas curvas entonativas prototípicas no se ajustaron a los primeros dos grupos, este grupo fue denominado sin patrón. Las curvas entonativas que lo componen presentan una gran variación al ser comparadas con el resto de los hablantes utilizados en el estudio. El grupo está compuesto por únicamente 2 hablantes, lo que representa el 13.3% de la población del estudio. Los contornos entonativos de estos hablantes pueden observarse a continuación en la figura 5.2.10.

Figura 5.2.10

Gráfica de líneas de las oraciones prototípicas del grupo no finales sin patrón.



La tabla 5.2.2 incluye un resumen de los hablantes que conforman cada uno de los grupos vistos anteriormente.

Tabla 5.2.2

Integrantes de cada grupo.

Tipo de oraciones	Grupo	Hablantes en el grupo
Oraciones finales	Descendentes	H1, H2, H3, H4, H5, H12, H15
	Ascendentes	H6, H7, H9, H10, H13
	Sin patrón	H8, H11, H14
Oraciones no finales	Rectas	H2, H3, H5, H9, H11, H13, H15
	Curvas	H1, H6, H7, H8, H10, H12
	Sin patrón	H4, H14

Una vez agrupados los hablantes con base en el comportamiento de sus curvas prototípicas finales y no finales, se obtuvieron los cinco parámetros del modelo de regresión polinomial de grado cuatro, los cuales describen la curva de forma matemática, para cada una de las 10 oraciones seleccionadas de cada uno de los 15 hablantes. Estos parámetros se utilizaron para calcular las distancias entre las diferentes oraciones haciendo uso del programa R⁵.

Se generaron matrices de distancias por cada grupo de hablantes, en las que se representa la distancia entre todas las oraciones de los hablantes clasificados en dicho grupo. Asimismo, R proporciona una gráfica con los valores de las distancias que calcula.

Al realizar este cálculo, se observó que algunas oraciones eran altamente contrastivas, tanto al compararlas contra las otras oraciones que el mismo hablante había producido, como al compararlas contra los de otros hablantes del grupo. Debido a lo anterior, se hizo una revisión manual de dichas oraciones y se pudo concluir que la razón por la cual eran tan contrastivas era que, en el caso específico de las oraciones finales, se trataba de oraciones con un contorno entonativo circunflejo. En el caso de las oraciones no finales, las oraciones que presentaron un alto contraste no presentan un contorno específico, simplemente se observa que su patrón es distinto. Estas oraciones fueron omitidas en cada grupo y deben ser analizadas de forma separada.

Con estas gráficas se puede observar la variación intrahablante, es decir, la distancia que presentan las oraciones de un hablante entre sí mismas, la cual está enmarcada por un cuadro negro; y la variación interhablante, la cual es representada por la distancia que presenta un hablante al comparar sus oraciones con las del resto del grupo. Esto se puede observar al tomar el cuadro de cada hablante y compararlo de manera vertical u horizontal con las demás oraciones (ver figura 6.1.1. como ejemplo).

⁵ R para Windows (versión 4.0.2 del 22 de junio del 2020)

Para calcular la variación interhablante de cada hablante de los 6 grupos (ver la tabla 5.2.2), se extrajo al hablante en cuestión del cálculo. Lo anterior significa que, por ejemplo, para calcular la variación interhablante del hablante 1 del grupo de oraciones finales descendentes se calculó la media del grupo sin los valores del hablante 1. Esta media grupal se comparó con la media del hablante 1. Dicho proceso se repitió para todos los hablantes de todos los grupos.

El cálculo de distancias y la media son las medidas propuestas para resolver la problemática de calcular hasta qué punto se considera que la variación intrahablante o interhablante deja de formar parte de la variación estándar presentada por todos los hablantes y puede empezar a considerarse como poca o mucha variación. Los resultados que se muestran a continuación representan la puesta en práctica de mi propuesta metodológica.

6. Resultados

Las gráficas obtenidas de la matriz de distancias muestran de manera visual la distancia entre las oraciones de cada uno de los hablantes. Cuentan con etiquetas para identificar cada una de las oraciones, las cuales son visibles del lado izquierdo y en la parte superior de la gráfica. El primer número de las etiquetas corresponde al hablante en cuestión, mientras que el segundo corresponde a la oración de dicho hablante, es decir, la etiqueta “1_4” refiere la oración cuatro del hablante uno.

Cada gráfica contiene también una escala que muestra el color asignado a cada valor de distancia. Esta escala corresponde a las distancias presentes en la matriz, lo que significa que el rango de poca a mucha distancia depende de los datos con los que la matriz es generada. El rango de color de la escala se puede observar en los círculos, entre más grande y más azul sea un círculo, mayor es la distancia que existe entre las oraciones comparadas. Por el contrario, si el círculo es

pequeño y anaranjado, significa que la distancia es poca. En otras palabras, entre más distancia, mayor variación.

6.1 Oraciones finales

En este apartado se muestran las gráficas de la matriz de distancias, así como la media correspondiente a los 3 grupos de oraciones finales.

La figura 6.1.1 muestra la gráfica correspondiente al grupo de oraciones finales descendentes. Este grupo está compuesto por los hablantes 1, 2, 3, 4, 5, 12 y 15 (ver tabla 5.2.2). Recordemos que este es el contorno entonativo más utilizado en la Ciudad de México. La escala de distancias va desde el 0 hasta 1179, donde el 0 se obtiene al comparar una oración contra sí misma. En la gráfica se observa que los hablantes 2, 5 y 12 presentan menos variación intrahablante, mientras que los hablantes 1, 3, 4, y 15 presentan mayor variación intrahablante.

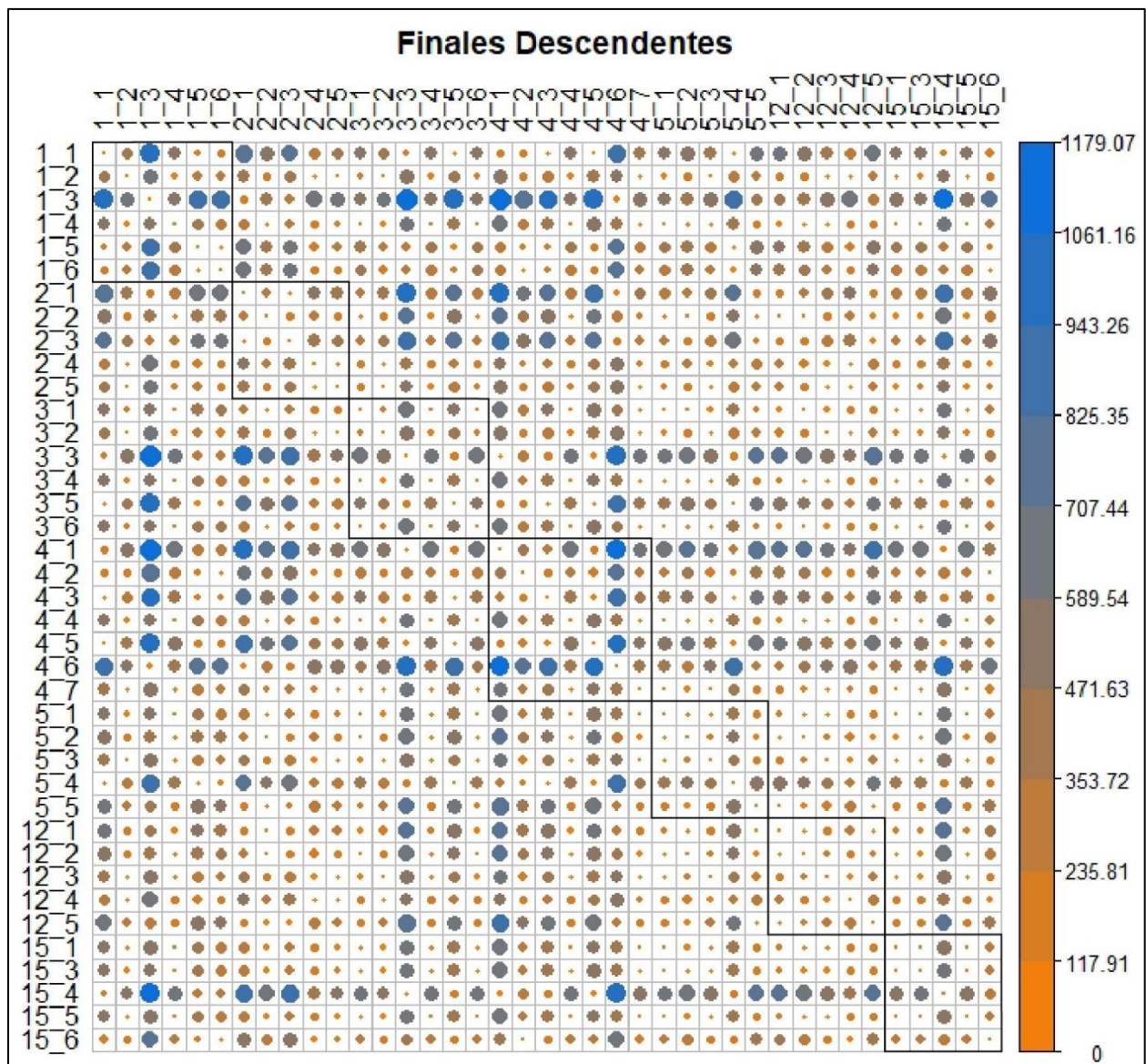
Con respecto a la variación interhablante, podemos ver que aquellos hablantes que presentan poca variación intrahablante, es decir los hablantes 2, 5 y 12, presentan también poco contraste al compararlos entre sí mismos. Por su parte, los hablantes con mucha variación intrahablante mantienen, de igual forma, un alto contraste contra el resto de los hablantes del grupo.

Por último, es necesario mencionar que se omitió una oración del grupo debido al alto contraste que presentaba contra el resto de las oraciones. Esta oración es la 15_2, la cual tiene un contorno entonativo circunflejo. La oración no se omitió del grupo de oraciones del hablante

utilizado para obtener la media debido a que se consideró que esta oración forma parte de la variación intrahablante estándar del hablante en cuestión.

Figura 6.1.1

Gráfica de distancias del grupo finales descendentes.



La tabla 6.1.1 que se muestra a continuación contiene una lista de los hablantes del grupo, seguida por la media del hablante en cuestión (H), la cual representa la variación intrahablante, y la media

del grupo sin hablante en cuestión (G), la cual representa la variación interhablante.

Se puede observar que los hablantes 2, 5, 12 y 15 presentan una media(H) menor a la del resto de los hablantes. Por su parte, los hablantes con valor de media (H) más altos son el hablante 1 y 4, lo cual se ve reflejado en la gráfica. La media (G) de este grupo es bastante homogénea, se puede observar que se encuentra en un rango de 322.95 a 328.79. El hecho de que al sacar un hablante la media (G) permanezca en el mismo rango significa que la variación interhablante es estable.

Tabla 6.1.1

Media individual y grupal del grupo finales descendentes.

Hablante	Media (H)	Media (G)
H1	431.52	324.52
H2	274.94	327.40
H3	325.73	326.71
H4	454.92	322.95
H5	260.96	327.59
H12	173.44	328.79
H15	294.20	327.13

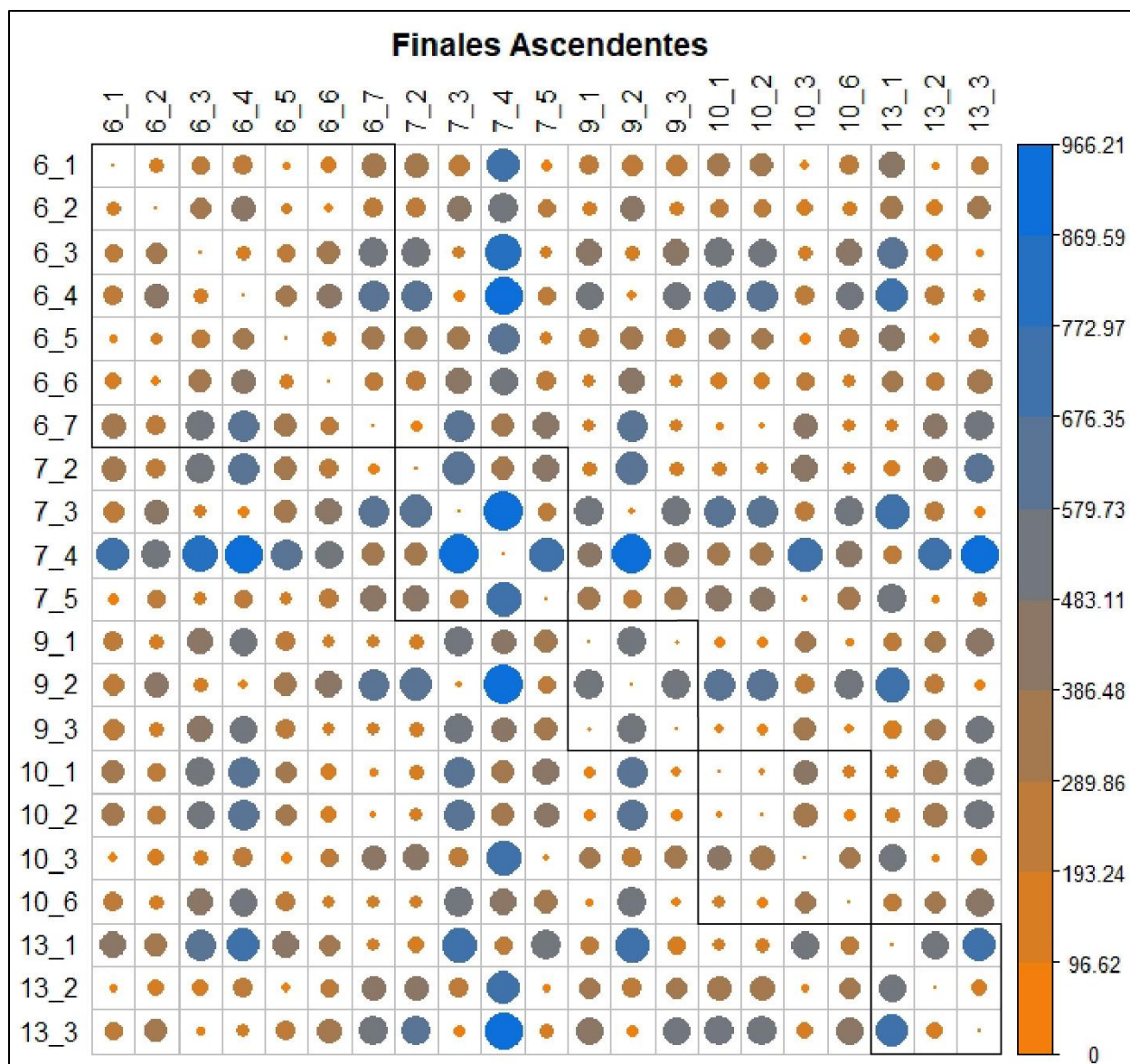
La gráfica correspondiente al grupo de oraciones finales ascendentes se muestra en la figura 6.1.2. Este grupo se compone por los hablantes 6, 7, 9, 10 y 13 (ver tabla 5.2.2). La escala correspondiente a las distancias presentadas entre los hablantes de este grupo tiene un rango del 0 al 966.21, siendo la escala con el rango más pequeño dentro del grupo de las oraciones finales. En la gráfica se

observa que los hablantes 6 y 10 presentan menos variación intrahablante que interhablante, mientras que los hablantes 7, 9 y 13 presentan más variación intrahablante e interhablante.

De este grupo se omitieron las oraciones 7_1, 10_4, 10_5 y 13_4 debido a que eran muy contrastivas. Al igual que en el grupo anterior, las oraciones omitidas resultaron ser circunflejas, dichas oraciones fueron tomadas en cuenta al momento de calcular la media de los hablantes que las produjeron.

Figura 6.1.2

Grafica de distancias del grupo finales ascendentes.



En la tabla 6.1.2 se observa que los hablantes 6 y 10 presentan las medias (H) más bajas del grupo. Con respecto a la media (G), los valores van de 319.87 a 333.66, lo que nuevamente muestra que la variación interhablante se mantiene estable.

Tabla 6.1.2

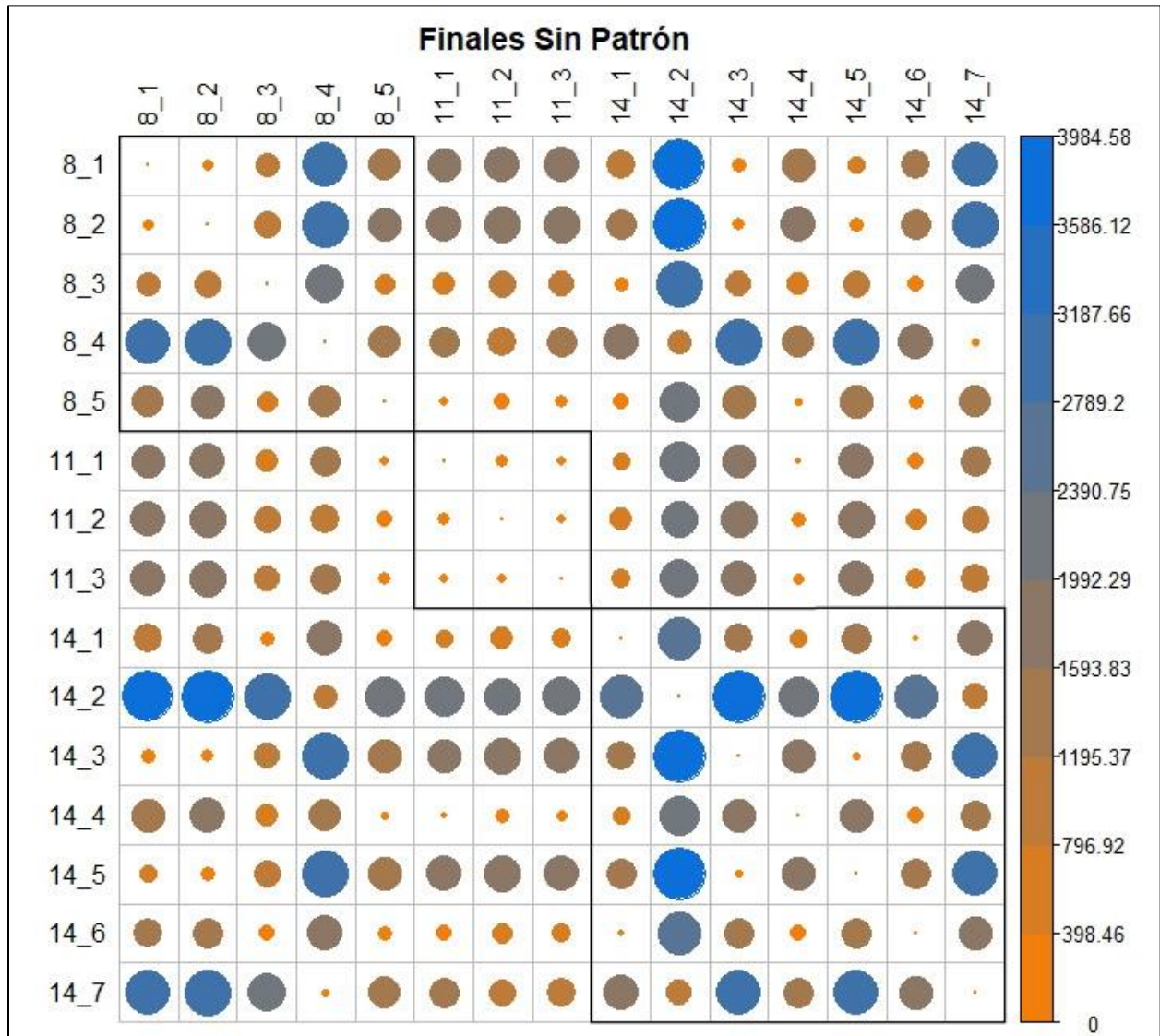
Media individual y grupal del grupo finales.

Hablante	Media (H)	Media (G)
H6	265.10	333.66
H7	562.46	319.87
H9	371.50	326.16
H10	218.94	329.98
H13	452.00	324.99

El tercer grupo de las oraciones finales denominado sin patrón está compuesto por los hablantes 8, 11 y 14 (ver tabla 5.2.2). Cabe recordar que este grupo se formó debido a que dichos hablantes presentaban curvas entonativas con poca similitud al resto de los hablantes. Debido a lo anterior, es el grupo que presenta mayor variación, esto se ve reflejado en la escala de la gráfica, la cual abarca un rango 0 a 3984.58. Siendo así la gráfica con la escala más grande dentro del grupo de las oraciones finales (ver figura 6.1.3). De los hablantes que componen este grupo, el hablante 11 muestra poca variación intrahablante y mayor variación interhablante, en contraste con los hablantes 8 y 14, los cuales muestran una gran variación tanto intrahablante como interhablante. No se omitieron oraciones de este grupo.

Figura 6.1.3

Gráfica de distancias del grupo finales sin patrón.



La tabla 6.1.3 presenta una lista de los hablantes del grupo, la media (H) y la media (G). El valor más bajo con respecto a la media (H) pertenecen al hablante 11. Sin embargo, la media (G) de este hablante es la más alta de los miembros del grupo. El hablante 8 presenta una media (H) más alta que la media (G) correspondiente. Finalmente, el hablante 14 presenta la media (H) más alta del grupo y la media (G) más baja.

A diferencia de los dos grupos anteriores, en este se aprecian diferencias en la media (G) correspondiente a cada hablante, abarcando un rango de 1969.13 a 2143.76, es decir, 174.63 puntos.

Tabla 6.1.3

Media individual y grupal del grupo finales sin patrón.

Hablante	Media (H)	Media (G)
H8	2288.39	2068.51
H11	242.68	2143.76
H14	2570.73	1969.13

6.2 Oraciones no finales

En este apartado se hablará de las gráficas de la matriz de distancias, así como la media correspondiente a los 3 grupos de oraciones no finales.

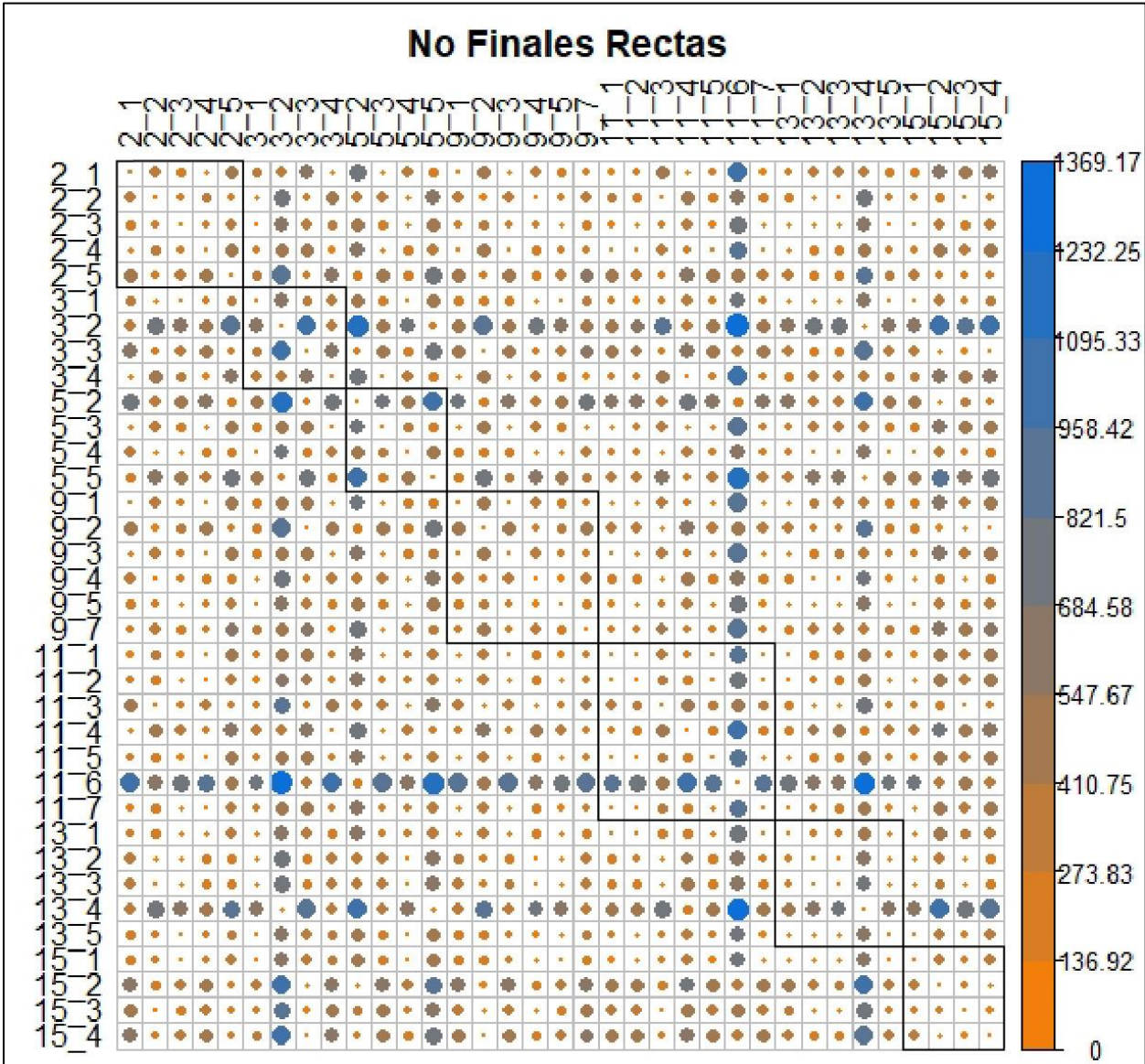
El primer grupo de las oraciones no finales es el de las oraciones rectas, el contorno más utilizado en la Ciudad de México. Dicho grupo está compuesto por los hablantes 2, 3, 5, 9, 11, 13 y 15 (ver tabla 5.2.2). La escala de distancias abarca un rango de 0 a 1369.17. Los cuadros de los hablantes 2, 9 y 15 presentan menos variación intrahablante, mientras que los hablantes 3, 5, 11, y 12 presentan mayor variación intrahablante.

Con respecto a la variación interhablante, podemos ver que los hablantes 2 y 9 presentan poca variación interhablante al compararlos con el resto de los hablantes y, en especial, al compararlos entre ellos. Por su parte, los hablantes con mucha variación intrahablante, es decir, los hablantes 3, 5, 11, y 12 mantienen su nivel de variación contra los demás hablantes del grupo.

De este grupo se omitieron las oraciones 5_1 y 9_6 debido al alto contraste que presentaban contra el resto de las oraciones, como se mencionó en el capítulo anterior. Las oraciones no fueron omitidas del grupo de oraciones de cada hablante utilizado para obtener la media debido a que se consideró que estas oraciones forman parte de la variación intrahablante estándar del hablante en cuestión. Lo mencionado anteriormente se puede ver en la figura 6.2.1.

Figura 6.2.1

Gráfica de distancias del grupo no finales rectas.



La tabla 6.2.1 presenta una lista de los hablantes del grupo, seguida por la media de cada hablante (H) y la media del grupo sin el hablante en cuestión (G). Los valores más bajos con respecto a la media (H) pertenecen a los hablantes 2, 9 y 15. Por otro lado, los valores más altos corresponden a los hablantes 3 y 5.

Este grupo presenta estabilidad en la media (G), conservándose esta en un rango de 283.92 a 286.67, lo que significa que es estable.

Tabla 6.2.1

Media individual y grupal del grupo no finales rectas.

Hablante	Media (H)	Media (G)
H2	239.00	286.41
H3	452.56	283.92
H5	451.52	283.93
H9	222.18	287.26
H11	309.86	284.73
H13	264.16	285.98
H15	182.01	286.67

El segundo grupo es el de las no finales curvas, dicho grupo está compuesto por los hablantes 1, 6, 7, 8, 10 y 12 (ver tabla 5.2.2). La escala de distancias abarca un rango de 0 a 1420.09. Los cuadros de los hablantes 6 y 12 presentan menos variación intrahablante, mientras que los hablantes 1, 7, 8, y 10 presentan mayor variación intrahablante.

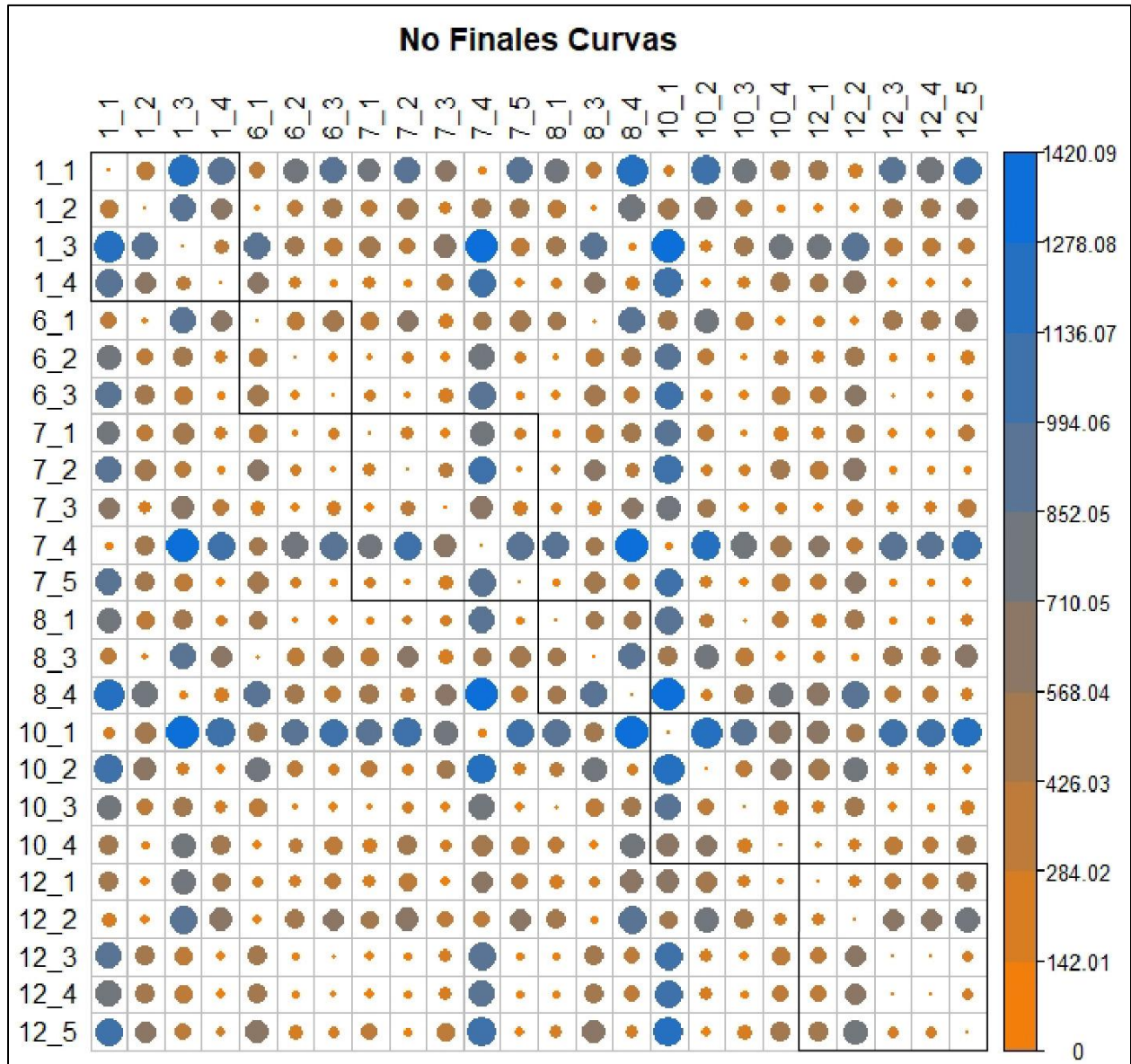
Con respecto a la variación interhablante, podemos ver que los hablantes 6 y 12 presentan poca variación interhablante al compararlos con el resto de los hablantes y contra ellos mismos. Por su parte, los hablantes con mucha variación intrahablante son el 1, 7, 8, y 10, quienes también mantienen su nivel de variación contra los demás hablantes del grupo.

De este grupo se omitieron las oraciones 8_2 y 8_5 debido al alto contraste que presentaban contra el resto de las oraciones, como se mencionó en el capítulo anterior. Las oraciones no fueron omitidas del grupo de oraciones de cada hablante utilizado para obtener la media y la desviación

estándar, debido a que se consideró que estas oraciones forman parte de la variación intrahablante estándar del hablante en cuestión. Esto puede observarse en la figura 6.2.2.

Figura 6.2.2

Gráfica del grupo no finales curvas.



La tabla 6.2.2 presenta una lista de los hablantes del grupo de oraciones no finales curvas, seguida por la media de cada hablante (H) y la media del grupo sin hablante en cuestión (G). Los valores

de la media (H) más bajos son de los hablantes 6 y 12. Por otra parte, los valores de media (H) más altos del grupo pertenecen a los hablantes 1 y 10.

Este grupo presenta diferencias mayores con respecto a la media (G) correspondiente a cada hablante comparándolo con el grupo anterior. Dichos valores se encuentran en un rango de 444.09 a 453.23, lo que se puede interpretar como valores estables.

Tabla 6.2.2

Media individual y grupal del grupo no finales.

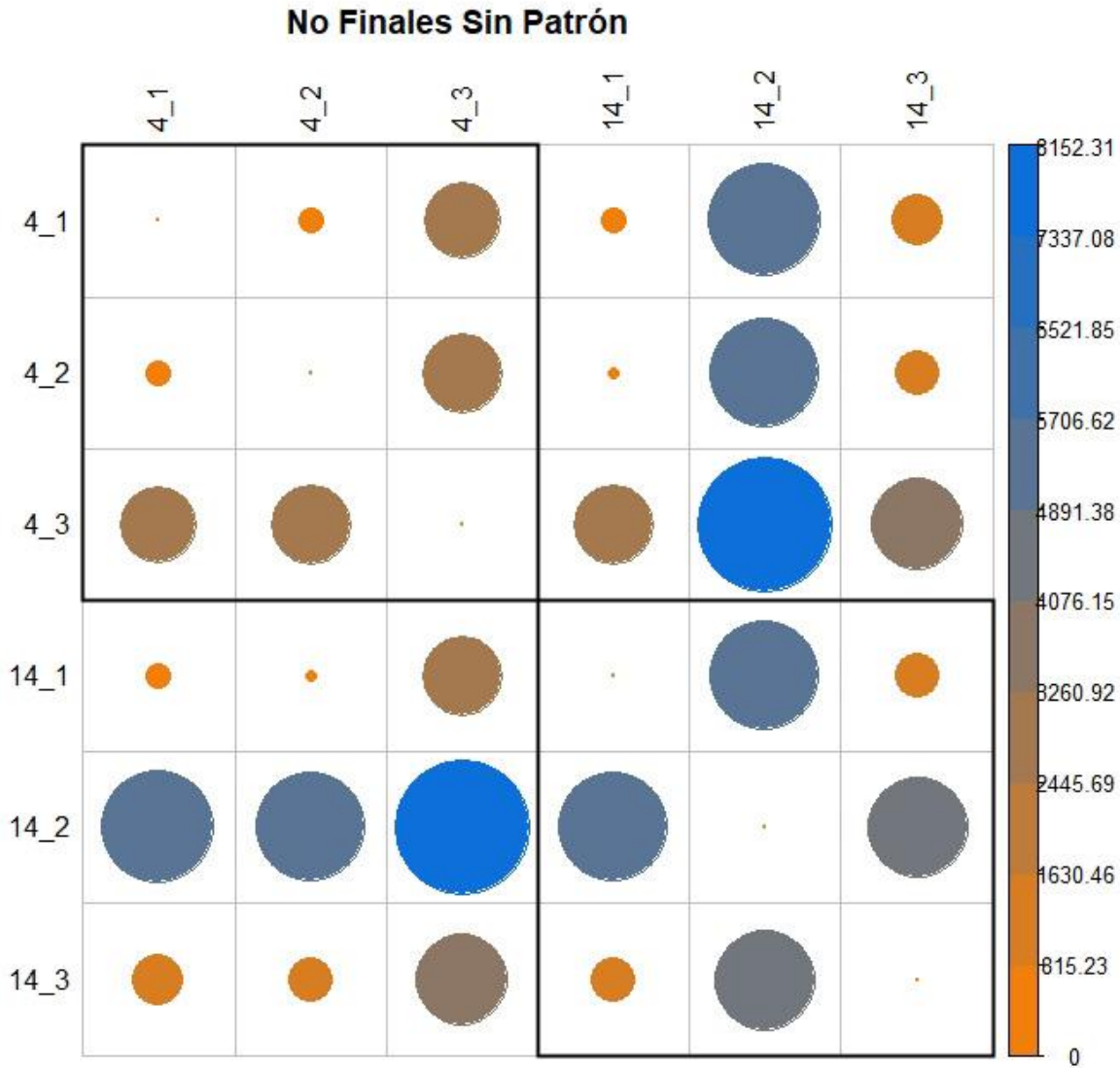
Hablante	Media (H)	Media (G)
H1	724.90	444.09
H6	362.00	451.16
H7	464.28	449.66
H8	583.50	448.73
H10	660.22	445.53
H12	369.33	453.23

El tercer grupo es el de las no finales sin patrón. Este grupo está compuesto por los hablantes 4 y 14. La escala de distancias abarca un rango de 0 a 8152.31, el valor más alto de todos los grupos de esta investigación. El cuadro del hablante 4 presenta menos variación intrahablante que el del hablante 14.

Con respecto a la variación interhablante, podemos ver que ambos hablantes presentan mucha variación interhablante al compararlos entre sí. Su gráfica se puede ver en la figura 6.2.3.

Figura 6.2.3

Gráfica del grupo no finales sin patrón.



La tabla 6.2.3 presenta una lista de los hablantes del grupo de oraciones no finales sin patrón, seguida por la media de cada hablante (H) y la media del grupo sin hablante en cuestión (G). Hay que recordar que los hablantes que conforman este grupo presentan una variación intrahablante e interhablante alta, lo cual se puede ver reflejado en los valores numéricos de la media (H) y la media (G).

De los dos hablantes que componen el grupo, el valor de media (H) más bajo corresponde al hablante 4 equivalente a 1919.38. Por su parte, la media (H) del hablante 14 tiene un valor de 3524.03.

Este grupo sí presenta diferencias entre la media (G) correspondiente a cada hablante, al igual que el grupo de sin patrón finales. En el caso del hablante 4, la media (G) tiene un valor de 3212.46 y para el hablante 14 es de 2811.30.

Tabla 6.2.3

Media individual y grupal del grupo no finales sin patrón.

Hablante	Media (H)	Media (G)
H4	1919.38	3212.46
H14	3524.03	2811.30

7. Análisis de resultados

El objetivo principal de esta investigación consiste en explorar el comportamiento del contorno entonativo de las oraciones declarativas neutras del español de la Ciudad de México como posible caracterizador de hablantes en el ámbito forense, dicho objetivo fue alcanzado parcialmente al demostrar que ciertos hablantes presentan una variación intrahablante menor a la variación interhablante.

La manera en la que evaluamos si la curva entonativa es un rasgo que podría servir como identificador de hablantes es que cumpla uno de los principios de la lingüística forense que dice que la variación intrahablante debe ser menor que la variación interhablante. Para medir este principio, se propuso calcular las distancias entre las oraciones de un mismo individuo y las

oraciones entre diferentes individuos.

Tomemos como ejemplo a los hablantes 2, 5, 12 y 15 del grupo de oraciones finales descendentes. En la gráfica de distancias se pueden distinguir los grupos de oraciones de cada uno de estos hablantes debido a la poca variación intrahablante que poseen. En los cuatro casos, la distancia entre las oraciones producidas por el hablante es reducida. En contraste, al comparar sus contornos contra los del resto de los hablantes se puede observar que la distancia es mayor, especialmente al compararlos contra los hablantes 1, 3 y 4 (Ver figura 6.1.1). Lo anterior también puede observarse de forma numérica en la media (H) y la media (G).

En el caso del hablante 2, su media (H) tiene un valor de 274.94 mientras que su media (G) es de 327.40. Para el hablante 5 la media (H) es de 260.96 y la media (G) de 327.59. Por su parte, el hablante 12 tiene una media (H) de 173.44 y una media (G) de 328.79. Finalmente, el hablante 15 presenta una media (H) de 294.20 y una media (G) de 327.13. Los datos anteriores nos permiten llegar a conclusiones más objetivas, ya que todos los hablantes tienen una media (G) más alta que su media (H). Por lo tanto, se cumple uno de los criterios de identificación: la variación intrahablante debe ser menor que la variación interhablante. Con estos datos se puede concluir que el hablante que es más consistente en sus producciones es el hablante 12 con una diferencia de 155.35 entre su media (H) y su media (G).

Es importante recordar que los grupos utilizados en la investigación son homogéneos con la finalidad de asemejarse a una población relevante. Todos los hablantes son hombres que nacieron en la Ciudad de México, tienen una edad de 20 a 35 años y cuentan con estudios de licenciatura. Adicionalmente, fueron separados en 3 grupos dependiendo el tipo de contorno entonativo que presentaban e incluso bajo estas circunstancias es posible hacer distinciones entre ellos. Dicho en otras palabras, es posible identificar a ciertos hablantes que presentan una variación

intrahablante pequeña dentro de un grupo de hablantes con un comportamiento similar en sus contornos entonativos y que comparten un trasfondo sociocultural parecido.

Los hablantes de todos los grupos de análisis que pueden ser identificados con este método debido a las diferencias entre su variación intrahablante e interhablante están enlistados en la tabla 7.1.

Tabla 7.1

Hablantes identificables.

Grupo	Hablante	Media (H)	Media (G)
Finales descendentes	2	274.94	327.40
	5	260.96	327.59
	12	173.44	328.79
	15	294.20	327.13
Finales ascendentes	6	265.10	333.66
	10	218.94	329.98
Finales sin patrón	11	242.68	2143.76
No finales rectas	2	239.00	286.41
	9	222.18	287.26
	13	264.16	285.98
	15	182.01	286.67
No finales curvas	6	362.00	451.16
	12	369.33	453.23

No finales sin patrón	4	1919.38	3212.46
-----------------------	---	---------	---------

Por otro lado, tenemos hablantes que no pueden ser identificados debido a que su variación intrahablante es mayor que su variación interhablante o porque los valores de ambas son cercanos, lo que significaría que la variación intrahablante y la variación interhablante son similares. Muestra de lo anterior serían los hablantes 1,3 y 4 del grupo de oraciones finales descendentes. (Ver figura 6.1.1). La lista de estos hablantes puede verse en la tabla 7.2.

Tabla 7.2

Hablantes no identificables.

Grupo	Hablante	Media (H)	Media (G)
Finales descendentes	1	431.52	324.52
	3	325.73	326.71
	4	454.92	322.95
Finales ascendentes	7	562.46	319.87
	9	371.50	326.16
	13	452.00	324.99
Finales sin patrón	8	2288.39	2068.51
	14	2570.73	1969.13
No finales rectas	3	452.56	283.92
	5	451.52	283.93
	11	309.86	284.73

No finales curvas	1	724.9	444.09
	7	464.28	449.66
	8	583.50	448.73
	10	660.22	445.53
No finales sin patrón	14	3524.03	2811.30

Adicionalmente, utilizando la división de contornos necesaria para formar grupos se puede observar que, a nivel poblacional, los hablantes utilizan los contornos predominantes en la Ciudad de México con mayor frecuencia, estos son el contorno final descendente y el no final recto. En el caso de las oraciones finales, 7 de los 15 hablantes utilizados en el estudio producen un contorno descendente en sus oraciones, 5 tienen un contorno ascendente y 3 se encuentran fuera de los patrones comunes. Por su parte, para las oraciones no finales, 7 hablantes presentan un contorno recto, 6 un contorno curvo y sólo dos hablantes quedan fuera de los patrones comunes. Lo anterior puede verse de forma detallada en la tabla 7.3.

Tabla 7.3

Porcentaje del uso de los contornos a nivel poblacional.

Tipo de contorno	Número de hablantes que lo producen	Porcentaje a nivel poblacional
Final descendente	7	46.7%
Final ascendente	5	33.3%
Final sin patrón	3	20.0%

No final recto	7	46.7%
No final curvo	6	40.0%
No final sin patrón	2	13.3%

Otro aspecto interesante relacionado a la tipicidad de los diferentes contornos entonativos utilizados en este trabajo es que 4 de los hablantes producen tanto el contorno prototípico final, es decir, el contorno descendente, como el prototípico no final, el cual se denominó recto. Los hablantes en cuestión son el 2, 3, 5 y 15. Esto significa que un 26.7% de los hablantes utilizados en el estudio produce las configuraciones entonativas más comunes de la Ciudad de México. Asimismo, tres de los cuatro hablantes presentan poca variación intrahablante (2, 5 y 15) en sus contornos entonativos finales y dos presentan poca variación intrahablante (2 y 15) en sus contornos entonativos no finales (ver tabla 7.1).

Como objetivo específico de esta investigación se busca presentar un modelo de análisis de los contornos entonativos que funcione en el ámbito forense al calcular la variación intrahablante e interhablante de los hablantes de una población relevante. Es necesario que este modelo sea objetivo y fácil de replicar sin depender en gran medida del investigador que lo use. Dicho objetivo se cumplió parcialmente debido a que el modelo de análisis calcula la variación intrahablante e interhablante en una población relevante de manera exitosa en aquellos casos en los que el hablante cumple con el principio de variación, muestra de esto puede encontrarse al observar a los hablantes presentes en la tabla 7.1 en las gráficas de distancias de sus grupos. Cada uno de estos hablantes puede identificarse de manera visual sin necesidad de otros análisis.

Sin embargo, para cumplir con el criterio de objetividad se necesitó buscar una forma de llevar lo observado en las gráficas de distancias a cantidades numéricas haciendo uso de la

estadística. La medida que reflejó de manera fehaciente lo presentado en estas gráficas fue la media, al calcularla podemos utilizar los valores numéricos para evitar la subjetividad inherente en la percepción de poca o mucha variación intrahablante sin tener datos que respalden dicha opinión.

Recordemos que para esta investigación se utilizó la media de las distancias entre las oraciones producidas por un mismo hablante como su variación intrahablante y la media del grupo sin el hablante en cuestión como variación interhablante. Para determinar si la variación intrahablante era menor que la interhablante, y por lo tanto determinar que el contorno entonativo puede ser un rasgo para identificar hablantes, se buscó que la media de las oraciones de un hablante fuera menor a aquella de las oraciones del grupo. Si, por otro lado, la media del hablante era igual o mayor a la media del grupo, el principio no se cumple.

Finalmente, el tercer criterio que debía cumplir el modelo de análisis era la facilidad de ser replicado sin que los resultados dependan en gran medida del investigador. Lo anterior se cumplió debido a que el modelo requiere una participación mínima por parte del investigador ya que se lleva a cabo utilizando scripts de *Praat* y *R* ya existentes, sin necesidad de que éste los modifique o interprete. Adicionalmente, al momento de utilizar *Curve Expert* sólo se necesita seleccionar la regresión correcta tomando como base el porcentaje de ajuste proporcionado por el mismo programa.

Las partes del modelo que requieren que el investigador utilice su conocimiento experto son (1) la selección de oraciones, donde se debe hacer la separación entre final y no final y buscar que las oraciones terminen con palabras con una acentuación similar, ya sea aguda, grave o esdrújula, (2) la selección de los segmentos silábicos de cada oración a utilizar para el análisis (ver tabla 5.2.1) y (3) la agrupación de los contornos entonativos prototípicos de los hablantes con base

en su similitud con otros hablantes del grupo.

Tomando en cuenta lo anterior, se puede concluir que ambos objetivos fueron cumplidos parcialmente, lo que resulta en un modelo de análisis de contornos entonativos para el ámbito forense exitoso siempre y cuando los hablantes utilizados en el estudio cumplan con el principio de variación. Es importante recordar que para realizar un análisis forense de voz de la manera correcta es necesario analizar la voz utilizando los rasgos más significativos para el hablante, es decir, los aspectos de su habla en los que presenta menor variación intrahablante y mayor variación interhablante.

8. Conclusiones

En esta tesis se puso a prueba la hipótesis de que el contorno entonativo que presenta un individuo puede servir como identificador de hablantes en el ámbito forense porque presenta variaciones específicas. La hipótesis anterior fue comprobada parcialmente, como se explica a continuación.

Dentro de los grupos de oraciones finales se encontraron siete hablantes que cumplen con el principio de variación y pueden ser identificados utilizando este método, esto es el 46.7% de la población utilizada en el estudio. La misma situación se presentó en las oraciones no finales, donde siete de los quince hablantes presentaron menor variación intrahablante que interhablante.

Un dato interesante es que cuatro hablantes de los 15 analizados presentan menor variación intrahablante que interhablante tanto en sus oraciones finales como en las no finales.

Respecto a las oraciones finales, el modelo propuesto es capaz de identificar con mayor certeza a aquellos hablantes dentro del grupo de oraciones finales descendentes pero no así a los hablantes clasificados en los otros dos grupos (finales ascendente y finales sin patrón) pues en estos casos presentaron una variación intrahablante e interhablante alta.

Con respecto a las oraciones no finales, se puede observar que el grupo que mejor se adapta a este modelo es el de no finales rectas pues la mayoría de sus hablantes presentan poca variación intrahablante y mucha variación interhablante. El grupo de oraciones no finales curvas también presentó un buen ajuste al modelo de análisis, aunque no obtuvo resultados tan favorables. El grupo de no finales sin patrón no presentó buenos resultados al presentar mucha variación.

Con base en los resultados se pudo observar que el modelo propuesto se ajusta mejor a contornos entonativos curvos, esto debido a que la regresión polinomial de grado cuatro describe principalmente curvas, por lo que es necesario buscar un modelo de regresión que presente mayor ajuste a contornos más rectos.

De forma general, se puede decir que la hipótesis fue parcialmente confirmada dado a que el modelo sí se ajusta con cierto tipo de oraciones y esto nos permite ver que el contorno entonativo de algunos hablantes sí presenta variaciones específicas que lo diferencian de otros hablantes con un contorno entonativo similar. Por lo tanto, este modelo de análisis puede funcionar como una característica más al realizar comparaciones de voz en el ámbito forense cuando los hablantes cumplen con el principio de variación.

Tomando en cuenta los resultados obtenidos, se puede señalar que el patrón de contornos entonativos de oraciones declarativas neutras finales más utilizado en la CDMX es el descendente, mientras que el contorno entonativo para las oraciones no finales es el contorno recto. Cada grupo estaba formado por un 46.7% de los hablantes utilizados en el estudio.

Dado que esta investigación es una propuesta de análisis, las pruebas de validez y fiabilidad del método no formaron parte del objetivo, pero es importante que en investigaciones futuras se revise el desempeño del modelo al comparar hablantes de grupos distintos, por ejemplo, de finales descendentes y ascendentes, con la finalidad de verificar si los resultados cambian y en qué grado.

También resulta de interés probar el modelo con otros grupos de hablantes, en especial con hablantes con un nivel de instrucción diferente y con grupos más numerosos. De ser posible, lo ideal sería probar el modelo bajo condiciones similares a las de una investigación forense.

De igual forma, resulta necesario hacer investigación para generar valores respecto a lo que es poca o mucha variación intrahablante e interhablante, debido a que la literatura no presenta información al respecto, simplemente se atiende a mencionar que es uno de los criterios indispensables para considerar un modelo exitoso, sin detenerse a dar una escala, un rango o alguna metodología para calcularla.

Por último, siguiendo lo presentado en los resultados, es necesario trabajar en una comparación utilizando únicamente contornos entonativos circunflejos, debido al contraste que presentaron durante el análisis.

9. Bibliografía

Akagi, M., & Ienaga, T. (1997). Speaker individuality in fundamental frequency contours and its control. *Journal of the Acoustical Society of Japan (E)*, 18(2), 73–80.
doi:10.1250/ast.18.73

Ariani, M. G., Sajedi, F., & Sajedi, M. (2014). Forensic Linguistics: A Brief Overview of the Key Elements. *Procedia - Social and Behavioral Sciences*, 158, 222–225.
<https://doi.org/10.1016/j.sbspro.2014.12.078>

- Atal, B. S. (1972). Automatic Speaker Recognition Based on Pitch Contours. *The Journal of the Acoustical Society of America*, 52(6B), 1687–1697. <https://doi.org/10.1121/1.1913303>
- Battaner, E., Gil, J., Marrero, V., Carbó, C., Llisterri, J., Machuca, M., Madrigal, M., De la Mota, C., Riera, M., & Wellington, A. (2003). VILE: estudio acústico de la variación inter e intralocutor en español. *Procesamiento del lenguaje natural*, ISSN 1135-5948 (35), 435-436.
- Beckman, M.; Díaz-Campos, M.; McGory, J.; Morgan, T. (2002). Intonation across Spanish, in the Tones and Break Indices framework. *Probus* 14. 9-36
- Bricker, P., Gnanadesikan, R., Mathews, M., Pruzansky, S., Tukey, P., Wachter, K. & Warner, J.(1971). Statistical Techniques for Talker Identification, *Bell Syst. Tech. J.* 50, 1427-1454.
- Caicedo, M. (1991). *Introducción a la sociolingüística*. Cali: Universidad del Valle.
- Cantera, J. (2002). *Teoría y análisis de la entonación*. Barcelona: Publicaciones de la Universidad de Barcelona.
- Congosto Martín, Y. (2015). Entonación idiomática y fonética forense: la información dialectal y sociolingüística de la entonación en español. *Normas*, 0(5), 31–49. <https://doi.org/10.7203/Normas.5.6820>
- Coulthard M, Grant T, and Kredens K (2010) Forensic linguistics. En: Johnstone B, Wodak R, and Kerswill P (eds.) *The SAGE Handbook of Sociolinguistics*. Londres: Sage, 529–544.
- Das, S. & Mohn, W. (1969). Pattern Recognition in Speaker Verification, *AFIPS Conf. Proc., Fall Joint Computer Conf.* 35, 721-732 .
- De la Mota, C., Martín Butragueño, P., & Prieto, P. (2010). Mexican Spanish intonation. En Prieto, P. & Roseano, P. (eds.) *Transcription of Intonation of the Spanish Language*, 319–350, *Lincom Studies in Phonetics* (6), Munich: Lincom Europa.

- Doddington, G. (1971). A New Method of Speaker Verification, *J. Acoust. Soc. Amer.* 49, 139(A).
- Estebas - Vilaplana, E., & Pilar, P. (2008). La notación prosódica del español: una revisión del Sp-ToBI. *Estudios de Fonética Experimental*, 17, 264–283. Retrieved from <http://www.raco.cat/index.php/EFE/article/view/140072/191251>
- (2010). Castilian Spanish Intonation. En Prieto, P. & Roseano, P. (eds.) *Transcription of Intonation of the Spanish Language*, 17- 48, *Lincom Studies in Phonetics* (6), Munich: Lincom Europa.
- Hargreaves, W. & Starkweather, J. A. (1963). Recognition of Speaker Identity, *Language and Speech* 6, 63-67.
- Henríquez Ureña, P. (1938). Mutaciones articulatorias en el habla popular, en *El español en Méjico, los Estados Unidos y la América Central*, Buenos Aires: Universidad de Buenos Aires.
- Hollien, H. (2012). Forensic Phonetics. In *Encyclopedia of Forensic Sciences: Second Edition* (2nd ed., pp. 164–173). Elsevier Ltd. <https://doi.org/10.1016/B978-0-12-382165-2.00029-5>
- Hughes, V., & Foulkes, P. (2015). The relevant population in forensic voice comparison: Effects of varying delimitations of social class and age. *Speech Communication*, 66, 218–230. <https://doi.org/10.1016/j.specom.2014.10.006>
- Künzel, H. J. (1989). How Well Does Average Fundamental Frequency Correlate with Speaker Height and Weight? *Phonetica*, 46(1–3), 117–125. <https://doi.org/10.1159/000261832>
- (1995). The relation between speech tempo, loudness, and fundamental frequency : an important issue in forensic speaker recognition. *Science and Justice*, 35(4), 291–295.

- Kvavik, K. (1974). An analysis of sentence-initial and final intonational data in two Spanish dialects, *Journal of Phonetics*, 2, 351-361.
- Lass, N. (1980). A comparative study of speaker height and weight identification from voiced and whispered speech, *J. Phonet*, 8, 195-204.
- Lass, N., Beverly, A., Nicosia, D. & Simpson, L., (1978). An investigation of speaker height and weight identification by means of direct estimation, *J. Phonet*, 6, 69-76.
- Lass, N., Almerino, C., Jordan, L. & Walsh, J., (1980). The effect of filtered speech on speaker race and sex identifications, *J. Phonet*, 8, 101-112.
- Li, K., Dammann, J. & Chapman, W. (1966). Experimental Studies in Speaker Verification Using an Adaptive System, *J. Acoust. Soc. Amer.* 40, 966-978.
- Luck, J. (1969). Automatic Speaker Verification Using Cepstral Measurements, *J. Acoust. Soc. Amer.* 46, 1026-1032 .
- Martín Butragueño, P. (2008). La estratificación sociolingüística de la entonación circunfleja mexicana. *Realismo en el análisis de corpus orales: primer Coloquio de Cambio y Variación Lingüística*, 93-121.
- Matluck, J. (1965) Entonación hispánica, *Anuario de Letras*, 5, 5-32.
- Morrison, G. S. (2009). Forensic voice comparison and the paradigm shift. *Science and Justice*, 49(4), 298–308. <https://doi.org/10.1016/j.scijus.2009.09.002>
- (2014). Distinguishing between forensic science and forensic pseudoscience: Testing of validity and reliability, and approaches to forensic voice comparison. *Science and Justice*, 54(3), 245–256. <https://doi.org/10.1016/j.scijus.2013.07.004>
- Olsson, J. (2008). *Forensic Linguistics: Second Edition*. Continuum International Publishing Group, 1-16 y 155-187. <https://doi.org/10.1007/s007690000247>

- Orozco, L. (2016). Aproximación a la entonación de enunciados declarativos en Guadalajara. *Estudios de Lingüística Aplicada*, 0(63), 13-35. doi: <https://doi.org/10.22201/enallt.01852647p.2016.63.635>
- Perkins, R., & Grant, T. (2013). Forensic Linguistics. In *Encyclopedia of Forensic Sciences* (2nd ed., pp. 174–177). Elsevier Ltd. <https://doi.org/10.1016/B978-0-12-382165-2.00030-1>
- Pruzansky, S. (1963). Pattern-Matching Procedure for Automatic Talker Recognition, *J. Acoust. Soc. Amer.* 35, 354-358.
- Quilis, A. (1981). *Fonética acústica de la lengua española*. Madrid: Gredos.
- (1985). Entonación dialectal hispánica. *Lingüística Española Actual*, 7, 145-190.
- (1999). *Tratado de Fonología y Fonética Españolas*. Madrid: Gredos. 409-483
- Rose, P. (2002). *Forensic Speaker Identification*. London: Taylor & Francis.
- Saks, M.J., Koehler, J.J. (2005) The coming paradigm shift in forensic identification science, *Science*, 309, 892–895
- Segundo, E. S., Alves, H., & Trinidad, M. F. (2013). CIVIL Corpus : Voice Quality for Speaker Forensic Comparison. *Procedia - Social and Behavioral Sciences*, 95, 587–593. <https://doi.org/10.1016/j.sbspro.2013.10.686>