

La terminología para indizar documentos en trabajo asalariado de la mujer

Camelia Romero Millán*

Catalina Naumis Peña**

*Artículo recibido:
16 de febrero de 2015*

*Artículo aceptado:
12 de enero de 2016*

RESUMEN

El problema a analizar en este trabajo es la respuesta de los lenguajes controlados utilizados para indizar los documentos en estudios de género en la vertiente de la participación de la mujer en el trabajo asalariado. Los estudios de género en México han logrado un espacio en el campo de las ciencias sociales para explicar y cambiar la construcción ideológica que prevalece en la sociedad. Durante más de dos décadas, como resultado de las investigaciones, los autores han acuñado un lenguaje para denominar características y condiciones de la vida de las mujeres. En este artículo se analizan las representaciones de los contenidos documentales en los catálogos de las bibliotecas de instituciones de

* Biblioteca Daniel Cosío Villegas, El Colegio de México. millianr@gmail.com

** Instituto de Investigaciones Bibliotecológicas y de la Información, UNAM. naumis@unam.mx

educación superior en México a través del análisis de los documentos, los términos de indización en los catálogos, el tratamiento del tema en los lenguajes controlados y el uso de las frecuencias de aparición de los términos significativos en los documentos para obtener términos de indización. Se supone que el análisis de la terminología desde estos puntos de vista permitiría proponer términos representativos de un ámbito emergente.

Palabras clave: Estudios de Género; Trabajo asalariado de la mujer; Lenguajes controlados; Modelo matemático de Zipf; Punto de transición de Goffman.

Terminology for indexing documents in wage labor of women

Camelia Romero-Millán and Catalina Naumis-Peña

ABSTRACT

This paper examines the response of the controlled language used to index gender studies documents, specifically in the area of women working for wages. Gender studies in Mexico has earned influence in the field of the social sciences by providing new explanations, while changing prevailing ideological constructs in society. As a result of two decades of research, the authors have developed a language to describe the characteristics and living conditions of women. This paper examines indexing terms used to represent content of documents in the catalogs of libraries of institutions of higher education in Mexico. This approach includes, the analysis of documents, indexing terms in catalogs, indexing terms in controlled languages and the use of the frequencies of the occurrences of the terms in the documents. The study assumes that the analysis of terminology from these standpoints will allow researchers to propose representative terms for this emerging area.

Keywords: Gender Studies; Controlled Languages; Zipf Mathematical Model; Goffman Transition Point.

INTRODUCCIÓN

El desarrollo de la ciencia marca el ritmo y orientación en las colecciones de las bibliotecas universitarias; los acervos son parte del sustento teórico de la investigación y para retroalimentarla es necesario hacer accesibles los documentos que los conforman. La incorporación de los estudios de género (EG) en la academia es una muestra de la apertura de nuevos campos de conocimiento en las ciencias sociales y las humanidades. Las instituciones mexicanas de educación superior (IMES) apoyaron el campo de los estudios de género, primero adquiriendo colecciones para apoyar los inicios de la nueva línea de investigación y posteriormente incorporando las publicaciones que se editaron por las propias IMES.

El avance de los EG en diferentes campos de estudio y el interés cada vez más amplio de investigadores y alumnos han incrementado la consulta de estas colecciones; sin embargo, las búsquedas especializadas enfrentan dificultades en la recuperación de los documentos generados en el tema. En este sentido el proceso de indización es fundamental y para llevarlo a cabo se requiere de herramientas como los lenguajes controlados. Como parte de la sinergia entre el avance de la ciencia y la organización de las colecciones, las repercusiones se deberían ver reflejadas tanto en los lenguajes controlados como en los sistemas de clasificación.

El objetivo del artículo es mostrar los primeros resultados de una investigación para determinar la precisión en la indización hecha sobre los documentos especializados en uno de los aspectos de los EG y la inclusión de los términos usados en la especialidad en los lenguajes controlados. El estudio de la representatividad de los catálogos comprobará cómo responden a la terminología. El campo de los EG es muy amplio porque comprende varias aristas de convergencia con aspectos sociales, políticos y humanísticos. En este trabajo sólo se aborda el estudio de la representación documental en los catálogos por temas y por razones metodológicas se limita a estudiar la terminología del trabajo asalariado en el cual participan las mujeres.

Los documentos que se analizan en el estudio fueron localizados en bibliotecas de las IMES. Los catálogos escogidos para analizar fueron aquellos donde la institución mantiene un programa de docencia o una línea de investigación en este campo de estudios.

Para responder a la inquietud de conocer el comportamiento terminológico en los catálogos de las bibliotecas y los lenguajes controlados en materia de EG se retomó la metodología utilizada en la indización automatizada, como es la aplicación del modelo matemático de Zipf y el cálculo del punto de transición de Goffman.

LOS ESTUDIOS DE GÉNERO EN LA ACADEMIA MEXICANA

La investigación dedicada al análisis de las desigualdades a las que se enfrentaban las mujeres en los diferentes ámbitos de su vida como la relación de pareja, el espacio de trabajo, la familia y la escuela se hizo en un principio por organizaciones no gubernamentales como CIDHAL (Comunicación, Intercambio y Desarrollo Humano en América Latina, A. C.), fundada por Betsie Hollants en 1969 y que albergó el primer centro de documentación para mujeres en América Latina (<http://www.cidhal.org/CURRICULUM%20CARPETA.pdf>).

El CIDHAL comenzó a difundir los primeros escritos elaborados por intelectuales interesadas en conocer a fondo las causas y consecuencias de la desigualdad hacia las mujeres. Los estudios de la mujer (EM), en un inicio, estuvieron enfocados al conocimiento de sus condiciones de vida (Barbieri, 1998: 103-139); en el marco de la Década de la Mujer de las Naciones Unidas (1976-1985) proliferó la investigación y tuvo un espacio en la agenda de fundaciones y agencias de financiamiento (Cardaci, Goldsmith y Parada-Ampudia, 2002: 247-261).

En México, en la década de 1970, la Universidad Nacional Autónoma de México (UNAM) transmitía el programa de radio *Foro de la Mujer*, en el que participaba Alaíde Foppa,¹ quien además impartía la cátedra Sociología de la Mujer en la Facultad de Ciencias Políticas y Sociales. En la Facultad de Psicología se abrió el Centro de Estudios de la Mujer (Cardaci, Goldsmith y Parada-Ampudia, 2002: 247-261). Las profesoras e investigadoras universitarias que incursionaron en el tema se propusieron extender, impulsar y crear espacios propios de los estudios en las universidades (Barbieri, 1998: 103-139).

En 1983 se creó el Programa Interdisciplinario de Estudios de la Mujer (PIEM) en El Colegio de México, encabezado por Elena Urrutia.² A finales de esa década los EM habían ganado un espacio en la academia, creció el número de cursos impartidos y las alumnas interesadas habían elaborado sus primeras investigaciones plasmadas en trabajos de tesis. En la Universidad Autónoma Metropolitana, unidad Xochimilco, se creó el área de investigación Mujer, identidad y poder; en la unidad Azcapotzalco de la misma universidad, el Taller de la mujer y en la Facultad de Psicología de la UNAM se abrió el Centro de Estudios de la Mujer (Cardaci, Goldsmith y Parada-Ampudia, 2002: 247-261).

1 Alaíde Foppa nació el 3 de diciembre de 1914 en Barcelona, España. Durante su exilio en México trabajó en la Facultad de Ciencias Políticas y Sociales y fue integrante del grupo de mujeres que fundaron la revista *Fem* en 1975. De visita en Guatemala, fue secuestrada el 18 de diciembre de 1980 (Rossi, 2000: 104-108).

2 Quien también participó con Alaíde Foppa en la fundación de la revista *Fem* (Urrutia, 1987: 9-11).

En 1986 se hizo un recuento del impacto del campo disciplinario en la academia. Para entonces ya existían en el Distrito Federal 17 bibliotecas que tenían colecciones especializadas en el tema (Bustos Romero, 1989). En 1999 se reportó un incremento del 53 % en los programas y centros de estudios de la mujer (Cardaci, Goldsmith y Parada-Ampudia, 2002: 247-261).

Lo que en un primer momento se denominó “feminismo” y después “estudios de las mujeres” hoy se llama estudios de género (EG), como se denomina en la mayoría de las universidades y así se utiliza en el presente escrito. El campo de los EG se consolidó después de haber avanzado por cuatro etapas (Montecino, 1996):

- constituirse como un campo de estudio en la academia,
- insertarse en las principales disciplinas del área social y humanística,
- generar interés curricular donde se incorporó la diversidad, e
- impulsar el avance por la globalización e internacionalización de los EG.

Los EG son un “segmento de la producción de conocimientos que se ha ocupado de ese ámbito de la experiencia humana” (Gomáriz, 1992: 2), han avanzado en forma transversal en el análisis del entramado social y se han articulado con factores como edad, estado civil, educación, etnia y clase social; su alcance en las investigaciones llegó a la economía, la antropología y la literatura, entre otras disciplinas.

En 20 entidades de todo el país las IMES (entre las que se encuentran universidades y centros de investigación) han incorporado en sus programas de estudio el área de investigación en EG. En la UNAM se ha profundizado en el estudio y al mismo tiempo se han desarrollado colecciones en las bibliotecas de algunas facultades y escuelas, entre ellas la Facultad de Psicología, la Facultad de Economía, la Escuela Nacional de Trabajo Social y especialmente en el Programa Universitario de Estudios de Género (PUEG), fundado en 1992, como un espacio académico especializado en el área.

En sus 22 años de trayectoria y además de impartir diplomados, el PUEG ha publicado investigaciones y traducciones. El Colegio de México también ha desarrollado una colección especializada a partir de la creación del Programa Interdisciplinario de Estudios de la Mujer, el cual tiene un camino recorrido de 31 años de investigación y como resultado de sus actividades un fondo editorial propio. Este tipo de instituciones son las que desarrollan investigación en el tema y que por lo tanto promueven colecciones que respaldan los estudios y crean la expectativa de una representatividad eficiente en los catálogos de las bibliotecas académicas.

Impulsar el desarrollo de las actividades académicas en EG requiere de sustento teórico y retroalimentación, motivo suficiente para incrementar la bibliografía en las colecciones de las bibliotecas. Debido a la extensión y profundidad en el tratamiento del tema, los profesionales de la bibliotecología se enfrentan con mayor frecuencia a la necesidad de organizar y hacer recuperable información cada vez más especializada. Sin embargo, la revisión temática de los catálogos muestra un vacío que no se corresponde con la existencia de documentos en las colecciones.

La descripción temática carente de representatividad en los catálogos, aunado a la práctica en el uso de los encabezamientos de materia, donde se comprueba un vacío terminológico en el tema, permiten suponer que se pueden encontrar mecanismos para hacer visible la terminología del área emergente y proponer términos a incorporarse en los lenguajes controlados.

Como se explicó líneas arriba, para realizar la investigación se decidió trabajar únicamente con el tema del trabajo asalariado en el cual participan las mujeres, una problemática vigente a la fecha y uno de los temas que cuestionaron los grupos feministas desde sus primeras investigaciones, en las que se buscaba reivindicar el trabajo que realizan las mujeres dentro y fuera de su casa. A partir de este cuestionamiento y como resultado del análisis del mismo se acuñó el término “doble jornada laboral” (Bonaccorsi, 2007).

La inserción de las mujeres en la economía ha tenido efectos transformadores en las actividades productivas y las oportunidades de trabajo para ambos sexos en el mercado laboral. Los indicadores de ocupación y empleo en México, al segundo trimestre de 2014, informan que la población ocupada por sector de actividad económica refleja una participación activa de las mujeres:

- Sector primario (agricultura, ganadería, silvicultura, pesca, minería): 6 074 812 hombres y 698 093 mujeres.
- Sector secundario (industria manufacturera, construcción): 8 973 719 hombres y 3 047 879 mujeres.
- Sector terciario (servicios financieros, médicos, educativos, servicios a empresas como los administrativos, procesamiento de datos y consultoría): 15 539 612 hombres y 14 932 805 mujeres. (Inegi, 2014)

Se observa en el sector terciario que la brecha en la ocupación de hombres y mujeres es cada vez más cerrada, como lo muestran diversas investigaciones que explican y analizan de qué manera se han trastocado los estereotipos de género en los ámbitos del trabajo, como el caso de las mujeres policías y en cargos de dirección, por mencionar unos ejemplos.

Desde la década de 1970 se empezaron a publicar las primeras investigaciones en torno a la participación de las mujeres en la economía del país. Entre las autoras destacadas se encuentran Elizabeth Jelin, Mercedes Pedrero, Helen Safa, Brígida García y Orlandina de Oliveira. En su desempeño laboral fuera de la casa, las mujeres como asalariadas se han enfrentado a la discriminación por su condición biológica, esto se refiere a los embarazos, la necesidad de cuidar a los hijos y a las personas adultas o enfermas en su familia.

Alberti Manzanares, Vázquez García y Zapata Martelo (2001) mencionan que las nuevas formas de acercarse a la realidad social se construyen a partir de nuevos conceptos, contruidos a su vez a partir de que los EG han visibilizado la presencia de las mujeres en los diferentes ámbitos sociales, como el educativo y la familia, por mencionar algunos.

METODOLOGÍA

Para seleccionar la muestra documental se analizaron otros trabajos de investigación terminológica con la sorpresa de que no existen propuestas en este sentido. Sin ser demasiados específicos en su planteamiento, Lara, Ham Chande y García Hidalgo (1979) señalan que el lenguaje de los investigadores es parte de la lengua culta, por lo que permite un entendimiento amplio y tiene la capacidad de servir como modelo de corrección de los hablantes. Además, los autores sostienen que para tener una muestra significativa del patrón de los hablantes se debe consultar una diversidad de textos, sin especificar cuántos.

En atención a este planteamiento se consultaron los catálogos (libros, índices de revistas, tesis) de las entidades académicas enlistadas en el siguiente apartado, que son las que tienen colecciones especializadas sobre EG y que contemplan la línea de investigación sobre trabajo asalariado de la mujer. Se diseñó una estrategia de búsqueda por tema en los catálogos. Los documentos obtenidos bajo el criterio temático fueron 110 títulos producto del trabajo académico, mismos que fueron organizados en una base de datos que incluye los documentos digitalizados y transformados a formato legible por computadora.

Se decidió hacer una prueba piloto sobre algunas de las fuentes seleccionadas. La selección de los documentos fue realizada con base en entrevistas a investigadoras del área presentándoles la base de datos con los registros bibliográficos de los documentos y pidiéndoles que escogieran los 10 más representativos de la terminología del trabajo asalariado de la mujer. Por último, se aplicaron dos software para recuperar la terminología de los documentos seleccionados y se les aplicaron mediciones para establecer la frecuencia.

Selección de las fuentes

Las fuentes analizadas se tomaron de los catálogos de las bibliotecas de las IMES. Las instituciones, además de hacer investigación en EG, trabajan el tema del trabajo asalariado de la mujer y se encuentran en 20 estados de la república (*Tabla 1*).

Baja California	Universidad Autónoma de Baja California
Baja California Sur	Universidad Autónoma de Baja California Sur
Campeche	Universidad Autónoma de Campeche
Chiapas	Centro de Investigaciones y Estudios Superiores en Antropología Social (CIESAS) Sureste Colegio de la Frontera Sur, Unidad San Cristóbal Universidad Autónoma de Chiapas
Chihuahua	Universidad Autónoma de Ciudad Juárez
Coahuila	Universidad Autónoma de Coahuila
Distrito Federal	CIESAS D.F. El Colegio de México Escuela Nacional de Trabajo Social, UNAM Programa Universitario de Estudios de Género, UNAM (PUEG) Facultad Latinoamericana de Ciencias Sociales, sede México Catálogo colectivo de la UNAM
Estado de México	Universidad Autónoma del Estado de México
Guanajuato	Universidad de Guanajuato
Hidalgo	Universidad Autónoma del Estado de Hidalgo
Jalisco	CIESAS Occidente Universidad de Guadalajara
Michoacán	El Colegio de Michoacán
Morelos	Centro Regional de Investigaciones Multidisciplinarias (CRIM) de la UNAM
Nuevo León	CIESAS Noreste
Oaxaca	CIESAS Pacífico Sur
Querétaro	Universidad Autónoma del Estado de Querétaro
San Luis Potosí	Universidad Autónoma de San Luis Potosí
Sonora	El Colegio de Sonora
Veracruz	CIESAS Golfo
Yucatán	CIESAS Peninsular

Tabla 1. Instituciones académicas que incluyen colecciones sobre EG en sus acervos

Para conocer y analizar la magnitud en la que se recuperaban o perdían los documentos se realizaron búsquedas con las palabras generales “estudios de género”, “mujer” y “trabajo”. Después de haber obtenido resultados con estos términos se utilizaron frases terminológicas del lenguaje de los autores como “estudios de género y trabajo”, “familia y trabajo”, “discriminación sexual en el trabajo”, “discriminación sexual en el empleo”, “división del trabajo por sexo”, “roles sexuales en el trabajo”, “papel del género en el trabajo” y “violencia en el lugar de trabajo”. Esta segunda búsqueda se hizo principalmente por las frases, las cuales incorporan palabras o conjuntos de palabras de una especialidad, y mediar así la comunicación profesional (Cabré, 1993).

Lo que se observó es que al ser más precisas las palabras arrojan escasos resultados. Se estableció la presunción de que en este caso se debía a un trabajo de indización de los catalogadores en que además de las listas de encabezamientos de materia habían consultado alguna otra herramienta lingüística.

En otras bibliotecas, esos mismos escasos documentos recuperados con términos más específicos tenían asignados términos de indización muy generales que no los hacían recuperables con el verdadero enfoque de los contenidos. Este resultado podría deberse a una mala asignación del catalogador, pero en varios casos se comprobó que no existían los términos que contemplaran los temas tratados ni siquiera en tesauros especializados. Se observó entonces la necesidad de sistematizar los datos en la medida en que se obtuvieron para analizarlos en conjunto.

Mediante la recuperación temática se obtuvieron 110 documentos, entre los cuales hubo libros, capítulos de libros, artículos de revistas arbitradas, tesis de maestría y tesis de doctorado. En la *Figura 1* se muestra la proporción de los diferentes tipos de documentos examinados: 23 artículos, 70 libros y capítulos de libros, nueve tesis de doctorado y ocho tesis de maestría.

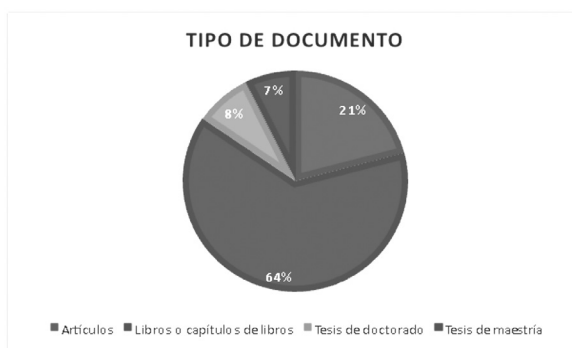


Figura 1. Proporción del tipo de fuente recuperada en los catálogos de las bibliotecas

ANÁLISIS DEL CONTENIDO DE LOS DOCUMENTOS SELECCIONADOS

El trabajo de investigación contempla el análisis terminológico de los 110 documentos identificados y constituye la tesis de doctorado de la autora. En virtud de que era necesario hacer una prueba piloto se consultaron 10 especialistas en el tema (investigadoras de El Colegio de México y del PUEG) y se les presentó la base elaborada con las 110 referencias para escoger las 11 más representativas en relación a la inclusión de la terminología del trabajo asalariado de la mujer. Las mayores coincidencias que fueron seleccionadas para presentar los primeros resultados fueron un libro publicado por El Colegio de Tlaxcala y una tesis de El Colegio de Sonora. La prueba piloto resultó tener elementos de interés que se presentan en este trabajo.

Los documentos se digitalizaron para transformarlos en archivos de texto y encontrar las palabras de mayor recurrencia. El software empleado fue WordSmith, el cual genera archivos en Excel donde ordena las palabras por frecuencia, de esta forma se ubican las palabras significativas utilizadas por los investigadores en sus escritos. Este software se complementó con el software *Notepad* para ubicar las palabras en el contexto de las frases terminológicas. Para obtener mejores resultados se retiraron de los textos las palabras vacías (Gil Leiva, 2008: 329); es decir, preposiciones, conjunciones y artículos, además se incluyeron en este grupo los nombres de personas y nombres geográficos.

En los archivos de Excel se aplicó el modelo matemático de Zipff (Urbizagástegui Alvarado, 1999):

$$r \cdot f = C$$

Donde:

r = Es el orden la palabra en la lista

f = Es la frecuencia o el número de veces que se repite una palabra en el documento

C = Es la constante

También se aplicó la fórmula para calcular el punto de transición de Goffman (Pao, 1978: 121-124) en los mismos archivos (Urbizagástegui Alvarado y Restrepo Arango, 2001: 83):

$$\underline{n} = \frac{-1 \pm \sqrt{1 + 8I_1}}{2}$$

Para conocer la brecha entre el lenguaje natural utilizado por los autores y el lenguaje controlado se hizo un comparativo con los encabezamientos de

materia propuestos para el tema en la lista del Library of Congress Subject Headings (LC), ya que son los utilizados en las bibliotecas académicas como apoyo para indizar documentos. Posteriormente se buscaron los encabezamientos equivalentes en español en las mismas fuentes que utilizan estas bibliotecas. La fuente de los encabezamientos de materia en español es el proyecto de catálogo bilingüe (inglés-español) en el que participa Mike Kreyche de la Universidad de Kent, Ohio (<http://lcsb-es.org/util.html?l=es>). Los encabezamientos utilizados en las diversas bibliotecas que conforman el sitio son las siguientes: Biblioteca Nacional de España (NE); Bilindex (BX); Catálogo de Autoridades del CSIC (CS); Queens Borough Public Library (Nueva York) (QB), y San Francisco (California) Public Library (SF). En la consulta de los encabezamientos de materia se pudo verificar la existencia o ausencia de los términos en las herramientas utilizadas en la indización.

Otras fuentes consultadas de lenguaje controlado para realizar el análisis comparativo fueron el *Tesaurus de género: lenguaje con equidad* (Tesaurus, 2006), publicado en 2006 (es la única edición hasta la fecha) por el Instituto Nacional de las Mujeres de México, y el *Tesaurus de la Red de Centros de Documentación y Bibliotecas de Mujeres del Estado Español* (Red de Centros de Documentación y Bibliotecas de Mujeres, 2014), actualizado hasta 2014. El primero se tomó en cuenta porque entre las fuentes utilizadas para su construcción se incluyeron términos usados por el centro de documentación de la institución que lo elaboró, tesauros de género así como enciclopedias y diccionarios. El segundo fue seleccionado por ser un documento actualizado, por el idioma y por su utilización por una red de bibliotecas y centros de documentación especializados en España.

Los objetivos de la consulta y comparación de las fuentes estuvieron centrados en varios objetivos específicos: el primero, identificar los términos asignados a los documentos en los catálogos donde fueron ubicados; el segundo, verificar la existencia de términos adecuados en los lenguajes controlados para describirlos; el tercero, encontrar la terminología utilizada por los autores que no aparecían en los encabezamientos. Los términos que hoy no forman parte de los lenguajes controlados podrían ser candidatos a incorporarse a partir de dos elementos: la garantía literaria y el uso del lenguaje como ya lo había mencionado Svenonius (2003: 824).

RESULTADOS

Entre los argumentos en contra de la indización automatizada está el hecho de que los sistemas automatizados no son capaces de identificar frases (Gil

Leiva, 2008); por ejemplo, el software *WordSmith* ordena sólo palabras. Sin embargo, esta aseveración ha cambiado porque existen herramientas de software que permiten asociar términos. Por este motivo, después de obtener las frecuencias se reconstruyeron las frases terminológicas con un software libre llamado *Notepad*. Este proceso complementario permite tener certidumbre respecto a los términos relevantes. A continuación se presentan dos ejemplos para mostrar los hallazgos.

a) *Maquila domiciliaria y subcontratación en México en la era de la globalización neoliberal*. Libro publicado por Plaza y Valdés y El Colegio de Tlaxcala. La *Tabla 2* presenta las palabras de mayor frecuencia utilizadas por el autor hasta el punto de transición de Goffman:

9R	Word	F	R*F	Goffman
1	TRABAJO	209	209	32904
2	INDUSTRIA	202	404	181.3945975
3	MAQUILA	187	561	180.3945975
4	DOMICILIARIA	150	600	90.19729875
5	PRODUCCIÓN	122	610	
6	DOMICILIO	111	666	
7	VESTIDO	103	721	
8	CAPITALISTA	94	752	
9	INDUSTRIAL	81	729	

Tabla 2. Resultado de las palabras de mayor frecuencia en el texto

En la aplicación del modelo matemático de Zipf se calculó la constante y se elaboró la gráfica a partir de la frecuencia más alta hasta donde se ubica el punto de transición de Goffman. La línea con inclinación hacia el eje de las X se debe a que hay palabras con menor frecuencia, las palabras en el eje de las Y son las de mayor frecuencia y representan parte del contenido del documento.

Para reconstruir los términos del discurso de los autores se utilizó el software *Notepad*, que permite hacer búsquedas por palabra y encontrar las frases terminológicas. La *Figura 3* muestra los resultados al trabajar con “trabajo a domicilio”, “trabajo por encargos” y “trabajo domiciliario”.

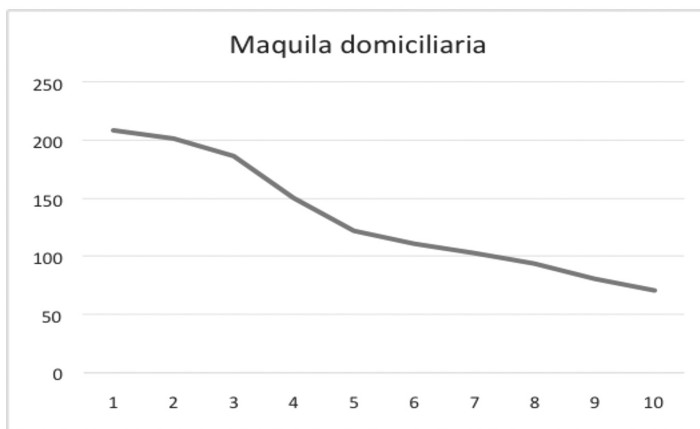


Figura 2. Modelo de Zip, el gráfico corresponde a la fórmula

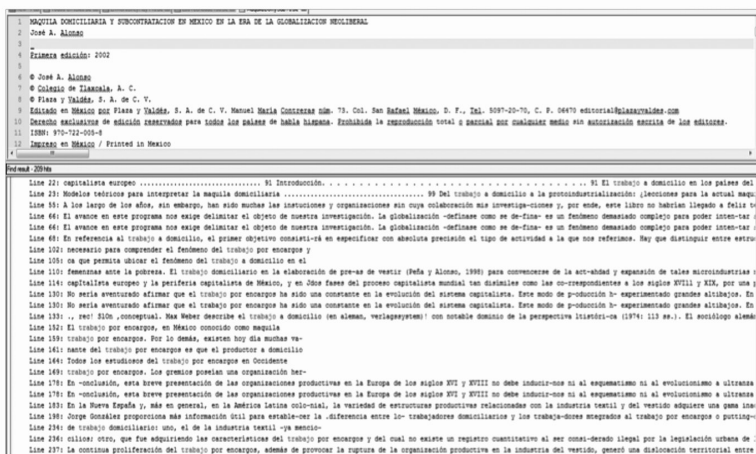


Figura 3. Uso del software Notepad para reconstruir las fases teminológicas utilizadas por el autor

Con los elementos presentados se seleccionaron los términos considera-dos relevantes para la indización, con los cuales se elaboró el cuadro comparativo en el que se incluyen como elementos de referencia los encabezamientos que se le asignaron al documento cuando se ubicó en el catálogo, los encabezamientos de materia de LC equivalentes en español y dos tesauros especializados. La *Tabla 3* presenta la comparación entre los lenguajes controlados y el lenguaje natural del autor.

Encabezamientos asignados	Enc. de LC equivalentes en español	Tesoro de género de INM	Tesoro de género de la red española	Lenguaje del autor
Mujeres trabajadoras de la industria del vestido -- México	Industria del vestido (BX)	Industria textil	Industria textil	Industria del vestido Industria textil
Trabajo a domicilio – México Mujeres -- Empleo – México	Trabajo a domicilio (BX, NE, CS); sector informal (NE)	Trabajo a domicilio	Trabajo a domicilio	Trabajo domiciliario Trabajo a domicilio Trabajo por encargos
	Trabajadoras de maquiladoras (SF) Industria maquiladora (NE)			Maquila domiciliaria Talleres de maquila Mujeres en la actividad maquiladora Producción textil domiciliaria Producción artesanal

Tabla 3. Comparativo de los lenguajes controlados y el vocabulario del autor

En el ejemplo se observa lo siguiente:

- a) Existen algunos encabezamientos de materia para describir con mayor precisión el contenido, como *trabajadoras de maquiladoras*.
- b) El tesoro de género no incluye terminología de la industria maquiladora, en tanto que en LC *trabajadoras de maquilas* ya forma parte del lenguaje controlado.
- c) El lenguaje utilizado por el autor es representado en el lenguaje controlado pero no es utilizado por el bibliotecólogo.

b) Tesis para obtener el grado de maestría por El Colegio de Sonora: *Los techos de cristal. Barreras de ascenso en la carrera laboral de las gerentas del sector privado en Hermosillo*.

En la *Tabla 4* se presentan las palabras de mayor frecuencia ubicadas en el texto hasta el punto de transición de Goffman.

En la aplicación del modelo matemático de Zipf se calculó la constante y se elaboró la gráfica, la cual se construyó desde la primera posición hasta donde se ubica el punto de transición. De acuerdo al comportamiento del uso del lenguaje y la frecuencia de los términos se dibuja otra vez una línea que desciende hacia el eje de las X.

R	Word	F	R*F	Goffman
1	MUJERES	538	538	73.34848328
2	TRABAJO	356	712	146.4618595
3	HOMBRES	298	894	73.23092975
4	SECTOR	148	592	
5	LABORAL	143	715	
6	PUESTOS	111	666	
7	MUJER	108	756	
8	BARRERAS	100	800	
9	SEGREGACIÓN	97	873	
10	SERVICIOS	97	970	
11	SEXO	96	1056	
12	GÉNERO	93	1116	
13	EMPRESA	87	1131	
14	ACTIVIDAD	86	1204	
15	DISCRIMINACIÓN	86	1290	
16	MAYOR	80	1280	
17	PERSONAL	80	1360	
18	CASO	75	1350	
19	TRABAJADORES	73	1387	

Tabla 4. Resultado de las palabras de mayor frecuencia en el texto

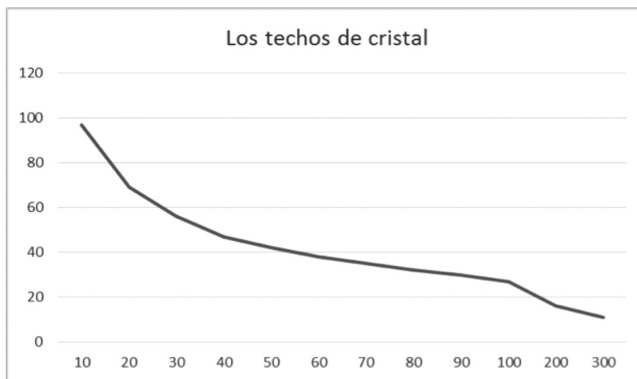


Figura 4. Modelo de Zipf, el gráfico corresponde a la fórmula

Se procede a la reconstrucción de los términos, en este caso “trayectorias laborales de mujeres”.



Figura 5. Uso del software Notepad para reconstruir las fases terminológicas utilizadas por el autor

En la *Tabla 5* se presenta la comparación entre los lenguajes controlados y el lenguaje natural del autor.

Encabezamientos asignados	Encabezamientos de LC equivalentes en español	Tesoro de género de INM	Tesoro de género de la red española	Lenguaje del autor
Mujeres -- Empleo	Mujeres trabajadoras (NE, CS) Condiciones de trabajo (NE,CS) Trabajo y familia (NE, SF, QB)		Personal directivo Altos cargos de la administración Profesiones no tradicionales	Puestos de dirección Condiciones de trabajo
				Trayectorias laborales de mujeres

		Mujeres empresarias/ ejecutivas Techo de vidrio	Empresarias Techo de cristal	Mujeres en posiciones de dirección
Discriminación sexual en el trabajo		Discriminación en el trabajo Segregación en el trabajo/empleo Segregación vertical	División sexual del trabajo	Discriminación laboral de género Segregación ocupacional por sexo División sexual del trabajo
				Hostigamiento laboral
Discriminación sexual	Discriminación sexual contra las mujeres (QB)	Discriminación sexual		Discriminación por sexo Desigualdad de género
	Mujeres - Salarios (NE) Discriminación salarial de la mujer (NE) Salarios (BX, NE, CS, SF, QB)	Discriminación salarial	Salarios (UP discriminación salarial)	Discriminación salarial

Tabla 5. Comparativo de los lenguajes controlados y el vocabulario del autor

Para este caso se observa lo siguiente:

- a) Los encabezamientos por los cuales fue localizado el documento en el catálogo se quedan en la descripción general de la discriminación sexual en el trabajo.
- b) No se describe el tema central, la discriminación en el trabajo hacia las mujeres en puestos de dirección.
- c) En ambos tesauros de género se tienen los descriptores *techo de vidrio* y *techo de cristal* o bien *mujeres en posiciones de dirección*, que son términos acuñados específicamente para describir las condiciones de trabajo de las mujeres con esas características; es decir, que no son reconocidas sus capacidades y no pueden rebasar una posición por el hecho de ser mujeres.

- d) Otro término es el de *discriminación salarial de la mujer*, el problema se presenta debido a que a pesar de existir un encabezamiento preciso no es utilizado por los catalogadores.
- e) No existe un término para describir las *trayectorias laborales de mujeres*.

CONCLUSIONES

En la muestra de los documentos examinados se pudo comprobar que existen frases terminológicas que no están incorporadas en los encabezamientos de materia; respecto a los tesauros, aun cuando se consultaron dos especializados se tienen algunos huecos en los términos de indización.

Otro de los hallazgos fue que a pesar de la existencia de términos especializados no son utilizados por los bibliotecólogos, por ejemplo, *discriminación salarial de la mujer y trabajadoras de maquiladoras*.

En algunos casos los encabezamientos de materia son más precisos que los términos en el tesauro para representar el mismo contenido conceptual.

Al aplicar el modelo de Zipf se comprobó que la frecuencia con la que aparecen las palabras en el texto puede emplearse para deducir los términos que contribuyen a representar los contenidos de un documento.

Respecto al cálculo del punto de transición de Goffman, se comprobó que también es un indicador a partir del cual se puede identificar una región de términos más recurrentes y los de mayor probabilidad para ser utilizados en la indización.

La identificación de términos simples no es problemática para la computadora; sin embargo, las frecuencias por sí solas no arrojan resultados definitivos. Para localizar frases es necesario recurrir al software *Notepad* con la finalidad de recuperar las frases terminológicas, acercarse más al lenguaje de los autores y hacer propuestas reales para la indización.

Otra comprobación es que se requiere de un análisis humano previo y posterior a los resultados para identificar los términos que describan el significado del contenido en los documentos.

Para el caso de la investigación respecto a la participación de las mujeres en el trabajo asalariado, se pudieron obtener términos significativos para representar con mayor precisión los documentos, lo que prueba la importancia y la posibilidad de replicar el proceso en otros espacios de interés en los que han incursionado las mujeres y han sido motivos de investigación.

Si bien los lenguajes controlados han sido actualizados para responder al desarrollo de las ciencias, no ha sido suficiente y se confirma que aún existen

vacíos terminológicos por cubrir ante el vertiginoso desarrollo de los EG, por lo que se requiere de una revisión exhaustiva y permanente ante fenómenos específicos.

Este artículo es una pincelada de lo que se puede profundizar en el campo de la representación de los contenidos documentales en EG ante la transversalización del tema en el ámbito social.

REFERENCIAS

- Alberti Manzanares, Pilar, Verónica Vázquez García y Emma Zapata Martelo. 2001. "Introducción", en *Género, feminismo y educación superior. Una visión internacional*, 13-29. México: Colegio de Posgraduados en Ciencias Agrícolas, Estado de México, British Council, ANUIES.
- Alonso Herrero, José Antonio. 2002. *Maquila domiciliaria y subcontratación en México en la era de la globalización neoliberal*. México: Plaza y Valdés, El Colegio de Tlaxcala.
- Barbieri, Teresita de. 1998. "Acerca de las propuestas metodológicas feministas", en *Debates en torno a una metodología feminista*, Eli Bartra (comp.), 103-139. México: Universidad Nacional Autónoma de México, Programa Universitario de Estudios de Género, Universidad Autónoma Metropolitana.
- Bejarano Celaya, Zenaida Margarita. 2001. "Los techos de cristal. Barreras de ascenso en la carrera laboral de las gerentas del sector privado en Hermosillo". Tesis de Maestría en Ciencias Sociales, El Colegio de la Frontera Norte, Hermosillo, Sonora.
- Bonaccorsi, Nélica. 2007. "Doble jornada laboral", en *Diccionario de estudios de género y feminismos*, Susana Beatriz Gamba, Tania Diz, Dora Barrancos, Eva Gilberti y Diana Maffía (coord.), 102-103. Buenos Aires: Editorial Biblos.
- Bustos Romero, Olga. 1989. "Los estudios sobre la mujer (y de género) en la UNAM: investigaciones y tesis", en *Estudios de Género y feminismo I*, Patricia Bedolla Miranda, Fátima Flores Palacios y Blanca E. García y García (comps.), 123-147. México: Fontamara.
- Cabré, María Teresa. 1993. *La terminología. Teoría, metodología, aplicaciones*. Barcelona: Antártida.
- Cardaci, Dora, Mary Goldsmith y Lorena Parada-Ampudia. 2002. "Los programas y centros de estudios de la mujer y de género en México", en *Feminismo en México: revisión histórico crítica del siglo que termina*, Griselda Gutiérrez Castañeda (coord.), 247-261. México: Universidad Nacional Autónoma de México, Programa Universitario de Estudios de Género.
- CIDHAL, A. C. S. f. *Comunicación Intercambio y Desarrollo para América Latina*. Fecha de consulta: 10 de febrero de 2014. <http://www.cidhal.org/CURRICULUM%20CARPETA.pdf>
- Gil Leiva, Isidoro. 2008. *Manual de indización: teoría y práctica*. España: Trea.
- Gomáriz, Enrique. 1992. *Los estudios de género y sus fuentes epistemológicas: periodización y perspectivas*. Santiago de Chile: Facultad Latinoamericana de Ciencias Sociales.

- Inegi (Instituto Nacional de Geografía Estadística e Informática). 2014. *Ocupación y empleo*. Fecha de consulta: 20 de agosto de 2014. <http://www3.inegi.org.mx/sistemas/temas/default.aspx?s=est&c=25433&t=1>
- Lara, F., Roberto Ham Chande y María Isabel García Hidalgo 1979. *Investigaciones lingüísticas en lexicografía*. Jornadas, 89. México: El Colegio de México, Centro de Estudios Lingüísticos y Literarios.
- Library of Congress Authorities. S. f. Fecha de consulta: 10 de Agosto de 2015. <http://authorities.loc.gov/>
- Library Subject Headings in Spanish. S. f. Fecha de consulta: 13 de julio de 2014. <http://lcsj-es.org/util.html?l=es>
- Maestría en Estudios de Género, Programa Interdisciplinario de Estudios de la Mujer, El Colegio de México. S. f. Fecha de consulta: 12 de noviembre de 2013. <http://piem.colmex.mx/index.php/maestria>
- Maestría en Estudios de la Mujer, Universidad Autónoma Metropolitana, Unidad Xochimilco. S. f. Fecha de consulta: 12 de noviembre de 2013. <http://dchx.xoc.uam.mx/estudiosmujer/>
- Montecino Sonia. 1996. "De la mujer al género: implicaciones académicas y teóricas". *Excerpta* (Centro de Estudios "Miguel Enríquez" Archivo Chile, CEME) 2 (abril): 1-11. Fecha de consulta: 11 de diciembre de 2014. <http://www.archivo-chile.com>
- Moreiro González, José Antonio. 2004. *El contenido de los documentos textuales: su análisis y representación mediante el lenguaje natural*. España: Trea.
- Notepad ++, GNU General Public License. S. f. Fecha de consulta: 12 de julio de 2014. <http://notepad.softonic.com/>
- Pao, Miranda Lee. 1978. "Automatic text analysis based on transition phenomena of Word occurrences". *Journal of The American Society for Information Science* 29 (3): 121-124.
- Programa Universitario de Estudios de Género, Formación académica. S. f. Fecha de consulta: 12 de noviembre de 2013. <http://www.pueg.unam.mx/index.php/formacion-academica/diplomado>
- Red de Centros de Documentación y Bibliotecas de Mujeres. 2014. *Tesaurus de género: tesaurus de la Red de Centros de Documentación y Bibliotecas del Estado Español*. Fecha de consulta: 12 de enero de 2013. http://www.juntadeandalucia.es/iam/catalogo/doc/web/tesaurus_genero.pdf
- Rossi, Annunziata. 2000. "Una semblanza de Alaíde Foppa". *Debate feminista* 22 (11): 104-108.
- Svenonius, Elaine. 2003. "Design of controlled vocabularies", en *Encyclopedia of Library and Information Science*, 822-838. New York: Marcel Dekker. Fecha de consulta: 19 de enero de 2014. http://polaris.gseis.ucla.edu/gleazer/260_readings/svenonius.pdf
- Tesaurus de género: lenguaje con equidad*. 2006. México: Instituto Nacional de las Mujeres.
- Urbizagástegui Alvarado, Rubén. 1999. "Las posibilidades de la Ley de Zipf en la indización automática". *B3: Revista Electrónica de Bibliotecología*. Fecha de consulta: 23 de agosto de 2013. http://www.academia.edu/480398/LAS_POSIBILIDADES_DE_LA_LEY_DE_ZIPF_EN_LA_INDIZACION_AUTOMATICA

- Urbizagástegui Alvarado, Rubén y Cristina Restrepo Arango. 2001. “La ley de Zipf y el punto de transición de Goffman en la indización automática”. *Investigación Bibliotecológica* 25 (54): 71-92.
- Urrutia, Elena. 1987. “Una publicación feminista Alaíde Foppa, siempre entre nosotras”. *Fem* 10 (49): 9-11.
- WordSmith Tools Lexically Analysis Software* version 6.0. 2011. <http://www.lexically.net/wordsmith/version6/>

Para citar este texto:

- Romero Millán, Camelia y Catalina Naumis Peña. 2017. “La terminología para indizar documentos en el trabajo asalariado de la mujer”. *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 73 (31): 191-211.
<http://dx.doi.org/10.22201/iibi.24488321xe.2017.73.57853>