



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE FILOSOFÍA Y LETRAS

COLEGIO DE LETRAS HISPÁNICAS

Propuesta metodológica para la elaboración de una lista de
términos para conformar un glosario de términos de uso
frecuente en Ciencia Forense

TESIS

QUE, PARA OPTAR POR EL GRADO DE
LICENCIADA EN LENGUA Y LITERATURAS HISPÁNICAS,

PRESENTA

PAOLA GRANADOS ARISTA

DIRECTORA

Dra. Fernanda López Escobedo

Ciudad Universitaria, Cd. Mx.

2 0 1 8



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.



Esta tesis fue elaborada con apoyo del proyecto PAPIIT IA303216 Glosario de
Términos de Uso Frecuente en Ciencia Forense.

*Para Eloisa y Genaro
quienes en el camino se quedaron*

AGRADECIMIENTOS

Este trabajo ha sido posible gracias al apoyo de muchas personas, Espero que, aunque no estén presentes todos en estas líneas, sepan que siempre se encuentran en mi memoria.

A mis padres.

Mamá, simplemente no tendría la oportunidad de escribirte esto si no hubieras estado a mi lado todo este tiempo. Gracias por todos los días soleados, nublados y lluviosos que hemos pasado juntas. Te debo todo lo que soy y en lo que me he convertido, porque me has visto crecer y extender mis alas, y me has apoyado cuando me he caído y roto en mil pedazos quedándote conmigo hasta que nuevamente pueda emprender mi vuelo. Papá, gracias por enseñarme la importancia del esfuerzo y la dedicación y por enseñarme que uno puede aprender de sus errores si tiene algo por qué pelear, tu amor incondicional y tu protección es lo que me ayuda a ser fuerte día con día. Abuela Jose, gracias por el amor que me sigues dando, por demostrarme que el camino es fácil si lo ves desde otra perspectiva, por demostrarme que se puede ser fuerte ante las adversidades, pero que también se vale detenerse un poco y llorar.

A mis hermanos

Pipa, eres el espejo en el que quiero verme reflejada siempre, gracias por el amor incondicional de hermana mayor y el apoyo como amiga, nunca estarás tan orgullosa de ti como yo siempre lo estaré de ti. Gracias por tu interés en mi vida personal y profesional, por ser honesta y directa, pero, sobre todo, por ser mi compañera de vida. Mari, gracias a tu amor por los libros yo estoy aquí, eres una de las razones por las que decidí dedicarme a lo que soy ahora y eso me hace afortunada porque tomé la decisión correcta, gracias por demostrarme el mundo de la literatura de la manera más hermosa posible, a través de la docencia. Avi, gracias por protegerme y cuidarme siempre, por ser el hermano mayor que no esperaba,

pero que necesitaba, porque me has demostrado que con amor las cosas son más sencillas.

A mis amigas y amigos.

A José, porque he encontrado en ti todo aquello que no puedo ver en nadie más, por ser mi persona y demostrarme que puedo soñar y volar alto a pesar de todo, por estar conmigo en los buenos y malos momentos y por creer en mí, aun cuando ni yo podía hacerlo.

Alondra, gracias por estar a mi lado en los momentos difíciles, por esta amistad que se hace cada día más fuerte y duradera, porque hemos crecido y avanzado juntas contra todo. A Janet, pues, aunque nuestros caminos han sido diferentes, siempre estaremos la una para la otra en todo momento. Jimena, gracias por los momentos duros y suaves, por las charlas de todo y nada, por la deconstrucción mutua de la palabra amiga y aliada, porque eres mi persona medicinal, y porque contigo comencé a ver todo con gafas púrpuras. Susana, porque has sido un apoyo para mí en los días difíciles, porque me has demostrado que la dedicación es muy importante, y que a veces uno debe optar por aquello que lo hace feliz a uno más que a los demás, porque tu sino para mí es perfecto. A las cuatro, por ese camino que nos falta.

A Marisol, Dalia, Vanesa y Michelle, por estar conmigo desde los primeros días de la carrera, las clases, así como gran parte de mi vida, serían muy diferentes de no haberlas tenido a mi lado. A Mariana, Ángeles, Fernando y Uriel, porque son una parte de mi vida que no quiero ni necesito soltar, porque cada uno ha aportado un poquito a lo que soy ahora. A mi novena generación de correctores de estilo, que espero de corazón sigamos teniendo esa buena vibra que nos caracteriza, porque no necesitamos parecernos mucho para llevarnos bien. A mis amigos godinez, que me han enseñado la importancia del trabajo en equipo, del apoyo y el valor de una amistad verdadera. Que el amor no se quede en cuatro paredes.

A Fernanda, porque más que una asesora, te has convertido en una buena amiga, una persona a la que estimo y que agradezco todo el apoyo que me ha dado, tanto de manera académica como personal.

A mis sinodales: Jeanett, Zoraida, Teresita y Octavio, pues sus comentarios, sugerencias y aportaciones han nutrido este trabajo desde cada una de sus

trincheras de investigación. Gracias, también, a Jennifer Hincapié por el apoyo recibido para la formación del corpus y de las listas de términos.

A la Licenciatura en Ciencia Forense, a su personal y al proyecto PAPIIT IA 303216, por todo el apoyo brindado para esta investigación.

A todos, gracias.

ÍNDICE

AGRADECIMIENTOS	4
ÍNDICE DE ILUSTRACIONES Y TABLAS	9
CUESTIONES PRELIMINARES Y DESCRIPCIÓN DEL TRABAJO	11
INTRODUCCIÓN	13
1. TERMINOLOGÍA	18
1.1 CONCEPTO DE TERMINOLOGÍA	18
1.2 TERMINOLOGÍA Y LINGÜÍSTICA (LEXICOGRAFÍA)	21
1.3 CONCEPTO DE TÉRMINO	26
1.4 PASADO Y PRESENTE DE LA TERMINOLOGÍA	29
1.4.1 ANTECEDENTES	29
1.4.2 TEORÍA GENERAL DE LA TERMINOLOGÍA (TGT)	31
1.4.3 TEORÍA DE LAS PUERTAS	33
1.5 EL TRABAJO DE TERMINÓLOGO	35
2. MÉXICO Y SU NUEVO SISTEMA DE JUSTICIA PENAL	38
2.1 ANTECEDENTES EN MÉXICO	38
2.2 DERECHOS HUMANOS EN MÉXICO	41
2.3 REFORMA CONSTITUCIONAL EN MATERIA DE SEGURIDAD Y JUSTICIA	43
2.4 SISTEMA DE JUSTICIA PENAL ACUSATORIO	45
2.5 TRILOGÍA DE LA INVESTIGACIÓN	49
3. METODOLOGÍA	53
3.1 CORPUS: SELECCIÓN Y CARACTERÍSTICAS	54
3.2 HERRAMIENTAS DE LINGÜÍSTICA COMPUTACIONAL UTILIZADAS PARA EL PROYECTO	56
3.2.1 GESTOR DE CORPUS GECO	57
3.2.2 ETIQUETADOR DE TEXTOS FREELING	58
3.3 PROCEDIMIENTOS PARA LA EXTRACCIÓN TERMINOLÓGICA	61
3.3.1 MÉTODOS DE EXTRACCIÓN	61
3.3.2 EXTRACTORES TERMINOLÓGICOS	62
3.3.3 EXTRACTOR DE TÉRMINOS TERMEXT	63
4. PROPUESTA METODOLÓGICA	67
4.1 SELECCIÓN DE CORPUS Y PRIMEROS PROCEDIMIENTOS	67
4.2 PRIMEROS RESULTADOS	74
4.3 LIMPIEZA DE CANDIDATOS A TÉRMINO	74
4.4 EVALUACIÓN	77
BIBLIOGRAFÍA	89

ÍNDICE DE ILUSTRACIONES Y TABLAS

Ilustraciones

ILUSTRACIÓN 1 VOCABULARIO GENERAL (BALDINGER, 1952)	25
ILUSTRACIÓN 2 LENGUAJE NACIONAL (HOFFMANN, 1995)	26
ILUSTRACIÓN 3 FASES INTERNAS DEL SISTEMA DE JUSTICIA PENAL	46
ILUSTRACIÓN 4 VISTA PRINCIPAL DE GÉCO, GESTOR DE CORPUS ELABORADO POR EL GRUPO DE INGENIERÍA LINGÜÍSTICA DE LA UNAM.	72
ILUSTRACIÓN 5 VISTA DEL EXTRACTOR TEMEXT TRABAJANDO (2017).	73
ILUSTRACIÓN 7 ESQUEMA DE ENTREVISTA PARA PERITOS. ELABORACIÓN DE LÓPEZ-ESCOBEDO (EN PRENSA).	80

Tablas

TABLA 1 CONCEPTOS Y FUNCIONES DE LA TERMINOLOGÍA CON DIFERENTES ÁREAS.....	20
TABLA 2 DIFERENCIAS ENTRE LINGÜÍSTICA Y TERMINOLOGÍA.	22
TABLA 3 DIFERENCIAS DE LOS SISTEMAS DE JUSTICIA PENAL MEXICANOS.	41
TABLA 4 CLASIFICACIÓN DE ADJETIVOS A PARTIR DE EAGLES, (UNIVERSIDAD POLITÉCNICA DE CATALUÑA, 2017).	59
TABLA 5 ETIQUETAS DE ADJETIVOS A PARTIR DE EAGLES, ELABORACIÓN DE LA UNIVERSIDAD POLITÉCNICA DE CATALUÑA (2017).	59
TABLA 6 TEXTOS NORMATIVOS QUE COMPRENDEN EL CORPUS MODIFICADO. BASADA EN LÓPEZ-ESCOBEDO; ET. AL. (EN PRENSA).	71
TABLA 7 CARACTERÍSTICAS PARA LA LIMPIEZA DE CANDIDATOS A TÉRMINO	76
TABLA 8 ESTRUCTURA TERMINOLÓGICA. ELABORACIÓN PROPIA (2017).....	77

CUESTIONES PRELIMINARES Y DESCRIPCIÓN DEL TRABAJO

Este trabajo forma parte del proyecto PAPIIT IA 303216, uno de los que se desarrollan por parte de la Licenciatura en Ciencia Forense de la Universidad Nacional Autónoma de México. El objetivo principal de este proyecto fue elaborar un glosario que contenga términos especializados para el área de las ciencias forenses y una aplicación en el sistema de justicia penal acusatorio en México, razón que explicaré a lo largo de esta investigación. Para llegar a ello es necesario establecer una metodología que permita seleccionar aquellos términos que deben incluirse en el glosario, a través de una propuesta que incluya herramientas computacionales y extracciones manuales que permitan facilitar y objetivar el trabajo.

INTRODUCCIÓN

La lengua ha sido receptora de grandes cambios a lo largo del tiempo, y el estudio de la variación lingüística nos ha presentado, con los años, las diferentes maneras de expresarnos en distintas lenguas, pero también cómo los diferentes hablantes de una misma lengua pueden hablar de un único objeto en formas diferentes. Estos cambios están definidos por varios factores externos; uno de ellos es el que hace el hablante en los diferentes contextos sociales (Halliday, 1990).

La variedad de lenguas se presenta como una barrera para la adquisición de conocimientos y también dentro de éstas pues los lenguajes especializados dificultan la comunicación entre los hablantes. Esto supone que recurren a elementos lingüísticos distintos para expresar contenidos distintos y, a la vez, que se pueden usar elementos lingüísticos diferentes para decir unas mismas cosas. (Moreno, 1999: 4) Estos lenguajes especializados son tan varios que un solo hablante no tendría la capacidad de memorizarlos, por lo que no es posible lograr que el lenguaje común pueda integrar dentro de sí un lenguaje de ciencia, un lenguaje de ley, un lenguaje de programación y demás lenguajes de especialidad. Este tipo de lenguajes y su variabilidad genera problemas a partir de que se requiera una comunicación entre hablantes de distintas áreas, pues más allá de la falta de información son la ambigüedad y exactitud del éste lo que produce la mala comunicación entre las partes, porque no es sólo que un hablante no entienda lo

que otro dice, sino que dentro de su lenguaje especializado el significante puede corresponder a algo que es completamente diferente a lo que el primer hablante hacía referencia.

Presento en este trabajo un ejemplo claro de los problemas que pueden generar los lenguajes especializados, específicamente en el contexto de la reforma en materia de seguridad y justicia en México. Esta reforma devino en la formación de un nuevo sistema de justicia penal en el país, que busca brindar soluciones rápidas y eficaces a los problemas de impartición de justicia a lo largo de la República Mexicana.

Como característica principal se observa la mudanza de un sistema escrito a uno oral, en el que las distintas partes que participan en el proceso de impartición de justicia establecerán diálogos unos a otros para dar a conocer sus pruebas. Además de que trabajarán en conjunto con diferentes áreas técnicas, científicas y jurídicas, por lo que será necesaria y obligatoria la buena comunicación entre estos actores, de los cuales dependerá la adecuada impartición de justicia. Por esta razón, se busca la elaboración de un glosario que facilite a los partícipes de este sistema romper la barrera comunicativa semántica que presentan los lenguajes especializados. Este trabajo consiste en facilitar a los hablantes las herramientas para que estos tipos de lenguajes sean entendibles entre los distintos especialistas.

Además, se infiere que la implementación del nuevo sistema de justicia penal conlleva el uso de términos propios del ámbito jurídico y de otras especialidades, que todos los actores deben conocer, entender y saber comunicar. La ambigüedad, la polisemia y la escasez de información para definir un significado que actualmente presentan muchos de esos términos complica, y en algunos casos,

impide el correcto flujo de información dentro del sistema de comunicación. Por lo que el consenso entre los actores sobre el significado de un término permitirá el flujo de comunicación y su adecuado entendimiento.

Por ello, el objetivo general del presente trabajo es establecer las bases metodológicas para la elaboración de un glosario de términos de uso compartido entre las diversas áreas de la Ciencia Forense.

A lo largo de este trabajo presentaré el planteamiento, la metodología y elaboración que culminó en una primera lista de términos para conformar un glosario de uso compartido en las distintas áreas de la ciencia forense en México.

En el capítulo primero parto del estado de la cuestión de este trabajo, presento la terminología como la base y la Teoría Comunicativa de la Terminología de María Teresa Cabré (2002) como nuestra aliada en este proceso, pues es nuestra base para justificar la participación de las distintas áreas de acceso para conocer y manejar un término. En la segunda parte se presenta el contexto histórico del trabajo. Presento los antecedentes del sistema de justicia penal de nuestro país, al mismo tiempo que presento el nuevo sistema de justicia penal en nuestro país, conocido coloquialmente como los juicios orales. Asimismo, mencionaré sus características principales, las ventajas y sus desventajas; también mencionaré a los partícipes de las diferentes áreas, los cuales se encargarán de brindarnos un mejor sistema de seguridad y, por ello, presento la necesidad de la comunicación entre pares dentro de este proceso.

Posteriormente, en la tercera parte del trabajo, presento la metodología con la que se trabajó la investigación, presento cuáles fueron los criterios para la selección de un corpus ideal y qué herramientas fueron utilizadas.

Finalmente, explicaré todo el primer desarrollo de la metodología, puesta en acción con nuestro corpus final y el tratamiento técnico y manual de la información. Además, doy a conocer los resultados obtenidos y los criterios que se utilizaron para seleccionar los términos; también presento el trabajo posterior y cómo es que estos resultados formaron la base para la construcción de un glosario.

CAPÍTULO I

1. TERMINOLOGÍA

1.1 Concepto de Terminología

Establecer un concepto concreto para terminología es arriesgado, considerando que es una disciplina que, en los últimos años, ha avanzado debido al progreso en las distintas áreas del conocimiento, principalmente en áreas especializadas, por lo que su definición ha ido evolucionando para cubrir todas las necesidades que ha requerido. Partamos entonces, del concepto básico: "conjunto de términos o vocablos propios de determinada profesión, ciencia o materia." (*Diccionario de la lengua española*, 2014 sv. terminología). Si bien es un concepto acertado, no nos brinda mayor información. Cardero además de la definición indica el uso y funciones la terminología:

permite identificar el vocabulario de una especialidad en forma sistemática en una situación comunicativa específica en los textos propios de la especialidad y entre los profesionales del área, analizarlo desde la lingüística y, si es necesario, crearlo entre el especialista y el terminólogo, además de normalizarlo para funcionamiento concreto con la finalidad de responder a las necesidades de expresión de sus usuarios (2004: 28).

Además de que esta disciplina estudia los subconjuntos de palabras conocidas por un grupo especializado, también se encarga, con base en lo anterior, de la organización de los términos y de su agrupación en glosarios, siendo la terminografía la encargada principal de esta tarea y, con ello, su metodología y teorías de clasificación. Para muchos investigadores, la terminología no sólo está

relacionada con la lingüística, es por ello que llegar a su denominación actualmente resulta complicado si se viera desde una sola área, por ello se considera a la ciencia terminológica como transdisciplinaria, ya que “guarda relaciones especialmente estrechas con la lingüística, el conjunto de las ciencias y sus aplicaciones, la filosofía, la información y documentación, la lingüística computacional y la ingeniería del conocimiento” (Arntz y Picht, 1995: 22). Por lo que, enfocarse sólo en los conceptos de conjuntos especializados y normalización provocaría incongruencias en la presente investigación.

Para Cabré (1995), la terminología tampoco es sólo sinónimo de estandarización o normalización, sino representativa de la diversidad. Depende de los diferentes conceptos en que se aloje la disciplina, aquello que la forma y en las aplicaciones dadas en las áreas de investigación. Con base en lo que Cabré dice, depende de la disciplina en cuestión para obtener un concepto conveniente de terminología, no obstante, también se contará con puntos en común, así, usando como ejemplo tres áreas de conocimiento (lingüística, filosofía y disciplinas científico-técnicas) se trata de establecer puntos en común y de quiebre del concepto que pueda abarcar todos los campos en cuestión. Es importante establecer, entonces, que la terminología no sólo depende de un área especializada sino de un conjunto de disciplinas que van complementando su definición y aplicación. A manera de comprender la diversidad que tiene la terminología en áreas mayores de estudio (las anteriormente mencionadas), elaboro un esquema comparativo con base en la información que Cabré proporciona:

Terminología	Lingüística	Filosofía	Disciplinas científicotécnicas
Es	Conjunto de signos lingüísticos que constituyen un subconjunto dentro del componente léxico de la gramática del hablante.	Conjunto de unidades cognitivas que representan conocimiento especializado.	Conjunto de unidades de expresión y comunicación que permiten transferir el pensamiento especializado.
Función prioritaria	Significación	Cognición y representación	Denominación para la transferencia
La terminología es una manera de:	Saber	Conocer	Comunicar
Relación poliédrica (unidades lingüísticas, cognitivas y sociocomunicativas)	Nombre-significado	Cosa-significado	Cosa-nombre
Concepción global de 'término'	Signo lingüístico	Concepto o representación de la realidad	Unidad de expresión y denominación

Tabla 1 Conceptos y funciones de la terminología con diferentes áreas.

Los tres apartados interpretan a la terminología como un conjunto de unidades, que poseen una finalidad esencial y con un valor funcional (Cabré, 1995), aunque éstas no coincidan unas con otras. Así pues, el concepto más acertado es aquel que envuelva estas tres áreas.

La terminología para las tres disciplinas es el conjunto de términos concebidos como unidades sánicas de tres caras (cosa, nombre y significado) que hacen referencia a la realidad especializada. Las diferencias que podemos percibir se limitan simplemente a posiciones o priorizaciones distintas: distintas en cuanto al marco situacional (la gramática, la ciencia cognitiva o el dominio especializado), distintas en cuanto a la priorización de la perspectiva con la que se enfoca el término (signo lingüístico, unidad

conceptual o signo sin más), distintas en cuanto a la función prioritaria que atribuyen al término (la significación, la representación o la denominación) (Cabré, 1995).

El concepto de Cabré además posiciona a la terminología como una disciplina independiente de todas las demás, que ha adquirido elementos específicos de otras áreas (lingüística, lógica, tecnologías, etc.) para formar su propia teoría, y que además está ligada necesariamente a la comunicación y a las tecnologías emergentes; ya que "en cualquier área especializada, las terminologías constituyen la base para la comunicación especializada escrita y oral" (Arntz y Picht, 1995: 23).

1.2 Terminología y lingüística (lexicografía)

Si bien, es posible concebir a la terminología dentro del campo de la lexicografía, ya que el objeto de estudio es diferente para ambas: término y palabra. No obstante, la línea que divide a una de otra es difusa, pues "incluso los propios diccionarios de la lengua general recogen acepciones de palabras que se podrían clasificar como términos, dado que los conceptos que designan pertenecen a dominios de especialidad." (García, 2001). Esta falta de claridad se debe principalmente a que no se consideran otros factores más que la misma unidad léxica (García, 2001).

Estudiosos como Cabré incluso consideran que la terminología no pertenece en un cien por ciento al campo de la lingüística, sino que se trata de una interdisciplina, constituida por elementos lingüísticos (Cabré, 1995). A consideración de Cabré, el tener bases teóricas delimitadas, objeto de estudio, área teórica y aplicada, y al no poseer todos los fundamentos lingüísticos, la convierten

en una disciplina de carácter autónomo. A esto último, tomo en consideración los puntos más importantes que Cabré (1995) recalca para establecer su postura.

Lingüística	Terminología
Tiene como objeto de estudio el lenguaje, para dar a conocer su función parte de una muestra idealizada que viene de un hablante también idealizado.	Parte de un lenguaje real (adquiere datos documentados –orales y escritos-) para explicar la función del lenguaje especializado.
El significado debe de estar relacionado a la palabra.	El concepto es independiente de aquello que lo designa
Las palabras son descritas teniendo en cuenta su contexto.	El contexto no se considera.
Distingue y engloba aspectos sincrónicos y diacrónicos de las palabras.	Sólo se interesa por la sincronía de sus términos.
Descriptivista	Prescriptivista
Lexicología: estudio de las palabras.	Estudio de los términos propios dentro de un capo especializado.
Estudio de las palabras. Análisis y descripción de la competencia léxica del hablante.	Estudio de los términos (palabras propias de un campo de especialidad).

Tabla 2 Diferencias entre lingüística y terminología.

En efecto, la terminología posee más aspectos de la lingüística que de otras disciplinas, por lo que su separación de ésta -y especial de la lexicología- resultaría complicada, ya que, si bien lexicología y terminología no tienen criterios en común, sí responden a una misma necesidad: sistematiza una estructuración léxica basada en convencionalismos lingüísticos (García, 2001). Estas discusiones sobre la posición de la terminología dentro del área lexicográfica, y ahora lingüística, se dan a partir de que tampoco existe una línea tangible para separar el lenguaje común del especializado, por lo que hablaremos de esto a continuación.

El lenguaje especializado aparece debido a la necesidad de comunicar nuevos conocimientos que se van construyendo y se van actualizando de acuerdo con lo que la sociedad especializada requiere. Manuel José Morales (2004) le atribuye su reciente importancia a la globalización, ya que existe una mayor necesidad de investigación científica y, por ende, de un lenguaje que funcione como herramienta no sólo de etiquetaje, sino como herramienta de divulgación.

De acuerdo a las características que posee este lenguaje, éste entraría en la diferencia diatécnica, para Ettinger

la diferenciación técnica es aquella que se establece a partir de los lenguajes de grupo llamados 'tecnocetos' o 'lenguajes especializados', cuya principal característica, radica en el uso de vocablos propios y exclusivos de una profesión, un oficio o una especialidad, esto es: su terminología (Ettinger en Cordero, 2009: 76)

Hablar de lenguajes especializados es, indirectamente, hablar de lenguaje común, ya que la mayoría de sus definiciones, como se ve, deben hacer un contraste entre uno y otro. Por lo que es necesaria una definición clara de lo que es lenguaje común

El lenguaje común es una forma del idioma que tiene validez en todo el territorio donde se habla ese idioma y que es entendible por todos los que pertenecen a ese grupo lingüístico; se usa para el intercambio de pensamientos de índole general, es decir, sin una orientación específica dada por algún campo del saber (Fedor de Diego en Cordero, 2009: 76)

La norma DIN¹ 2342 define al lenguaje especializado como:

el área de la lengua que aspira a una comunicación unívoca y libre de contradicciones en un área especializada determinada y cuyo funcionamiento encuentra un soporte decisivo en la terminología

¹ La norma DIN del Instituto Alemán de Normalización, por sus siglas en Alemán (Deutsches Institut für Normung) se encarga de la normalización de terminología especializada, realizando las mismas funciones que las normas ISO (Organización de Estandarización Internacional, por sus siglas en inglés).

establecida. [...] El núcleo de la lengua del que participan todos los miembros de una comunidad lingüística se llama lengua común (Arntz y Picht, 1995: 28)

Esta definición, así como las anteriores, nos ayuda a entender también que, primero, no debemos hablar del lenguaje especializado como una unidad, sino todo lo contrario, que existe más de un lenguaje especializado y, segundo, que se hace una separación del lenguaje especializado y de lo que se conoce como lenguaje común. Por ello, la cuestión ahora es dónde se trazan las líneas entre lenguaje especializado y lenguaje común (Arntz y Picht, 1995: 28).

Se debe entender al lenguaje especializado como un subsistema, el cual depende de un lenguaje común. Para Rodríguez Díez (2009) existe también una clasificación dentro del lenguaje especializado de acuerdo con determinados factores: la realidad y el pensamiento. Conforme al grado de dominio entre uno y otro, se llegan a distinguir tres clasificaciones (Cardero 2009: 77)

- Lenguajes científico técnicos: existe una correspondencia biunívoca entre significante y significado (Nomenclaturas, vocabularios específicos)
- Argots: Se manejan lenguajes cuya función expresiva es mayor que en las otras clasificaciones, además es un lenguaje cerrado intencionalmente. Lo que comúnmente conocemos como jergas.
- Lenguajes sectoriales: Éstos tienen un grado menor de especialización, pero no llegan a pertenecer al lenguaje común. Podría denominarse como un lenguaje semiespecializado.

Arntz y Picht también nos muestran algunos esquemas para poder comprender a qué se refieren con lenguaje especializado². A continuación, presento y explico dos de ellos que nos ayudarán a exponer mejor los conceptos proporcionados.



Ilustración 1 Vocabulario general (Baldinger, 1952)

El primer modelo que presento es de 1952, elaborado por Baldinger³, en éste se puede observar la relación que tiene el lenguaje común (llamado general) con el especializado y, además, un lenguaje intermedio entre ambos (como el sectorial antes mencionado). Baldinger especifica que el esquema se encuentra en círculo debido al constante proceso de compensación en ambas direcciones (Baldinger en Arntz y Picht, 1995:30), es decir, los tres lenguajes (general, medio y especializado)

² Cabe mencionar que estos esquemas los he elegido para mi marco teórico pues presentan a grandes rasgos cómo se estudia el lenguaje especializado partiendo desde el lenguaje general, sin llegar a separarlos de nuestro tema central; además con estos primeros estudios y esquemas, recopilados por Arntz y Picht, podemos entender el devenir de la terminología como la conocemos ahora.

³ "Die Gestaltung des wissenschaftlichen Wörterbuchs"/"Organización de un diccionario científico". Traducción castellana de 1984.

son directamente proporcionales, y se pueden encontrar en un mismo hablante, es por ello también que hace una división del lenguaje medio y de especialidad con base en disciplinas.



Ilustración 2 Lenguaje Nacional (Hoffmann, 1995)

Para complementar esta idea, y sin llevar el tema hacia otra dirección, explico e cuadro elaborado por Hoffmann en 1985. En dicho esquema el lenguaje común y especializado se encuentra en el mismo plano, no obstante, desarrolla un concepto nuevo: lengua global, la cual se considera como “una reserva de todos los medios lingüísticos”. (Hoffmann en Arntz y Picht, 1995:35).

1.3 Concepto de término

Una vez expuestas presentadas las diferentes posturas acerca de la terminología, se debe entender que el concepto de término, cabe mencionar que el término es esencial en el contexto especializado, pues nos brinda una definición más concreta y adecuada a nuestra necesidad en determinada disciplina, motivo por el cual se

elabora el presente trabajo. Dicho esto, doy a conocer acepciones que el Diccionario de la Lengua Española (2014) nos ofrece sobre la palabra:

1. m. Último punto hasta donde llega o se extiende algo.
2. m. Último momento de la duración o existencia de algo.
3. m. Límite o extremo de algo inmaterial.
4. m. mojón (|| señal para fijar los linderos).
5. m. Línea divisoria de los Estados, provincias, distritos, etc.
6. m. término municipal.
7. m. Plazo de tiempo determinado.
8. m. palabra (|| unidad lingüística).
9. m. Estado o situación en que se halla alguien o algo.
10. m. Forma o modo de portarse o hablar. U. m. en pl.
11. m. talle (|| traza, disposición o apariencia).
12. m. Elemento con el que se establece una relación. Término comparativo, de referencia.
13. m. Plano en que se representa alguien o algo ante la vista. Primer término de un cuadro, de una escena.
14. m. En una enumeración con los adjetivos primer, segundo y último, lugar que se atribuye a lo que se expresa.
15. m. Arq. Sostén o apoyo que termina por la parte superior en una cabeza humana, al modo que los antiguos figuraban al dios Término.
16. m. Fil. Aquello dentro de lo cual se contiene enteramente algo, de modo que nada de ello se halle fuera.
17. m. Fil. Cada una de las palabras que sustancialmente integran una proposición o un silogismo. Los términos de una proposición son dos: sujeto y predicado; los de un silogismo son tres: mayor, menor y medio.
18. m. Gram. Cada uno de los dos elementos necesarios en la relación gramatical.
19. m. Gram. Función sintáctica desempeñada por el sintagma introducido por la preposición.
20. m. Gram. Segmento sintáctico que desempeña la función de término.

21. m. Mat. Numerador o denominador de un quebrado.
22. m. Mat. En una expresión analítica, cada una de las partes ligadas entre sí por el signo de sumar o de restar.
23. m. Mús. punto (|| tono).
24. m. p. us. Hora, día o punto preciso de hacer algo.
25. m. p. us. Objeto, fin.
26. m. desus. Paraje señalado para algún fin.
27. m. pl. Condiciones con que se plantea un asunto o cuestión, o que se establecen en un contrato.

Como puede observarse, la definición que más se acerca a lo que se ha expuesto los especializados corresponde a la octava acepción (de 27) dentro del diccionario; sin embargo, no se puede definir al término como un sinónimo de palabra ya que poseen características si no contrastantes, sí únicas de cada concepto. A continuación, presentaré tres definiciones más:

Los términos son unidades léxicas que se utilizan para designar conceptos en un ámbito temáticamente restringido. En general, los términos no pueden distinguirse del lenguaje general si nos limitamos simplemente a mirar su forma (Vivaldi, 2001: 2)

Un término como elemento de una terminología, es una unidad constituida por un concepto y su denominación. (Norma DIN 2342 en Arntz y Picht, 1995:37)

Un término es una unidad de características lingüísticas similares, utilizada en un dominio de especialidad. Donde este punto de vista, una palabra que forme parte de un ámbito especializado sería un término. (Cabré, 1995)

Utilizo únicamente estas enunciaciones porque son suficientes para poder proporcionar una definición propia. Dada la discusión anterior, considerar al término como una unidad de carácter léxico resulta confuso. La terminología no

posee en su totalidad una representación en la lexicología o en la lingüística. Si bien estas acepciones tienen rasgos en común, principalmente al referirse al término como una unidad considero que la definición de Cabré es la más acertada para este trabajo, ya que no clasifica al término en alguna disciplina, además de que lo aproxima al campo especializado.

Dadas las acepciones y para fines de la presente investigación aportaré mi propio concepto sin ahondar en características o funciones. “El término constituye una unidad con fundamentos lingüísticos, que es representativa del campo de la especialización”.

1.4 Pasado y presente de la terminología

Terminología y lenguaje especializado han ido de la mano desde existen los avances tecnológicos. A continuación, presento sus antecedentes, así como sus ideologías principales, desde su surgimiento hasta nuestros días.

1.4.1 Antecedentes

Como ya hemos mencionado anteriormente, son los avances tecnológicos los que detonan la aparición de la terminología, por lo que, a partir del siglo XVIII, con la Revolución Industrial, hay una necesidad de la estandarización de conceptos para evitar ambigüedades (Barrón, 2007: 8). Sin embargo, en esa época la terminología tenía un matiz negativo –palabras que nadie entendía– (García, 2001). Durante todo este periodo no se consideró una disciplina sino hasta el siglo

pasado, favorecido por la globalización y la interacción entre lenguas (Barrón, 2007: 8). La Escuela Rusa, durante la tercera década del siglo xx, fue la primera en adentrarse en este nuevo campo, ya que el país tenía un enorme interés por el desarrollo tecnológico. (Cabré, 2001: 9) Lotte y Drezen (1926) como los principales exponentes, presentaron los resultados de sus investigaciones en su país, en ellas se afirmaba que la falta de estandarización de términos era una de las causas principales de la lentitud del desarrollo tecnológico, es aquí donde se habla por primera vez de una teoría y una metodología. Los trabajos de Lotte fueron el comienzo para que otros investigadores se le unieran, formando así el Comité para la Terminología Científica y Técnica (ΚΝΤΤ). Este comité trabajó sobre distintos ámbitos de especialización, además "se redactó una guía metodológica para la preparación y regulación de la terminología técnica y científica" (Cabré, 2001: 10). Fue Eugen Wüster quien vio el auge de la terminología a través de sus trabajos con el objeto de incluirla en los diferentes ámbitos sociales y académicos. Uno de sus grandes aportes fue establecer la terminología como una necesidad comunicativa de la normalización, por lo que surgió la Organización de Estandarización Internacional, ISO. Otro de los logros fue la aprobación de un programa por la UNESCO para la terminología, que dio lugar en Viena al centro internacional de documentación sobre terminología, INFOTERM (Cabré, 2005). Además, los trabajos y anotaciones de Eugen Wüster sirvieron para sentar la base de la que hoy en día conocemos como Teoría General de la Terminología.

1.4.2 Teoría General de la Terminología (TGT)

Esta teoría de los años treinta, al igual que la rusa, surge en Viena por Wüster y su necesidad de superar los obstáculos de la comunicación profesional, los cuales son causados por la imprecisión, diversificación y polisemia del lenguaje común. Dicha teoría estaba enfocada hacia la normalización de las unidades de comunicación con base en el estudio de los términos y del análisis de sus relaciones.

A continuación, enlisto algunas características de la teoría proporcionadas por Teresa Cabré (2005):

- La TGT se construyó sobre una base lingüística
- Hace una separación entre concepto y designación, así como denota prioridad del primero sobre el segundo.
- No retoma aspectos diacrónicos de conceptos y términos
- Resalta la importancia de la forma escrita de los términos sobre la oral, ya que la primera la considera como la única forma de normalización.
- Método de trabajo onomasiológico⁴
- Sugiere establecer normas internacionales para la estandarización de conceptos.

Sin embargo, la TGT presenta insuficiencias, ya que desde un principio es imposible lograr una comunicación sin ambigüedades y la búsqueda de un lenguaje universal, porque se tendría que pasar por alto todas las restricciones del lenguaje, además la prioridad científica opaca las características socioculturales,

⁴ Que va del concepto a su denominación (significado a significante). Definición propia y DEL (2014).

socioeconómicas, ubicaciones geográficas, y clasificación de las lenguas; y también, para la TGT, los términos no presentan valor pragmático ni variación semántica; es por ello que la normalización no daría lugar a la diversidad denominativa y conceptual de la realidad (Cabré, 2001).

Los cambios sociales acelerados ocurridos en la segunda mitad del siglo XX como la multiplicación de los temas de intercambio internacional, que son causados por un crecimiento acelerado del conocimiento, así como la diversificación de contenidos especializados y la interacción plurilingüe indujeron a un nuevo replanteamiento de los valores considerados como verdaderos durante la primera mitad del siglo.

Son estas razones las que provocan que la Teoría General de la Terminología se considere, en este punto, insuficiente, ya que está fundada en una serie de supuestos e ideales poco fiables. Cabré (2002) define a la terminología moderna como

el campo de encuentro entre la lingüística, la ciencia cognitiva, la ciencia de la información, la comunicación y la informática, establece un objeto de análisis y unas funciones de trabajo muy restrictivos, por cuanto limita a su objeto a las unidades unívocas normalizadas propias de los ámbitos científico-técnicos, reduce la actividad terminológica a la recopilación de conceptos y de términos para la normalización (fijación de nociones y denominaciones estandarizadas) de los términos circunscribe los ámbitos especializados a la ciencia y a la técnica y limita sus objetivos con el fin de asegurar la univocidad en la comunicación profesional, fundamentalmente en el plano internacional.

Esto quiere decir que el propósito de esta disciplina es tener una precisión en la comunicación especializada por medio de términos regulados.

La terminología moderna tiene como objeto de estudio las unidades terminológicas. Estos términos deben poseer un carácter lingüístico, si no lo

poseen no se considerarán términos, sino unidades con conocimiento especializado. El carácter específico de los términos va a depender de aspectos pragmáticos y del modo de significación; cabe mencionar que el resultado dependerá de la discusión experta del término dentro del discurso especializado a través de predicaciones.

1.4.3 Teoría de las puertas

Todo lo anterior conlleva a una nueva propuesta de una teoría de la terminología que pudiera satisfacer todas estas carencias que presenta la TGT en los tiempos actuales.

En primera instancia, el campo de conocimiento de la terminología es interdisciplinario, está formada por aspectos comunicativos, semióticos, cognitivos, y lingüísticos; además, el objeto de estudio deben ser las unidades terminológicas y se debe partir de ellas para conocer los resultados esperados en una investigación de términos especializados. El carácter específico de los términos va a depender de aspectos pragmáticos y del modo de significación; cabe mencionar que el resultado dependerá de la discusión experta del término dentro del discurso especializado a través de predicaciones.

De esta manera, la Teoría de las Puertas considera a la terminología como "un campo de conocimiento intrínsecamente interdisciplinar, que se ocupe de términos y que integre los aspectos cognitivos, lingüísticos, semióticos y comunicativos de las unidades terminológicas" (Cabré, 2002). Es una teoría que permite explicar la relación de los términos a través de diferentes disciplinas.

La Teoría de las puertas o Teoría Comunicativa de la Terminología (TCT) se presenta como un esquema donde el propósito es llegar a las unidades terminológicas, esto a partir de las distintas disciplinas (puertas) en las que se encuentra relacionada.

Las bases de esta nueva teoría son (Cabré, 2005)

- Las unidades terminológicas son el objeto central de la Terminología como campo de conocimiento;
- Las unidades terminológicas son poliédricas (lingüísticas, cognitivas y sociocomunicativas);
- Cada puerta de entrada exige una teoría propia, que debe compartir el mismo objeto central (las unidades terminológicas), y su concepción poliédrica tiene que ser coherente con las teorías propias de las demás puertas;
- El análisis de las unidades terminológicas desde la lingüística presupone entrar a través de los textos o producciones lingüísticas orales;
- En estos textos, las unidades terminológicas son las más prototípicas para la representación del conocimiento especializado;
- Son unidades denominativas y designativas que presentan variación (polisemia y sinonimia);
- Las unidades terminológicas comparten con otras unidades lingüísticas la expresión del conocimiento especializado;
- Las unidades terminológicas en una teoría del lenguaje natural no se conciben como unidades separadas de las palabras que comparten espacio

en el léxico del hablante, sino como valores especializados de las unidades léxicas contenidas en el lexicón del hablante;

- Sólo una teoría Lingüística de base cognitiva y funcional, es decir, que contenga semántica y pragmática, además de gramática es capaz de describir las unidades terminológicas en su especificidad, pero también dar cuenta de lo que comparten las unidades terminológicas con las unidades léxicas no especializadas. Y la pragmática, además, es indispensable para explicar la activación del valor terminológico de las unidades léxicas.

1.5 El trabajo de terminólogo

A lo largo del capítulo he hablado del desarrollo de la terminología desde sus inicios hasta la forma en que se concibe actualmente, ahondando en la historia de la disciplina presentando a sus exponentes e, incluso, las controversias que han llegado a presentarse con el paso del tiempo, logrando una evolución del concepto. Ahora bien, no me gustaría terminar este capítulo sin mencionar el papel del terminólogo. En esencia, para Cabré, el terminólogo debe “elaborar diccionarios especializados (denominados diccionarios técnicos), necesarios en un periodo de fuerte expansión y amplio desarrollo de la técnica y la tecnología” (2005). La terminología se considera multidisciplinaria por lo que el terminólogo, además de tener una formación de base lingüística, debe conocer otras áreas, esto se debe también, pues debe buscar claridad y precisión en cuanto a las unidades que deberá interpretar.

además de conocer la materia, debe expresarla, comunicarla a interlocutores precisos. Y estos interlocutores le exigen necesariamente precisión, pero también claridad y eficiencia comunicativa. El especialista se ve pues en la necesidad de adecuar su discurso sobre la materia a las necesidades receptivas de los distintos interlocutores, y para hacerlo debe ser capaz de buscar alternancias denominativas, fórmulas que sin transgredir el contenido comuniquen adecuadamente, metáforas explicativas, convincentes, etc. (Cardero, 2003)

Así, el quehacer del terminólogo consiste en identificar términos prototípicos, no para estandarizar el lenguaje, sino para facilitar la comunicación.

Finalmente, es deber de los mismos especialistas seguir nutriendo el campo de la terminología, ya que con sus investigaciones y procesos se verá reflejado el desarrollo de esta disciplina a nivel nacional e internacional.

CAPÍTULO II

2. MÉXICO Y SU NUEVO SISTEMA DE JUSTICIA PENAL

Esta tesis pretende aportar conocimiento a una situación que en el marco jurídico-social es de suma importancia para nuestro país: la implementación de un nuevo sistema de justicia penal. A lo largo de este capítulo presentaré el contexto jurídico-social en el cual se pretende incidir con este trabajo. Para ello, abordaré las necesidades históricas que fueron precisas para la construcción de esta nueva forma de impartir justicia, así como el impacto que tendrá en nuestra nación.

2.1 Antecedentes en México

Es importante conocer que la manera de impartir justicia en el país dependerá siempre de un sistema o proceso de enjuiciamiento, es decir, "la forma, modo o manera como un estado resuelve los conflictos interpersonales de sus gobernados, y que este sistema refleja con mayor exactitud los contenidos democráticos o autocráticos de su constitución" (Montes, 2009).

Para esto, y de acuerdo con María Alexandra Montes (2009), existe una relación directamente proporcional entre la autocracia y el sistema inquisitivo, así como la democracia y los sistemas acusatorios. A continuación, presento brevemente dos sistemas que han sido parte de la formación del proceso de impartición de justicia en el país.

- Sistema Inquisitivo.

Este sistema parte del principio de que la impartición de justicia recae en el juez, por lo que no existe una imparcialidad. El procedimiento es completamente escrito. Este proceso tiene sus orígenes en el segundo imperio romano a partir del siglo XIII. Llegó al continente americano debido a las guerras de conquista en el siglo XVI, posteriormente estuvo vigente en México desde 1886 hasta 1917 con Porfirio Díaz. (Cruz, 2004: 156). Se rige bajo los principios de: inmediatez, oficiosidad, secrecía, unidad de parte, escritura y sistema de valoración de tazado legal

- Sistema Mixto.

Este sistema parte de la estructura del sistema inquisitivo, se hace la modificación para que los acusados puedan tener una defensa participativa y así poder equilibrar el proceso, por lo que, idealmente, se prepara una acusación y una defensa, es por ello que se conoce como sistema mixto. Podría decirse que las revoluciones democráticas hacen que este sistema, en esencia pueda ser más justo para todas las partes, no obstante, tiene las mismas raíces del sistema inquisitivo. Este sistema llega al país en 1917 con la nueva constitución, presentando órganos institucionales, tanto para la acusación como para la defensa (Cruz, 2004: 287). No es sino con la implementación del Código de Procedimientos Penales para el Distrito Federal en la década de 1930 donde entró en su apogeo. El proceso se dividió en 3 etapas: averiguación previa, pre proceso o término constitucional y juicio. Además, se creó la figura del ministerio público.

Inicialmente se tenía considerado el uso de un jurado, sin embargo, con el paso de los años dejó de ser considerado debido a la manipulación de los abogados hacia el jurado.

Así, al llegar a los debates finales, toda vez que el Jurado decidiría sobre cuestiones de hecho y no de derecho, llegaba la oportunidad de los abogados de apelar haciendo gala de actitudes dramáticas, a la conciencia de sus miembros con argumentos no jurídicos, exponiendo los hechos desde el lado de las pasiones humanas y no de su tipificación en los códigos (Rabasa, 1998: 575)

El cambio de un proceso penal de carácter mixto con bases inquisitorias a un actual acusatorio inclinado hacia la democracia es un paso enorme en nuestro sistema de justicia, recordemos que la razón principal del cambio fue que se estaban violando las garantías individuales de los ciudadanos por lo que era imperativo hacer modificaciones en los procesos a nivel nacional. No obstante, las comparaciones entre uno y otro sistema son muchas, por lo que intentaré contabilizar las más importantes.

Sistema Mixto	Sistema Acusatorio
Procedimiento escrito	Procedimiento oral
Órgano judicial pasivo	Órgano judicial activo
La declaración tiene un valor casi total frente al Ministerio Público.	La declaración será ante el juez y en presencia de ambas partes, además estará presente la defensa técnica.
El promedio de tiempo es de 4 meses a 2 años.	La duración puede ser incluso de una semana.
El desahogo de pruebas se presenta	Las pruebas deben desahogarse en una

en más de una audiencia.	misma audiencia y en presencia del juez.
El Juez puede delegar su cargo a subordinados.	El Juez debe estar presente en las audiencias, además de ser partícipe de la investigación de las pruebas y en la declaración de las partes.

Tabla 3 Diferencias de los sistemas de justicia penal mexicanos.

2.2 Derechos Humanos en México

Como se ha podido apreciar, nuestro sistema penal mixto presentaba una variedad de problemas, e inconsistencias que provocaban un entorpecimiento en la demanda de justicia de los casos, así como abusos por parte de las autoridades del poder judicial. García menciona que “el proceso penal representa el principal campo de tensión entre la seguridad ciudadanas y el derecho a la libertad de quien es llamado a comparecer en juicio” (García, 2014: 13) cosa que no era reflejada en nuestro sistema penal.

No existe una única razón para una modificación en el sistema de impartición de justicia, Duce y Pérez (2001) consideran que hay causas generales en América Latina que pueden aplicarse a nosotros, las cuales son:

1. El proceso de democratización presente en varios países de América Latina, luego de décadas de gobiernos dictatoriales o autoritarios;
2. La crítica a sistemas políticos obsoletos;
3. Las presiones para modernizar al Estado en general;
4. La reevaluación del papel de los derechos humanos;

5. La percepción negativa sobre el sistema judicial y, en particular, sobre los juzgados penales.

Si bien todas estas razones son válidas, fue la presencia de organizaciones no gubernamentales dedicadas a la defensa de los derechos humanos las que más han impulsado el cambio en la forma de impartir justicia en América Latina.

La Organización de la Naciones Unidas, durante el gobierno de Vicente Fox, realiza una serie de recomendaciones para impedir más violaciones de los derechos humanos en el país.

Entre las observaciones hechas a México por tal organismo [Naciones Unidas] destaca la ausencia de un modelo acusatorio en el que imperen los principios de contradicción, oralidad, intermediación, concentración, publicidad y economía procesal; la existencia del sistema tutelar de sanción a los menores infractores, la ineficacia e ineficiencia del sistema de ejecución de sanciones, la falta de profesionalización de los defensores penales, y la ausencia de plena autonomía del Ministerio Público de la federación. (González, 2013: 187)

Estas recomendaciones hicieron evidentes las carencias de nuestro sistema de justicia, carencias que ya podían ser vistas por todos los involucrados en procesos anteriores. Carbonell y Ochoa (2007) mencionan que el sistema penal mexicano se encontraba en bancarrota, pues,

Ninguno de sus principales actores puede estar satisfecho con su funcionamiento actual [...] a) no sirve para atrapar a los delincuentes más peligrosos; b) permite la existencia de un altísimo nivel de impunidad y corrupción; c) no asegura los derechos fundamentales ni de víctimas, ni de acusados; d) no establece incentivos para una investigación profesional del delito; e) es sumamente costoso si se toman en cuenta sus pobres resultados.

Por esta razón había una falta de confianza y empatía tanto al sistema de impartición de justicia como a las instituciones. De acuerdo con cifras presentadas

de 2003 por Miguel Carbonell, el 85% de las víctimas no acuden a denunciar los delitos; 99% de los delincuentes no terminan condenados; 92% de las audiencias en los procesos penales se desarrollan sin la presencia del juez; 80% de los mexicanos cree que se puede sobornar a los jueces; 60% de las órdenes de aprehensión no se cumplen; 40% de los presos no ha recibido una sentencia condenatoria , además el 80% de los detenidos nunca habló con el juez que lo condenó (Bergman, 2003: 47).

El proceso penal es el elemento de cierre de la lucha contra la criminalidad y el instrumento de control social de la política pública de seguridad. Su diseño busca un balance entre la represión efectiva de las conductas delictivas y el debido respeto de las libertades individuales. Se trata de uno de los momentos de equilibrio de la mayor sensibilidad en el Estado democrático (García, 2014: 13)

Se demostró con esto, que nuestro sistema penal era nulo para las necesidades del país, además de que promovía altos niveles de corrupción. Las sugerencias propuestas por la ONU son el inicio de una serie de reformas de carácter constitucional, realizadas años más adelante por el gobierno de la República, con el propósito de cumplir las exigencias de la organización, así como "fortalecer el combate a la delincuencia organizada y brindar mayores derechos y oportunidades de participación a la víctima u ofendido por el delito" (González, 2013: 188)

2.3 Reforma constitucional en materia de seguridad y justicia

Con lo anterior, en 2008 se realizó una reforma constitucional de carácter normativo y estructural al sistema de justicia penal. En ella se modificaron los artículos 16, 17, 18, 20, 21, 22, 73, 115 fracción VII, y 123 apartado B fracción XIII de la Constitución Política de los Estados Unidos Mexicanos. Con esta reforma

se trata de contar con un aparato del más alto nivel normativo para enfrentar desde un punto de vista garantista a la delincuencia, establecer razonables condiciones de seguridad pública, reducir la impunidad, abatir la corrupción y mejorar, en general, el desempeño estatal en lo que respecta a la prevención del delito y la impartición de justicia penal (García 2014: 13)

Además de la reforma constitucional, se presentó, el 5 de marzo del 2014, el Nuevo Código Nacional de Procedimientos Penales. Este documento, junto con la Constitución Política de los Estados Unidos Mexicanos, sirve para abrirle paso al Sistema de Justicia Penal Acusatorio, por lo que se presenta una transición en la manera de impartir justicia en el país.

Se trata de un tránsito desde un esquema de resolución de conflictos mediante la estricta aplicación de la ley y el precedente (en el marco de un sistema inquisitorio y escrito), hacia una interpretación y aplicación del Derecho más dinámica e integradora. Se nutrirá no sólo del análisis normativo formal, sino de los principios y conexiones lógicas que lo hagan aplicable ante la particularidad de cada uno de los casos de resolución. (García, 2014: 14).

Para principios de 2017 todos los estados de la República implementaron de manera oficial este sistema. Los resultados de este sistema son notables, principalmente en cuanto al tiempo ya que

en términos generales, un proceso escrito dura tres o cuatro veces más que uno de carácter oral, lo que ha quedado corroborado con los estudios comparativos elaborados por el Instituto Latinoamericano de las Naciones Unidas para la Prevención del Delito y Tratamiento del Delincuente (ILANUD)." (González, 2013: 194-195)

No obstante, algunos legisladores han considerado que la duración de los juicios puede repercutir en el imputado, puesto que los casos se pueden realizar hasta en una sola audiencia. Eventualmente se comprobará, con el paso del tiempo, si no son violadas las garantías individuales de los ciudadanos y si favorece los principios de democracia en el país.

2.4 Sistema de Justicia Penal Acusatorio

En principio, el Sistema de Justicia Penal Acusatorio (también conocido como los Juicios Orales) es, para el país, un nuevo esquema de impartición de justicia que sustituye al sistema inquisitivo penal. Presume un avance contra la corrupción y las fallas de los sistemas anteriores.

Este sistema tiene sus orígenes en la Europa del siglo XVII, con el propósito de impedir los abusos y arbitrariedades de los antiguos sistemas de justicia con bases grecorromanas. Este nuevo sistema tiene como prioridad salvaguardar las garantías individuales de los ciudadanos, así como la protección de la libertad.

Si bien no en todos los países el sistema penal acusatorio posee las mismas características, una de las particularidades en la que todos coinciden es en la oralidad. Los juicios, donde serán presentadas las pruebas, los expedientes y las declaraciones de las partes serán a través de audiencias públicas. Además, el procedimiento es mucho más corto.

Existen, de acuerdo con el Código Nacional de Procedimientos Penales (2014: 89), tres fases internas dentro de este nuevo sistema.

- Investigación, dividida en dos etapas:
 - Investigación inicial: la cual parte desde que se presenta la denuncia y termina cuando el imputado queda a disposición del Juez.
 - Investigación complementaria: es la investigación realizada a partir de la imputación y hasta que se cierre la misma investigación.
- Intermedia: o de preparación del juicio, que viene a partir de que se formula dicha acusación y hasta la apertura del juicio.

- Juicio: parte desde que se recibe el auto de apertura y hasta que el tribunal de enjuiciamiento (de uno a tres jueces) dicta la sentencia.
- Finalmente, la fase de la ejecución de penas consiste en la vigilancia, por parte del juez, de que las sentencias se realicen.



Ilustración 3 Fases internas del sistema de justicia penal.

El sistema acusatorio tiene como base los siguientes principios (Constitución Política de los Estados Unidos Mexicanos)

- Principio de publicidad: Permite que las audiencias permanezcan abiertas al público, que toda la información sea de escrutinio público. Es decir que todas las partes estén enteradas de todos y cada uno de los pasos del proceso.
- Principio de contradicción: Permite a la defensa contradecir pruebas de cargo, es decir que "se ha establecido la prohibición de dictar sentencias si las pruebas no son desahogadas en el juicio" (García, 2014: 23), por lo que este principio no permite que se violen las garantías individuales de los ciudadanos.
- Principio de concentración: Es la reunión en un mismo acto de cuestiones determinadas, es decir, el análisis y verificación de las pruebas en un mismo momento que permitan la agilidad y eficacia del sistema
- Principio de continuidad: al estar concentrado en una audiencia se le dará un mejor seguimiento, ágil, rápido y continuo sobre el proceso en cuestión.

- Principio de inmediación: Es la vinculación del juez con las partes, por lo que el Juez tiene la obligación de “recibir en forma directa la prueba” (García, 2014: 24). Esto para que el juez conozca de manera directa el proceso. Este principio es esencial para este trabajo, ya que el desahogo de las pruebas se realiza a través de una comunicación de los jueces con los especialistas.

De manera secundaria, trabajan también los siguientes principios:

- Principio de Oralidad: indica que toda audiencia que se desarrolle a lo largo del proceso deberá realizarse de manera oral.
- Principio de Oficialidad: El inicio del proceso dependerá de que se dé el supuesto y se encuentre previsto en la norma
- Principio Acusatorio: Supone la existencia de una serie de limitaciones o condicionantes procesales, lo que impide agravar la situación del acusado. No obstante, se mantienen los términos de acusación generales, pero se trabajan de manera imparcial.
- Principio de Oportunidad: impide adelantar el proceso penal, con la finalidad de eliminar cualquier tipo de ventaja que se tenga contra el acusado.
- Principio del Juez establecido por la ley: Se elige al enjuiciador que ha sido seleccionado conforme a las reglas, y que no se debe delegar el cargo a otros.
-
- Libre valoración de la prueba: Indica que el Juez debe poseer un criterio lógico durante el juicio.

- Presunción de Inocencia: La persona acusada se debe presumir inocente hasta que su culpabilidad sea comprobada y legalmente establecida.

Uno de los personajes más importantes en este nuevo sistema es sin lugar a dudas el juez, ya que tiene una participación completamente activa dentro del sistema. Su función consiste en:

garantizar el cumplimiento de los derechos de las partes, para lo que es necesario transformar el método de trabajo basado en la construcción de un expediente que guarda registro formal de todo, en un sistema que genere un entorno de condiciones que permita que esa función se cumpla (García, 2014: 27)

Es decir que, como se menciona en el principio de inmediación, el juez deberá comprender, analizar y observar las declaraciones de ambas partes de una manera lógica, así como valorar libremente las pruebas que se desahogan en juicio.

2.5 Trilogía de la investigación

González (2013) afirma que este nuevo proceso

exige una transformación profunda e integral de leyes, procedimientos, metodologías de investigación técnicas jurídicas, infraestructura, recursos humanos (jueces, ministerios públicos, peritos, abogados policías investigadores), es decir, implica la creación y establecimiento de una cultura completamente distinta en materia de justicia penal

Con el propósito de dar una solución justa y rápida, el trabajo de aquellos que son partícipes del procedimiento penal será de suma importancia, por lo que sus metodologías serán modificadas, así como la participación entre ellos mismos para la resolución de casos. A esta colaboración en conjunto la conocemos como Trilogía de la investigación, en la que

la Procuraduría General de Justicia [...] se integra por actos de investigación que el ministerio público realiza en su papel de investigador jurídico, por su parte la Policía de Investigación asume la tarea de investigar fáctico, y los peritos son investigadores técnicos (Garrido, 2017).

Dentro de esta trilogía, cada participante es el encargado de desempeñar determinadas actividades. Teniendo como punto de partida la investigación, la trilogía deberá hacer lo que esté en sus manos para establecer un juicio justo, eficaz y rápido. La investigación es la etapa procedimental durante la cual la trilogía investigadora realiza todas aquellas diligencias necesarias para estimar la existencia de un hecho delictuoso y la probable intervención delictiva del sujeto y optar por formular imputación o no.

Los tres responsables deben cumplir con las siguientes tareas:

Ministerio Público: Es el líder de esta trilogía, su trabajo consiste en “implementar retos y formas radicalmente nuevas de trabajo, con el propósito de instaurar un sistema penal preponderantemente acusatorio” (Vázquez, 2010).

Peritos: Es el especialista en la investigación, por lo que se encuentra en diferentes áreas, “reúne pruebas y las explica en audiencia pública, mediante un dictamen y, en su caso, responderá al interrogatorio que le formulen las partes” (Garrido, 2017).

Policía de Investigación: sirve como un apoyo para el Ministerio Público, “recibe denuncias, realiza detenciones, recaba información y ejecuta en el campo la investigación” (Garrido, 2017).

Asimismo, estos actores colaboran directamente con el juez y con las partes.

Además

El trabajo coordinado de la Trilogía Investigadora es relevante en las tres etapas del procedimiento penal acusatorio. La etapa de investigación inicia con una denuncia o querrela, donde los peritos y la Policía recolectan elementos que permitan al Ministerio Público determinar si resulta procedente iniciar un proceso penal contra una persona para fundamentar su acusación. En la etapa intermedia se ofrecen y admiten pruebas y depuran hechos controvertidos y, de ser el caso, se aportan elementos para que las partes opten por alguna salida alterna de solución a su controversia. Finalmente, en la etapa de juicio oral, se presentan las pruebas, y con esta base el Tribunal de Juicio Oral resolverá el conflicto. (Garrido, 2017)

La trilogía de la investigación, y en conjunto el nuevo sistema de justicia penal, pretende revitalizar “la procuración e impartición de la justicia más atenta a las necesidades sociales por una solución rápida, efectiva digna y transparente a los conflictos penales dentro de un marco democrático que respeta los derechos fundamentales y las garantías de todos los actores que participan en un litigio” (González, 2014). Es decir que hay una transformación de modelos, metodologías y

elaboración de resultados para los casos en materia penal, así como en la procuración de justicia en nuestro país. Es aquí donde estará presente la participación de los distintos operadores dentro de este sistema: policías, investigadores, jueces, peritos, entre otros, laborarán en equipos en conjunto para establecer la resolución de casos; esto quiere decir que la información con la que trabajen debe ser lo más clara posible para que el mismo sistema no pierda su carácter de eficacia y agilidad.

Todos estos métodos, además, saltan a la oralidad en vez de a la escritura, y estará presente a lo largo de todo el proceso, con la asistencia personal de las partes y del juzgador.

Estos cambios demandan al mismo tiempo nuevas normas o métodos de trabajo en el que los especialistas puedan exponer de la mejor manera sus técnicas e investigaciones y que el lenguaje especializado no sea un impedimento para hacer un buen trabajo. Razón imperativa para el presente trabajo.

CAPÍTULO III

3. METODOLOGÍA

Con lo visto anteriormente debemos entender que, en este nuevo sistema de justicia, existirán equipos de trabajo, los cuales tendrán que interactuar en las tareas que se establezcan para la resolución de casos; estos equipos estarán formados por especialistas de diversas áreas (jurídica, policial, en ciencias forenses) y serán los responsables de realizar procesos que tienen alcances especializados y, en consecuencia, puede existir ambigüedad entre el desarrollo de los procesos y técnicas elaboradas por cada uno de ellos; pongamos un ejemplo: el especialista en toxicología puede no entender términos que utilizan los peritos de acústica y viceversa o incluso puede que existan palabras que se utilizan en ambas áreas pero cuyo significado sea completamente diferente. A su vez, para el juez también es imposible conocer el significado de todos los términos de las áreas científicas que intervienen en el análisis de indicios. Todo ello puede provocar la falta de entendimiento entre los diferentes actores.

Por lo anterior, es imperativo contar con un instrumento lingüístico que permita reunir un conjunto de términos que, a partir de su definición, contribuya a entender de manera adecuada la tarea que cada especialista ha realizado.

Por ello, reitero el objetivo principal del proyecto en el que se inscribe este trabajo de tesis: elaborar un glosario de los términos de uso común o compartido entre las diferentes áreas de la ciencia forense. De esta manera se podrán

normalizar los alcances terminológicos, dar certeza entre los operadores que interactúan en el campo forense y objetivar el trabajo de cada uno.

Para llegar a esto, es necesario establecer una metodología que nos permita definir el conjunto de términos que se utilizan comúnmente en esta convergencia entre ciencia y derecho. La propuesta metodológica para esta selección se desarrolla a continuación.

3.1 Corpus: selección y características

La primera etapa consiste en la recopilación de un corpus de textos especializados. A partir de dicho corpus se obtendrán los primeros candidatos a término. Para hacer la selección de textos no podemos prescindir de la lingüística de corpus, en palabras de Berber Sardinha (2004), esta disciplina

ocupa-se da coleta e exploração de corpora, ou conjuntos de dados lingüísticos textuais que foram coletados criteriosamente com o propósito de servirem para a pesquisa de uma língua ou variedade lingüística. Como tal, dedica-se à exploração da linguagem através de evidências empíricas, extraídas por meio de computador.⁵

Por lo tanto, el corpus lingüístico es un conjunto de textos o materiales existentes en una lengua, reunidos para llevar a cabo cierto análisis. El corpus debe poseer ciertas características, tanto generales como particulares, ambas establecidas por su compilador o compiladores.

⁵ se ocupa de la recolección de datos y exploración de corpus, o conjuntos de datos lingüísticos, que fueron colectados cuidadosamente con el propósito de ser útiles a una lengua o variante de la lengua. Como tal, se dedica a la exploración del lenguaje a través de la evidencia empírica extraída a través de la computadora. (traducción propia)

Respecto a las características generales, Da Silva establece criterios de información que el corpus debe contener (en cuanto a exploración de la lengua): "I) el corpus debe estar compuesto por textos auténticos en lengua natural, II) el contenido debe ser seleccionado a fin de garantizar que el corpus tenga sus características propias y III) el corpus debe ser representativo" (2006: 221). Para Diplomová (2008: 8) las características generales o básicas de un corpus lingüístico son representatividad, tamaño final, formato y anotaciones. A continuación, se explica cada una (Diplomová, 2008: 8):

Representatividad: Se considera incluso un objetivo de los corpus lingüísticos (Sinclair, 1991: 17) ya que deben presentarse los casos comunes de la lengua y así ser útil en el estudio de la lengua.

Tamaño final: El tamaño del corpus depende principalmente de la extensión y su uso, por ello, es importante establecer estos dos últimos puntos, pues no se podrán modificar a lo largo del trabajo. Actualmente existen varios tipos de corpus de acuerdo a su tamaño, entre ellos destacan:

- **Corpus de referencia:** Su tamaño es grande ya que debe ser una visión panorámica de la lengua de estudio, son importantes para la elaboración de gramáticas y diccionarios.
- **Corpus de fragmentos textuales:** Son sólo muestras representativas de textos más grandes, el tamaño del fragmento depende del compilador.
- **Corpus monitor:** Este tipo de corpus no posee un tamaño fijo debido a que con el tiempo se le agregan más textos y se eliminan otros.
- **Subcorpus:** El subcorpus es una parte proporcional del corpus principal y debe contener las mismas características que este.

- Corpus especial: Es una visión específica a un fenómeno de la lengua (grupo social, especialidad, zona, etc.)
- Corpus bilingüe: Su objetivo es la traducción automática, se divide en corpus comparable y corpus paralelo.

Formato: El formato electrónico es importante para su clasificación, debido a la facilidad que tiene para manipularse. Existen varios procesos de transcripción a formato electrónico, ya que no sólo es necesario que sea presentado a manera de archivo, sino que debe ser posible acceder a sus elementos para su análisis y estudio (Barrón, 2007: 36). No se ahondará en estos procesos, ya que, además de ser de varios tipos, más adelante explicaré el que se utilizó para nuestro corpus.

Anotaciones: Se les llama así a las etiquetas del texto establecidas para su clasificación; están las extratextuales (título del texto, autor o autores, año de publicación) y las intertextuales (etiquetado morfosintáctico y lematización).

3.2 Herramientas de lingüística computacional utilizadas para el proyecto

Entendamos la lingüística computacional como un área multidisciplinaria, que establece sus bases en el procesamiento del lenguaje natural, como “un campo muy complejo y vasto, en el que intervienen disciplinas tan diversas como la lingüística, la ingeniería con raíces de sus disciplinas, como el cómputo, las telecomunicaciones y la electrónica, entre otras” (Sierra, 2009: 13).

Esta disciplina, aparentemente nueva en el país, surge en Estados Unidos en la década de los cincuenta debido a la necesidad de elaborar traductores

automáticos de lenguas extranjeras. No obstante, es reconocida hasta treinta años después con el propósito de desarrollar teorías del lenguaje aplicables a computadoras, esto con la ayuda de lenguajes de programación, estadística, ingenierías, matemáticas, entre otras. Con el paso de los años, esta disciplina agrandó sus intereses y propósitos, pues ya no estaba sólo relacionada con la informática, sino también con la misma lingüística:

“en las décadas siguientes el interés se centró en los corpus textuales, especialmente en inglés, al desarrollo de distintos lenguajes de programación con insumos de la lingüística teórica y de distintos programas para el análisis morfológico y sintáctico” (Sierra, 2009: 13).

Después, y a raíz del desarrollo del internet, surgió una necesidad de mejorar las tecnologías del procesamiento automático del lenguaje, con ayuda de la fonética y la morfología.

Si bien los países angloparlantes han desarrollado más rápido esta disciplina, México ha dado pasos agigantados, desarrollando herramientas que puedan facilitar el trabajo lingüístico. A continuación, presentaré las herramientas partidarias de estos métodos que utilicé para la elaboración del trabajo.

3.2.1 Gestor de corpus GECO

El gestor de corpus GECO es una herramienta elaborada en el Grupo de Ingeniería Lingüística (GIL). Parte de la necesidad de contar con un unificador de textos para poder aplicar las herramientas necesarias para su análisis, por lo que nos permite, como usuarios, tener un mayor manejo de la información de nuestro corpus, manipular las técnicas de análisis para el mismo, obtener resultados más

precisos y, en rasgos generales, acceder a otros corpus de investigación de manera gratuita (Sierra y Solorzano, 2015).

Para que el gestor de corpus pudiese unificar de manera precisa y sin ningún error que altere nuestros resultados, fue necesario cumplir con ciertas características de formato, por ejemplo: eliminar cuadros de texto, viñetas, tablas, hipervínculos, encabezados, pies de página, fuente, referencias, imágenes, entre otros. Una vez realizado este procedimiento, se suben los archivos en la red para posteriormente unificarlos, es decir, homologar los textos para trabajarlo como uno solo. Una vez hecho esto, podemos utilizar el extractor terminológico TermExt, el cual se encuentra instalado en la plataforma GECO-UNAM.

3.2.2 Etiquetador de textos Freeling

Uno de los procesos principales que se realizan cuando estamos trabajando textos es el etiquetado morfosintáctico, que consiste en otorgar las posibles categorías léxicas de las palabras, asignando la etiqueta con base en el contexto de la información proporcionada, a esto también se le conoce como POS (*Part of Speech*) tagging. Este tipo de etiquetado también está propuesto en distintos niveles de estandarización presentando categorías, valores y atributos. Por ejemplo, las etiquetas EAGLES⁶ asignan un código único para clasificar las palabras a partir de una lematización⁷. Las etiquetas que asigna corresponden a la categoría gramatical de la palabra:

⁶ EAGLES, por sus siglas Expert Advisory Group on Language Engineering Standards

⁷ La lematización es un proceso de carácter lingüístico en el que se asigna una forma flexionada a un lema ya establecido. Es decir que se le asigna género, número, conjugación, a un representante de todas las formas de una misma palabra).

1. Adjetivos
2. Adverbios
3. Artículos
4. Determinantes
5. Nombres
6. Verbos
7. Pronombres
8. Conjunciones
9. Numerales
10. Interjecciones
11. Abreviaturas
12. Preposiciones
13. Signos de puntuación

Tomando como ejemplo la primera categoría numerada, el conjunto de etiquetas

EAGLES nos brinda la posibilidad de asignar la siguiente información:

ADJETIVOS			
Pos.	Atributo	Valor	Código
1	Categoría	Adjetivo	A
2	Tipo	Calificativo	Q
3	Grado	Apreciativo	A
4	Género	Masculino	M
		Femenino	F
		Común	C
5	Número	Singular	S
		Plural	P
		Invariable	N
6	Caso	-	O
7	Función	Participio	P

Tabla 4 Clasificación de adjetivos a partir de EAGLES, (Universidad Politécnica de Cataluña, 2017).

Forma	Lema	Etiqueta
alegres	alegre	AQOCP00
alegre	alegre	AQOCS00
bonitas	bonito	AQOFP00
bonita	bonito	AQOFS00
bonitos	bonito	AQOMP00
bonito	bonito	AQOMS00
quemada	quemado	AQOFSOP

Tabla 5 Etiquetas de adjetivos a partir de EAGLES, elaboración de la Universidad Politécnica de Cataluña (2017).

Como puede observarse en las tablas 4 y 5 los adjetivos, así como cada categoría gramatical del español, pueden clasificarse con este conjunto de etiquetas. Este tipo de etiquetado facilita el trabajo de manejar un texto en diferentes soportes, como lo es el electrónico.

Ahora bien, Freeling utiliza las etiquetas EAGLES para asignar la categoría gramatical de una palabra de manera automática. Freeling es una "suite de herramientas para análisis del lenguaje desarrollada en la Universitat Politècnica de Catalunya" (Sierra, 2016). Freeling es una "librería de código abierto para el procesamiento multilingüe, que proporciona una amplia gama de funcionalidades de análisis para varios idiomas" (Padró, 2013). Esta herramienta se ha ido modificando, ampliando así sus posibilidades de análisis de textos y agregando nuevos idiomas, actualmente presenta la posibilidad de analizar textos en más de 8 lenguas distintas.

El etiquetado morfosintáctico de un texto resulta relevante en la tarea terminológica, pues nos permite detectar patrones en los términos. Es decir, si establecemos un etiquetado de este tipo, podemos observar a detalle qué estructuras gramaticales pueden predominar dentro de nuestros textos, y si coinciden de una u otra manera con las estructuras predeterminadas en la lengua española. Una ventaja del extractor terminológico que utilizaremos para el trabajo es que, previo a sacar las frecuencias, realiza un etiquetado bajo los criterios de Freeling y utilizando la clasificación EAGLES.

3.3 Procedimientos para la extracción terminológica

A grandes rasgos, la extracción de términos consiste en el análisis de corpus y la contabilización de palabras o conjuntos de palabras dentro del mismo. A continuación, presentaré los distintos métodos de extracción de términos y cuál resulta más conveniente para el trabajo.

3.3.1 Métodos de extracción

Idealmente existen dos métodos de extracción terminológica (Barrón, 2007: 7):

- La Interacción directa: Se encarga de establecer comunicación directa con un conjunto de especialistas, esto a través de encuestas, cuestionarios, entrevistas, etc.
- Interacción indirecta: Consiste en la extracción de términos a través de un corpus escrito, relacionado con el tema en cuestión. Dentro de éste existe la extracción manual y la extracción automática. La primera consiste en una activa participación del lingüista con los textos, leyéndolos arduamente y recolectando las palabras o conjuntos de palabras con mayor presencia en cada uno; en la segunda, interviene la presencia de herramientas automáticas, en las que, al trabajar con determinados algoritmos y paquetes, la extracción resulta más fácil, eficaz y mucho más rápida.

No olvidemos que, al ser herramientas que son nuevas, o al trabajar con paquetería que puede no ser propia de nuestra lengua o nuestro tema, sino más bien una adaptación, se pueden presentar algunos errores, por ello es recomendable combinar extracciones manuales y automáticas.

3.3.2 Extractores terminológicos

Un extractor de términos consiste en, como su nombre lo dice, la sustracción de candidatos a términos a través de un corpus dado. Aunque Bowker, L. y Pearson, J. (2002) mencionan que “las herramientas de extracción de términos suelen considerarse herramientas automáticas”, es importante mencionar que se deben de llamar candidatos a término, esto se debe a que “con frecuencia, muchos candidatos seleccionados por la herramienta no serán términos en sí (ruido) y a su vez habrá otros que sí lo sean y no hayan sido detectados (silencio)”. (Bowker, y Pearson, 2002)

La razón de que esto suceda, según Barrón (2007), es debido a que “no ha sido posible desarrollar aún un sistema de extracción de términos que funcione perfectamente (es decir, que extraiga todos los términos de un documento sin obtener ninguno que en realidad no lo sea)” (2007: 21).

No es posible dar certeza de los extractores que existen hoy en día, y debemos recordar que no todos pueden resultar aplicables al tema o incluso a la lengua en sí. Si bien es un hecho que la extracción automática de términos facilita el trabajo, los resultados que ofrecen no serán definitivos.

Barrón (2007: 21) menciona que existen dos enfoques a los que se inclina uno u otro extractor: el lingüístico y el estadístico. Los primeros requieren cierto conocimiento lingüístico para obtener los candidatos a término, mientras que los segundos priorizan la frecuencia de palabras o sintagmas dentro de un texto, sea o no especializado, sin necesidad de recurrir a información lingüística.

Existe un tercer enfoque, conocido como híbrido, éste combina el enfoque estadístico y el lingüístico para dar mejores resultados, por lo que su procedimiento es más elaborado y tiene más de una fase.

Tomar una decisión sobre qué extractor es más conveniente utilizar es importante, ya que se debe de buscar herramientas que sean neutrales o especialistas de nuestro tema para disminuir el margen de error, es decir, que se debe procurar siempre, al elegir un extractor, que éste se encuentre orientado hacia los temas que vamos a abordar, pues existen extractores que fueron elaborados principalmente para áreas en específico, como la medicina o la informática y quizás no resultan eficaces para áreas como la biología o la literatura.

A esto se han revisado algunos extractores que se han utilizado para la conformación de glosarios en lengua española, la mayoría institucionales, en esta parte se tomaron como ejemplo dos extractores: Terminus (Universitat Pompeu Fabra) y Yate in the Web (Universidad de Cataluña), siendo el segundo el más atractivo puesto que se pueden combinar estrategias de trabajo y se prioriza al uso intensivo de información semántica, sin embargo no realiza la gestión de corpus, ventaja que posee el primero; una desventaja para ambos es que están nutridos de archivos en su mayoría de español peninsular. Ambas herramientas son libres y pueden trabajarse en línea desde cualquier región.

3.3.3 Extractor de términos TermexT

La justificación para elegir TermexT se debe a que es una herramienta mexicana, de fácil acceso (a través de GECO), que, en apariencia, al estar sobre una base

lingüística y no temática será imparcial en el momento de la extracción de candidatos a término; esto resulta muy conveniente debido a que las ciencias forenses son consideradas como un campo interdisciplinario.

TermexT es un extractor resultado del proyecto de maestría de Barrón (2007), cuya base no es más que de carácter lingüístico, por lo que favorece la investigación al no inclinarse hacia una especialidad.

El extractor trabaja a través del algoritmo C-value/NC-value. La primera parte, C-value, consiste en la generación de una lista de candidatos a término por medio de un filtro lingüístico basado en la búsqueda de patrones sintácticos y la aplicación de una lista de paro para eliminar candidatos que contengan palabras que no se espera que formen parte de los términos del área tratada (Barrón, 2007). Posteriormente se calcula el potencial de cada candidato por medio de su longitud y frecuencia en el corpus. El NC-value, la segunda parte, considera el contexto en el que aparecen los términos para favorecer la calidad de los resultados (Barrón, 2007). En síntesis, el C-value realiza una combinación de la estadística y la lingüística, usando de esta última un etiquetado morfosintáctico que ya está predeterminado dentro del extractor, por lo que el C-value nos dará criterios de frecuencia, basándose en la lematización de cada una de las palabras del corpus. Por otro lado, el NC-value también maneja la frecuencia de las palabras a partir de un etiquetado previo; sin embargo, éste trata de ser más específico en cuanto a sus resultados y lo que hace es utilizar el contexto sintáctico en el que se encuentra la palabra. Es decir, no tomará en cuenta la frecuencia del término por el término en sí, sino por lo que lo rodea. Cabe mencionar en este punto que entre el C-value y el NC-value se forma lo que conocemos como lista de paro, ésta es diferente en cada

extractor: en nuestro caso, la lista se maneja a través de criterios meramente lingüísticos y ya predeterminados, quiero decir que se elaboraron junto con el sistema del extractor, por lo que no puede verse, eliminarse o modificarse. Dentro de esta lista de paro hay palabras o clases de palabras que el extractor por su conocimiento en lengua española no considera candidatos a término, por lo que hace una exclusión automática de la lista final. Con ambos algoritmos de trabajo dentro de un mismo extractor, TermexT se convierte en una buena herramienta para el proceso de la extracción de candidatos a término para la elaboración del glosario en Ciencia Forense.

Es importante mencionar que todo este proceso de extracción se realizó con el gestor de Corpus GECO como mediador, ya que, dentro de este Gestor, TermexT ya se encuentra instalado, por lo que facilitó el trabajo que se presenta a continuación.

CAPÍTULO IV

4. PROPUESTA METODOLÓGICA

A continuación, presentaré el proceso que se llevó a cabo para la selección de los términos que se incluirán en un glosario de términos de uso frecuente en Ciencia Forense. Se mostrará el uso de herramientas de lingüística computacional para dicha propuesta. Asimismo, es importante recordar que el enfoque desde el que se presenta es basado en la Teoría de las Puertas propuesta por Teresa Cabré y que ya se abordó en el primer capítulo de esta tesis.

Cabe mencionar que los resultados de esta tesis formaran parte del proyecto PAPIIT IA 303216 Glosario de Términos de uso frecuente en Ciencia Forense. Los resultados que aquí se presentan constituyen la base metodológica para el proceso de extracción de candidatos a término y la selección de los términos que se incluirán en el glosario.

4.1 Selección de corpus y primeros procedimientos

El corpus que se utilizó fue seleccionado por especialistas en ciencia jurídica, juicios orales, lingüística forense y otras áreas de las ciencias forenses. Está compuesto de la siguiente manera:

En cuanto a representatividad, los textos están en español como lengua natural, con variable mexicana, es un corpus sincrónico, bajo un registro

puramente escrito y de especialidad, por lo que son textos de carácter diafásico, ya que, al realizar un corpus técnico se está consintiendo el uso de un registro lingüístico formal, condicionado al marco de producción, al tema y al hablante (quien escribe, en este caso).

En cuanto al contenido, es un corpus especial, puesto que posee un tema específico (textos de carácter jurídico, relacionados al nuevo sistema de justicia penal en México), por lo tanto, presenta características representativas propias, además de la lengua natural.

En tamaño, el corpus está compuesto por 26 textos escritos que en total contienen 573,631 palabras. Son de especialidad normativa y todos vinculantes, es decir que "su contenido es obligatorio para los operadores jurídicos (Ministerio Público, peritos, policías, jueces, etc.)" (López-Escobedo *et al.*, en prensa: 6).

En relación con el formato, los textos son electrónicos, en documento de Word para su fácil manipulación, separados por archivo y sin tablas, pies de página, citas, encabezados o imágenes que puedan alterar su lectura. También se realizó una copia en formato txt⁸ con codificación de UTF-8⁹ con el propósito de que los acentos se consideraran como caracteres y no como errores del formato, pues en un principio el extractor presentó constantes errores y cortes de palabra debido a los acentos.

Respecto a las anotaciones, si nos referimos a las extratextuales, cada uno de nuestros textos contiene subtema, título del documento, autor (en este caso se hizo

⁸ TXT es una extensión para archivo de texto simple, texto llano el cual contiene únicamente caracteres simples caracteres, careciendo de cualquier tipo de formato tipográfico, es decir un texto sin formato.

⁹ UTF-8 es un formato de codificación de caracteres Unicode (formato estándar de codificación de símbolos), utilizando símbolos de longitud variable.

la selección de un autor colectivo debido al riesgo de especificar a uno solo, ya que en textos normativos no tenemos una certeza de un autor individual), tipo de texto, año de publicación o modificación y número total de palabras, además de una numeración en cuanto al orden de selección principal. Si nos referimos a anotaciones intertextuales, todos nuestros documentos se encuentran lematizados y etiquetados morfosintácticamente, tanto por parte del extractor como por cuenta propia (ahondaré en esto más adelante).

Además de las características propuestas por Diplomová, da Silva y Barrón; Arntz y Picht (1995: 46) proponen un modelo de análisis de textos especializados con base en guías de Hoffman (1985) y Nord (1988). Presento aquí el modelo y su relación con el corpus seleccionado.

Apariencia óptica: Los textos no contienen soportes no verbales considerables para su interpretación, existen índices, referencias (en los protocolos) y pequeñas tablas que tuvieron que ser eliminadas por cuestiones de formato. No hay variedad de medios tipográficos más que en algunos títulos.

Organización de contenido: Al referirnos a leyes, reformas, decretos y códigos estamos hablando de vocabularios de una especialidad, los formatos son similares todos vienen divididos en secciones y apartados, incluso los títulos son largos y específicos.

Léxico especializado: Todos los textos son de especialidad jurídica, ya que fueron seleccionados con este fin.

Plano sintáctico: La longitud media de las oraciones en los textos varía de manera drástica, ya que no fue su longitud, sino su tema lo que se seleccionó para el trabajo.

Formas personales: Predomina la tercera persona del singular. La presencia de la primera y segunda persona es escasa.

Aspecto estilístico: Es importante recalcar que todos los textos son de carácter jurídico, por lo que, como mencioné con anterioridad, siguen un formato distintivo y poco variable.

Estos aspectos nos sirven para identificar textos especializados, así como su grado de complejidad, ya que se debe observar el uso de recursos lingüísticos y extralingüísticos para ello.

A continuación, presento la tabla con los textos seleccionados para conformar el corpus del presente trabajo.

	Título	Palabras
1	Acuerdo A/009/15 por el que se establecen las directrices que deberán observar los servidores públicos que intervengan en materia de cadena de custodia.	6222
2	Acuerdo A/009/2013 del C. Procurador General de Justicia del Distrito Federal, por el que se establecen los lineamientos que deberán cumplir todos los servidores públicos que intervengan en la preservación del lugar de los hechos o del hallazgo y en la cadena de custodia de los indicios, huellas o vestigios del hecho delictivo, así como de los instrumentos, objetos o productos del delito.	3458
3	Acuerdo A/078/12 de la Procuraduría General de la República. Directrices que deberán observar los servidores públicos para la debida preservación y procesamiento del lugar de los hechos o del hallazgo y de los indicios, huellas o vestigios del hecho delictuoso, así como de los instrumentos, objetos o productos del delito.	10508
4	Acuerdo 06/2012 del Secretario de Seguridad Pública, por el que se emiten los lineamientos generales para la regulación del procesamiento de indicios y cadena de custodia en la Secretaría de Seguridad Pública.	3720
5	Acuerdo A/017/2011 Del C. Procurador General de Justicia del Distrito Federal, por el que se emite el Protocolo de Investigación Ministerial, Policial y Pericial del delito de Femicidio.	2098

6	Código de Procedimientos Penales para el Distrito Federal	79373
7	Código Nacional de Procedimientos Penales	64098
8	Código Penal Federal	84169
9	Código Penal para el Distrito Federal	73574
10	Constitución Política de los Estados Unidos Mexicanos. (Artículos 19 a 20).	1897
11	Guía Nacional de Cadena de Custodia	6388
12	Ley de la Policía Federal	13633
13	Ley Federal contra la Delincuencia Organizada	10408
14	Ley Federal de Justicia para Adolescentes	24248
15	Ley Federal para Prevenir y Sancionar la Tortura	1990
16	Ley General del Sistema Nacional de Seguridad Pública	28125
17	Ley General de Víctimas	30960
18	Ley nacional de mecanismos alternativos de solución de controversias en materia penal	7475
19	Ley de Seguridad Nacional	6243
20	Ley orgánica de la Procuraduría General de Justicia del Distrito Federal	17767
21	Ley orgánica de la Procuraduría General de la República	21066
22	Protocolo Homologado para la Investigación del Delito de Tortura. Procuraduría General de la República	43660
23	Protocolo Nacional del primer respondiente DOF 5 octubre 2015	4601
24	Protocolo Nacional de policía con capacidades para procesar	5000
25	Protocolo Nacional de traslados	3658
26	Protocolo para el tratamiento e identificación forense. PGR	19292

Tabla 6 Textos normativos que comprenden el corpus modificado. Basada en López-Escobedo; et. al. (en prensa).

Estos textos, además se trabajaron de manera individual para eliminar caracteres que entorpecieran el trabajo del gestor de corpus y del extractor, se eliminaron encabezados, pies de página, tablas, imágenes, entre otras características que eran parte del formato del documento.

Todos los textos se subieron al sistema de Gestión de Corpus GECO, elaborado por el Grupo de Ingeniería Lingüística de la UNAM, para homologar todos los archivos y manejarlos a razón de que actuaran como uno solo, pero sin perder sus rasgos individuales.



Ilustración 4 Vista principal de GECO, gestor de corpus elaborado por el Grupo de Ingeniería Lingüística de la UNAM.

Una vez dentro del gestor, y cuando éste ya no nos ha arrojado ningún error de lectura, se procede a trabajar la extracción de términos automática dentro del mismo programa.



Ilustración 5 Vista del extractor Temext trabajando (2017).

Como se ha mencionado, se ha utilizado la herramienta TermexT pues, además de estar instalada como herramienta dentro del gestor, sus características son útiles para generar una lista de posibles candidatos a términos para conformar el glosario de términos de uso frecuente en Ciencia Forense.

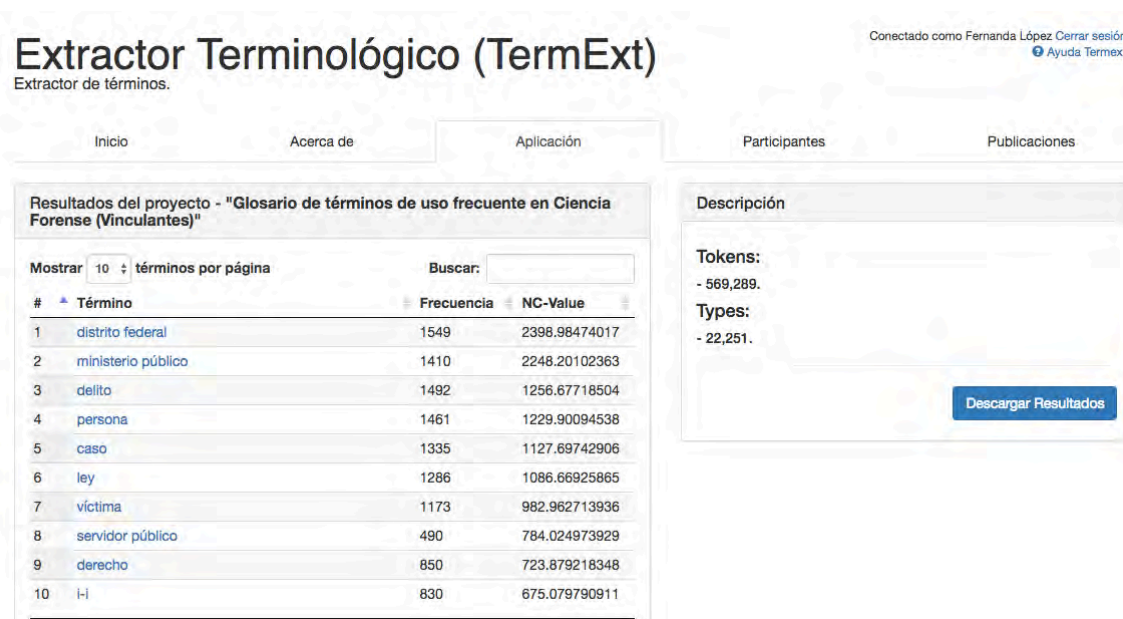


Ilustración 6 Vista de resultados del corpus que arrojó TermexT dentro del gestor (2017).

4.2 Primeros Resultados

Al subir los archivos y solicitar la extracción de los términos, la herramienta arrojó una lista de candidatos a término con los datos de C-value y NC-value previamente explicados, ordenándolos por frecuencia de aparición en el conjunto de textos.

El extractor automático arrojó un total de 2801 candidatos a término, es decir, aquellas unidades lingüísticas que cumplieron con los factores establecidos del C-value y el NC-value y que, además, no estuvieron presentes en la lista de paro predeterminada por el extractor. Sin embargo, dichos candidatos a término no necesariamente son términos, puesto que hay elementos ajenos a los algoritmos del extractor o como resultado del margen de error del mismo, que deben ser omitidos para hacer una lista más especializada y con características finas. Es por ello que se debe hacer una limpieza de estos candidatos. Esta primera lista arrojada por TermexT se revisó para comenzar con una depuración manual con base en criterios lingüísticos.

4.3 Limpieza de candidatos a término

Fueron varias las razones para la eliminación de varios candidatos a término de la lista original, pero las principales razones, y donde se eliminaba una considerable cantidad, fueron las siguientes.

Eliminación por	Explicación	Ejemplos
Diseño de norma	Los términos aparecían con frecuencia porque así lo solicitaba el formato del texto jurídico	capítulo, párrafo, artículo, I, II, III, IV, anexo 1, etc.
Paréntesis	El término contenía un paréntesis al final.	alcances), arrepentimiento),
Artículo	Aparecían los números de los artículos constitucionales y éstos no deben ser considerados candidatos a término, pero es entendible que aparecieran por la frecuencia de uso.	Artículo 3, artículo 40, artículo 55, artículo 64, artículo 70, artículo 82, artículo 9, etc.
Nombres de personas	Un nombre de persona no es candidato a término.	Enrique Peña Nieto, Andrés Manuel López Obrador, Ernesto Zedillo Ponce de León, Felipe de Jesús Calderón Hinojosa, Genaro García Luna, Jesús Murillo Karam.
Nombres de instituciones	No entran como término	Procuraduría General de la República, Secretaría de Seguridad Pública, Procuraduría General de Justicia del Distrito Federal, Consulado mexicano
Cargos públicos		Jefe de gobierno,

		presidente, procurador general, servidor público, secretario de seguridad pública.
Nombre de norma		Ley general de víctimas, Código nacional de procedimientos penales, ley federal contra la delincuencia organizada.
Siglas		df, dof, scjn, godf.
Artículo (determinado o indeterminado)		la, el, un.
Verbo	(en subjuntivo, de existencia, infinitivo, gerundio, etc.)	Señale, sea, son, identificar.

Tabla 7 Características para la limpieza de candidatos a término

Una vez analizada la lista de candidatos a término y de haber eliminado aquellos que aparecían con las características antes mencionadas, quedaron 1854, es decir se eliminaron 947 candidatos a término. El segundo depurado manual estuvo a cargo de especialistas en distintas áreas de la ciencia forense, quienes, a través de su conocimiento experto del área, realizaron una limpieza de términos, teniendo en cuenta que dentro de los textos se puede encontrar un lenguaje técnico y científico, propio de esta área (las ciencias forenses) que

son las ciencias de la aplicación al mundo jurídico para la solución de casos que requieren de la dictaminación de hechos que no se pueden descifrar con probanzas de apreciación común, como son las testimoniales o documentales, si no que requieren del conocimiento especializado (López-Escobedo, *et. al*, en prensa: 3).

La revisión consistió en el cruce de opiniones entre los expertos, tomando como referencia las proposiciones dadas en el lenguaje del derecho y la experiencia de los especialistas; por lo que, si existía un conceso del término, pasaba a la siguiente etapa y, si no era el caso, se eliminaba. Además, considerando que uno de los propósitos del glosario es resolver la ambigüedad de los términos, también fueron eliminados aquellos que no cumplían con esta característica, puesto que su significado resultaba claro. De esta manera se obtuvo un total de 177 términos.

4.4 Evaluación

Para finalizar, algo que fue posible detectar es que el margen de error del extractor de términos fue elevado, por lo que realicé un etiquetado manual de la lista final de los términos para observar si había una variación de estructuras sintácticas en ellos, la cual podría haber dificultado la extracción.

De acuerdo con Barrón (2007), existen ciertos patrones propios en la lengua que se verán representados en nuestros candidatos a término.

Obtuve las siguientes estructuras en cada una de las listas:

Estructura del término	Frecuencia
sus + adj + prep + sus	1
sus + prep + art + sus	1
sus + prep + sus + sus	1
sus + prep + verb	1
sus + sus	1
sus + sus + conj + sus	1
sus + adj + adj	2
sus + prep + sus + adj	10
sus + adj	33
sus + prep + sus	55
Sus	71
	177

Tabla 8 Estructura terminológica. Elaboración propia (2017).

Las estructuras coinciden con las que explica Barrón (2007), sin embargo, él indica que es necesario hacer una observación de aquellos términos que se presentan una vez y a qué corresponden, ya que pueden aparecer por el tipo de disciplina que se está analizando.

Los patrones más frecuentes en la lengua son (Barrón, 2007):

- Sustantivo
- Sustantivo + preposición + sustantivo
- Sustantivo + adjetivo

Como puede observarse, hubo compatibilidad entre las estructuras prototípicas del español y aquellas que presentaron los candidatos a término, a pesar de que se trabajó en un corpus de textos normativos cuyos conceptos en su mayoría son de carácter jurídico.

Esta primera etapa constituye la base para el futuro glosario, de la que hablaré más adelante. Una vez seleccionados los términos, el trabajo se muda al enfoque cualitativo, el cual consistió en una serie de entrevistas semiestructuradas a especialistas en el área de las ciencias forenses (peritos) quienes trabajan o trabajaron dentro del servicio público. Con base en la Teoría Comunicativa de la Terminología, coloquialmente conocida como La Teoría de las Puertas, de Teresa Cabré (2003), es pertinente conocer los términos a través de las mismas áreas científico-técnicas, pues “cada entrada presupone una teoría propia no contradictoria ni internamente ni con la condición interdisciplinaria de las unidades terminológicas” (Cabré, 2009), es decir que podemos acceder a los términos a partir de sus disciplinas y no desde las normas lingüísticas. Es por ello que estas entrevistas tuvieron el objetivo principal de

conocer términos que en el desempeño del ejercicio forense cotidiano les resultan confusos, o términos que utilizan coloquialmente con sus colegas para referirse a cuestiones técnicas o científicas de índole forense y que pudieran ser polisémicos o ambiguos. Así, también se buscaron términos que son referidos de forma diferente entre profesionales de su misma disciplina o por profesionales de otra disciplina. (López-Escobedo; et. al.: en prensa)

Al finalizar el procedimiento de las entrevistas, y realizar los análisis correspondientes, fue necesario hacer un nuevo análisis de nuestra primera lista de términos, por lo que se eliminaron y añadieron nuevos candidatos a término, teniendo como resultado una lista de 136 unidades terminológicas.

Dichas entrevistas semiestructuradas nos ayudan a conocer, a grandes rasgos, las inquietudes del personal responsable de trabajar y convivir en diferentes áreas.

I.

Formato de entrevistas para especialistas

Lugar de nacimiento
Edad
Género
Lugar de residencia en los últimos 5 años
Profesión
Especialidad
Años en servicio público

PREGUNTAS

En el desempeño de su ejercicio forense cotidiano ¿qué términos o palabras le resultan confusos?

En su trabajo diario ¿qué términos o palabras utiliza coloquialmente con sus colegas para referirse a cuestiones técnicas o científicas de índole forense?

¿Existen términos o palabras que sean referidos de forma diferente entre profesionales de su misma disciplina?

De la lista de términos que se le presentó ¿considera que hay otro u otros términos o palabras que por su ambigüedad derivan en una mala comunicación con los otros actores del proceso?

De la lista de términos que se le presentó ¿hubo algunos términos o palabras que no considerara ambiguos o confusos?

Podría definir los siguientes términos:

Acciones de coordinación

Acciones de investigación

Almacenamiento
Resguardo temporal

Análisis
Evaluación inicial

Identidad
Identificación

Método de búsqueda
Método criminalística

Cuando se solicita un peritaje, ¿qué términos o palabras considera usted que deberían ser claros para evitar ambigüedad entre lo que se solicita y lo que puede realizar?

¿Quisiera agregar un comentario?

Ilustración 6 Esquema de entrevista para peritos. Elaboración de López-Escobedo (en prensa).

II. CONCLUSIONES

Este trabajo partió de la premisa de que el nuevo sistema de justicia penal, al conllevar el uso de terminología especializada, producirá una barrera comunicativa entre las diferentes partes encargadas de impartir justicia en el país; esto quiere decir que debido a la cantidad de participantes especializados en diferentes áreas técnicas, científicas y jurídicas, el proceso puede llegar a entorpecerse debido a las particularidades del lenguaje que manejan los distintos grupos especializados dentro de las áreas antes mencionadas.

Esta investigación se elaboró con el fin de establecer las bases metodológicas para elaborar un glosario de unidades terminológicas que sean comunes o se compartan en los diferentes ámbitos donde se desarrollan las ciencias forenses, de esta manera se podrá elaborar un glosario con las acepciones necesarias o el significado que cubra todas las posibilidades que el significante desglosa. La Teoría Comunicativa de la Terminología (Teoría de las Puertas) propuesta por Cabré señala que se puede acceder y asimilar las unidades terminológicas partiendo de otras disciplinas de estudio ajenas a la lingüística, pues considera que las unidades terminológicas están compuestas de aspectos cognitivos, comunicativos, entre otros. Por esta razón, esta teoría es la base en la que se sustenta esta investigación, pues será mediante los textos de las diferentes áreas de las ciencias forenses, así como de sus actantes, que conoceremos a los candidatos a término y

posteriormente su significado o significados, dado que no se propone estandarizar sino desarrollar los significados de los términos, dentro de estas disciplinas.

El trabajo presentado demostró cómo los lenguajes legal y técnico-científico que participan en este nuevo proceso de impartición de justicia poseen una gran variedad de términos que tienen, dentro de sí, ambigüedad y polisemia, pues el significado puede variar e incluso no llegar a entenderse; es por ello que se procedió a trabajar con la elaboración de la lista de candidatos a término. Para ello, se realizó una serie de procesos técnicos, lingüísticos y jurídicos que van desde la selección de los textos ideales que puedan aportar mayor información, así como la limpieza de los candidatos que no cumplieran con los criterios establecidos para la elaboración de la lista.

Antes que esto fue necesario conocer los alcances de la terminología en nuestros días y cómo es que se relaciona con herramientas de lingüística aplicada, las cuales son responsables de apoyar en el trabajo de los investigadores al brindar los criterios para la correcta elaboración de corpus de trabajo. También fue necesario conocer los distintos programas encargados de la manipulación de textos digitales, etiquetadores morfosintácticos y, desde luego, los extractores terminológicos, sus características y modos de uso. Se hizo la elección de los programas que nos pudiesen ayudar a brindar los resultados esperados, no obstante fue necesario saber cómo trabajan estos sistemas de manera interna, cuáles son sus procedimientos, bajo qué criterios trabajan, sus bases de información y, por ello, si éstos son los más favorables para la elaboración del proyecto en cuestión, pues recordemos que el glosario se caracteriza por la variedad de unidades terminológicas de las distintas áreas del proceso penal en México; por

lo que se tuvo en consideración un extractor que no fuera de una sola especialidad, sino que pudiera manejar diferentes áreas técnicas y científicas.

Cuando se inició la investigación se tenía la idea de que, al trabajar con herramientas computacionales, el trabajo de la extracción de términos sería sencillo y eficaz; sin embargo, y aunque lo fue en su mayoría, no nos dio los resultados esperados en un primer momento, ya que la herramienta arrojó una lista que superó los 2000 candidatos a término, razón principal por la que se hizo más de una depuración manual, considerando diversos criterios (esquemáticos, lingüísticos, jurídicos, por decir algunos). Uno de los puntos a considerar a lo largo de la investigación fue la lista de paro con la que funcionaba el extractor, así como los textos de los que se alimentó en un principio; es decir, su base tanto léxica y sintáctica. El principal problema de la metodología surgió con la extracción de los candidatos a término, por ello se proponen dos puntos clave para mejorar las futuras investigaciones en corpus digitales con este u otros extractores:

Observación y modificación de la lista de paro que tiene por omisión el extractor terminológico.

Al ver y manipular –agregar u omitir– las unidades lingüísticas que son retiradas por omisión en el extractor facilitaría la selección de los candidatos a término ya que la lista de paro dependería del proyecto y corpus que se trabajaran, y así disminuir los errores en los resultados de los candidatos a término que arroje la extracción automática. Pongo un ejemplo: durante la investigación, el extractor arrojó los números romanos correspondientes a los capítulos, secciones o párrafos del texto que se estaba analizando, debido a su frecuencia de uso en el corpus, cuando estos números son utilizados de manera común en los formatos de ley u

otros textos de índole jurídico. Al poder modificar la lista de paro, se podrían agregar este tipo de unidades para reducir la lista de candidatos a término arrojada por el extractor.

Mejorar la probabilidad de contexto del extractor terminológico en cuestión.

Si bien el extractor contiene el algoritmo NC-value que mide los contextos de estructura sintáctica para considerar a un candidato a término y agregarlo a la lista, sólo informa al usuario sobre la frecuencia de uso, pero no qué clase gramatical es la que rodea al candidato; es decir, marca los porcentajes de frecuencia, pero no indica cuáles son las palabras que comúnmente encierran al término. Por ello se propone que el contexto arrojado por el extractor no sea sólo visible de manera frecuencial sino que, como usuarios, podamos visibilizar las unidades o palabras por las que se rodea el término trabajado, de esta manera será sencillo observar cómo se desenvuelve el término dentro de los textos del corpus y nos brinda una posibilidad para conocer si dentro del corpus se ha definido el término, de qué manera, y si hay alguna variación entre la definición encontrada en un texto y otro, de esta manera facilitaría la búsqueda de ambigüedad y polisemia en los términos para generar las acepciones debidas en los significados.

La lista final de términos para el glosario, al someterse a limpiezas manuales, disminuyó y con ella se realizaron entrevistas semiestructuradas a especialistas de las diferentes áreas de las ciencias forenses y que han laborado en la función pública para poder agregar o eliminar términos de la lista, para posteriormente presentar su significado a partir de la teoría de las puertas, donde el termino será explicado a través de las áreas comunicativas en las que se

desenvuelve, ya sea conociendo éste a partir de su contexto en los escritos del corpus u otros documentos de carácter lexicográfico, o a través de voces reales de los especialistas de cada disciplina, trabajo que está por finalizarse y que, gracias a la elaboración de esta lista, pudo realizarse de manera eficaz.

III. PROPUESTA

Este proyecto tuvo por objetivo construir una propuesta metodológica para realizar una lista de candidatos a término para la elaboración de un glosario de términos de uso compartido en las diferentes áreas de la ciencia forense. A continuación, presento los pasos de dicha metodología.

1. Selección del corpus: esta selección debe cumplir con ciertas características en común que permitan homologar la información.
2. Tratamiento del corpus:
 - o Limpieza del corpus digital, que consiste en eliminar criterios estilísticos del formato de los documentos que pudieran llegar a entorpecer la lectura digital del texto (tablas, encabezados, pies de página, entre otros).
 - o Concentración de los textos para trabajarlos como uno solo: Se recurre a la herramienta GECO para poder manejarlo adecuadamente. La carga de los textos es a través de internet y por documentos en formato .docx
3. Extracción automática de los candidatos a término: a través de una herramienta que cumpla con los objetivos de la investigación. En el caso de este trabajo TermexT fue la herramienta seleccionada pues su lista de paro no depende de ningún área de especialización. De esta manera, los

candidatos que arroja no están condicionados más que por caracteres lingüísticos básicos.

4. Depuración manual de los candidatos a término: se debe realizar considerando determinados parámetros (estos procesos se desarrollan a partir de especialistas en el área de la investigación, jurídica, lingüística y filosófica, en este caso).
 - A partir de criterios estructurales y lingüísticos: para eliminar todo aquello que no sea considerado término y que sólo se encuentre en la lista de candidatos por cuestiones de frecuencia en el texto (nombres propios, títulos de ley, dependencias gubernamentales, infinitivos, artículos definidos e indefinidos, por decir algunos).
 - A partir de criterios jurídicos: para eliminar todo aquello que no corresponda al área de estudio en cuestión, de esta manera sólo quedarán las unidades que produzcan ambigüedad y polisemia, y que podamos considerar unidades terminológicas.
5. Realización de entrevistas a especialistas del área: para complementar la selección de los candidatos a término. De esta manera se añaden aquellos términos que forman parte del lenguaje común entre los especialistas. Las entrevistas aplicadas en este trabajo se realizaron a aquellos ciudadanos que estuvieron o están laborando en la función pública, en un área relacionada con el proceso de impartición de justicia en México.
6. Selección de términos: con la información obtenida de la extracción de términos y las entrevistas a los especialistas se hizo una selección de los términos que conformarían el glosario, sobresaliendo aquellos que

presentaran ambigüedad, polisemia y falta de significado entre los hablantes.

El establecimiento de esta forma de trabajo convierte la metodología propuesta en una metodología híbrida, en la que el peso no recae en la extracción manual o automática, sino que se trabajarán en conjunto y sobre diferentes pasos para así obtener los mejores resultados. Conforme se fue avanzando el trabajo se comprobó que ambos métodos son necesarios en conjunto para mejorar los procesos de extracción de candidatos a término y para, finalmente, establecer una metodología aplicable a todas las áreas, incluyendo a las multidisciplinarias.

Del mismo modo se sugiere la elaboración de listas de paro por corpus, de esta manera se podrán eliminar aquellas palabras que por frecuencia entren al carácter de candidato a término y se puedan reducir los pasos y el tiempo de depuraciones manuales.

BIBLIOGRAFÍA

- ARNTZ, R., Picht, H. (1995). *Introducción a la terminología*, trad. Amelia de Irazábal (et al) Madrid: Fundación Germán Sánchez Ruipérez-Pirámide.
- BARRÓN, L. (2007). *Extracción automática de términos en contextos definitorios*, tesis de maestría inédita, México: Universidad Nacional Autónoma de México.
- BERGMAN, M. (2003). *Delincuencia, marginalidad y desempeño institucional. Resultados de la encuesta a población en reclusión en tres entidades de la República mexicana*, México: CIDE.
- BOWKER, L. y Pearson, J. (2002). "Extracción de Términos" en *Working with Specialized Language: A Practical Guide to Using Corpora*, Londres: Routledge.
- CABRÉ, M. T. (1995) *La terminología. Teoría, metodología, aplicaciones*, Barcelona: Antártida.
- CABRÉ, M. T. (1995). "La terminología hoy: concepciones, tendencias y aplicaciones" en *Ciência da Informação*, volumen 24, número 3.
- CABRÉ, M. T. (2001). "Importancia de la terminología para la comunicación en las lenguas europeas" en *Lingua e stile*, vol. 36, número 2.
- CABRÉ, M.T. (2002). "Terminología y lingüística: la teoría de las puertas" en *Estudios de Lingüística del Español*, número 16.

- CABRÉ, M. T. (2005). "La Terminología, una disciplina en evolución: pasado, presente y algunos elementos de futuro" en *Debate Terminológico*, número 1.
- CANALES, E. (2006). "Los juicios orales ante el sistema actual", en *Metrópolis*.
- CARBONELL, M.; Ochoa, E. (2007). "¿Necesitamos reformas en nuestro sistema de justicia penal? Algunos indicadores empíricos y teóricos" en *Revista de la Facultad de Derecho de México*, T LVII, Número 248, julio-diciembre, Universidad Nacional Autónoma de México.
- CARDERO, A. M. (2003). *Terminología y procesamiento*, México: Universidad Nacional Autónoma de México-Facultad de Estudios Superiores Acatlán.
- CARDERO, A. M. (2004). *Lingüística y Terminología*, México: Universidad Nacional Autónoma de México-Facultad de Estudios Superiores Acatlán.
- CONSEJO de coordinación para la implementación del nuevo sistema de justicia penal para el estado de Jalisco (S/A). *Cuadro comparativo sistema mixto vs sistema acusatorio*, Jalisco, [en línea] http://sistemadejusticiapenal.jalisco.gob.mx/sites/sistemadejusticiapenal.jalisco.gob.mx/files/cuadro_comparativo_sistema_mixto_vs_sistema_acusatorio.pdf [13/05/2017]
- CORDERO, S. (2009). "Algunas consideraciones sobre el lenguaje común y el lenguaje técnico" en *Revista Artes y Letras*, XXXIII.
- CRUZ, O. (2004). *Historia del derecho en México*, 2ª edición, México: Oxford University Press.
- DA SILVA, I. M. (2006). "La utilización de la lingüística de corpus en la enseñanza de español: usos de pero y sino" págs. 219-228. [en línea] https://cvc.cervantes.es/ensenanza/biblioteca_ele/publicaciones_centros/PDF/rio_2006/26_dasilva.pdf [4/03/2017]

- DIPLOMOVÁ, B. (2008) *Los corpus CREA y CORDE en el contexto de los corpus lingüísticos* [en línea] https://is.muni.cz/th/qct7p/Diplomova_Prace.pdf [10/06/2017]
- DUCE, M.; Pérez R. (2001). "Seguridad ciudadana y reforma de la justicia penal en América Latina" en *Boletín mexicano del derecho comparado*, número 102.
- FERNÁNDEZ, S. (2011). *Variación terminológica y cognición. Factores en la denominación del concepto especializado*, tesis doctoral, Barcelona: Universidad Pompeu Fabra.
- FISHMAN, G. (1972). *Sociology of language*. The Hague: Mouton [trad. esp.: *Sociología del lenguaje*. Trad. de Ramón Sarmiento y Juan Carlos Moreno. Madrid: Cátedra.
- GARCÍA, M. (2001). *Estructura definicional terminográfica en el subdominio de la oncología clínica*, Granada, Universidad de Granada [en línea]: <http://elies.rediris.es/elies14> [15/01/2017]
- GARCÍA, Z. (2014). *La argumentación judicial sobre hechos en el juicio acusatorio*, México, Bosch.
- GARRIDO, P. (2017). "La Trilogía Investigadora en el Sistema de Justicia Penal Acusatorio" en *24 horas, el diario sin límites*, 18 de septiembre 2017, [en línea] <http://archivo.24-horas.mx/la-trilogia-investigadora-en-el-sistema-de-justicia-penal-acusatorio/> [11/10/17]
- GONZÁLEZ, D. (2013). *Lógica, retórica y argumentación para los juicios orales*, México, Fontamara.
- HALLIDAY, M. (1990). *Language, context and text*. Oxford: Oxford University Press.
- LÓPEZ-ESCOBEDO, F.; et al, (en prensa). "En búsqueda de un Lenguaje Común entre la Ciencia y el Derecho: el Caso de la Ciencia Forense en México" en *Debate terminológico*, Número 17.

- MONTES, M. A. (2009). *Análisis teórico de la reforma constitucional del 18 de junio de 2008 en relación a la implementación en México del sistema penal acusatorio oral*, tesis de maestría, Hermosillo: Universidad de Sonora.
- MORALES, M. J. (2004). "Lenguaje y conocimiento común y especializado" en *Interam Bibliot*, vol. 27, número 1, enero-julio.
- Moreno F. (1999). "Lenguas de especialidad y variación lingüística" en *Lenguas para fines específicos (VI)*. Investigación y enseñanza, Alcalá de Henares, Universidad de Alcalá, 1999 pp. 3-14.
- MUNGO, F. (2004). *Teoría de la Interpretación jurídica*. Querétaro: Fundación Universitaria de Derecho, Administración y Política.
- PADRÓ, L. (2013). "Analizadores multilingües en Freeling", Centro de investigaciones TALP, Universidad Politécnica de Cataluña [en línea] <http://nlp.lsi.upc.edu/papers/padro11.pdf> [15/05/2017]
- RABASA, E. O. (1998). "El jurado popular en las constituciones de 1857 y 1917. Antecedentes, secuencia y consecuencia" en *Liber ad Honorem Sergio García Ramírez*, México: Universidad Nacional Autónoma de México.
- REAL ACADEMIA ESPAÑOLA (2009). *Nueva gramática de la lengua española*. Madrid: Espasa.
- REAL ACADEMIA ESPAÑOLA (2014). *Diccionario de la lengua española*, edición 23 [en línea],: <http://del.rae.es> [16/10/2017]
- SARDINHA, T. B. (2000). "Lingüística de corpus: histórico e problemática", en *DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada*, vol. 16, número 2.
- SIERRA, G. (2009). *Lingüística computacional en México. Investigación y desarrollo*, México: Universidad Nacional Autónoma de México.

- SIERRA, G.; Solórzano J. (2015) "GECO: Sistema de gestión de corpus" en *Actas del VII Coloquio de Lingüística Computacional*, México, Facultad de Filosofía y Letras-Universidad Nacional Autónoma de México [en línea]
<http://ru.ffyl.unam.mx/handle/10391/4775> [14/04/2017]
- SINCLAIR, J. M. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford: University Press.
- SUPREMA CORTE DE JUSTICIA DE LA NACIÓN (2016). *Principios rectores del sistema acusatorio* [en línea]
<https://www.sitios.scjn.gob.mx/cursoderechopenal/sites/default/files/Lecturas/Principios%20rectores%20del%20sistema%20acusatorio.pdf>
[13/08/2017]
- VÁZQUEZ, C. (2010). "El Ministerio Público y la Trilogía Investigadora", en *Revista semestral del consejo de coordinación para la implementación del sistema de justicia penal*, año 1, noviembre 2010, número 2 [en línea]
<http://www.accesocapacitacion.com/estudiantes-en-accion/44-recursos/recursos/159-ministerio-publico-trilogia-investigadora>
[15/06/2017]
- VIVALDI, J. (2001). *Extracción de candidatos a término mediante combinación de estrategias heterogéneas*, tesis doctoral, Barcelona, Universidad Politécnica de Cataluña.
- (2017) *Constitución Política de los Estados Unidos Mexicanos* [en línea]
http://www.diputados.gob.mx/LeyesBiblio/pdf/1_270818.pdf
- (2014). Código Nacional de Procedimientos Penales [en línea]
http://www.diputados.gob.mx/LeyesBiblio/pdf/CNPP_170616.pdf
[20/08/2018]

ANEXOS

ANEXOS

Lista final de Términos para la elaboración del glosario

Término	Término
acciones de coordinación	condiciones materiales y ambientales
acciones de investigación	confidencialidad
acreditación	consentimiento
acusación	conservación
agresión real	contacto directo
alegatos	contaminación
almacenamiento de indicios	datos confidenciales
alteración del indicio	datos sensibles
alteración de evidencia	depósito de indicios
alteración del lugar	descubrimiento probatorio
análisis de indicios	destrucción de evidencia
análisis de muestras	diagnóstico
análisis de evidencia	dictamen pericial
análisis del caso	disposición inmediata
aseguramiento de personas	disposición final
aseguramiento de indicios	disposición material
autenticidad	elemento material probatorio
bases de datos	elemento probatorio
bodega de indicios	elementos constitutivos del delito
búsqueda de indicios	elementos de prueba
cadena de custodia	embalaje
carácter de dato	entrega de indicios
ciencia forense	equipo de protección
comprobación de los hechos	especialista en la materia
comprobación de las proposiciones aseverativas de hechos	estándar de prueba
condiciones de preservación	estándar probatorio
condiciones de seguridad necesarias	etiqueta

evaluación inicial	métodos de observación
evidencia	métodos de fijación
examen de peritos	metodología criminalística
experticia	mismidad
fiabilidad	muestra
hallazgo	muestras representativas
hechos	muestreo
hipótesis de los hechos	ofendido
huella	ofrecimiento de pruebas
identidad	opinión científico técnica
identificación	opinión técnico científica
imputación	peritaje
indicio	perito
indicio principal	preservación de muestras
informe pericial	probabilidad de coincidencia al azar
informe policial homologado	probabilidad de error
inmediata disposición	procesamiento de indicios
instrumentos de derechos humanos	protección del lugar
integralidad	protocolo de cadena de custodia
integridad	protocolos de actuación
investigación preeliminar	protocolos de investigación
investigación complementaria	prueba
investigación criminalística	prueba atinente
investigación de delitos	prueba científica
levantamiento de evidencias	prueba pericial
levantamiento de indicios	pruebas pertinentes
localización de evidencias	recepción de indicios
localización de indicios	recolección
lugar de intervención	resguardo de indicios
lugar de los hechos	resguardo temporal
lugar de realización	sitio de intervención
lugar del hallazgo	técnicas de búsqueda
manual de investigación	teoría del caso
medio de comisión del delito	testigo experto
medios de prueba	traslado de indicios
métodos de búsqueda	trazabilidad
valoración de la prueba	valor probatorio
valoración de la prueba	
valoración probatoria	

vestigio	
víctima directa	
víctima indirecta	
víctimas potenciales	

IV.

V.