

**Tipo de documento:** Tesis de maestría



**Escuela de Negocios.** Master in Management + Analytics

# Análisis de estereotipos de género en videos de YouTube para niños

Autoría: García Michel, María Macarena  
Año: 2024

## ¿Cómo citar este trabajo?

García Michel, M. (2024). Análisis de estereotipos de género en videos de YouTube para niños". [Tesis de maestría. Universidad Torcuato Di Tella]. Repositorio Digital Universidad Torcuato Di Tella.

<https://repositorio.utdt.edu/handle/20.500.13098/12903>

El presente documento se encuentra alojado en el Repositorio Digital de la Universidad Torcuato Di Tella bajo una licencia Creative Commons Atribución-Compartir igual 4.0 Internacional Deed  
Dirección: <https://repositorio.utdt.edu>



UNIVERSIDAD  
TORCUATO DI TELLA

Master in Management + Analytics

***ANÁLISIS DE ESTEREOTIPOS DE  
GÉNERO EN VIDEOS DE YOUTUBE  
PARA NIÑOS***

TESIS

María Macarena García Michel

Mayo 2024

Tutora: Laura Alonso Alemany



## Resumen

Desde la infancia, la exposición a producciones audiovisuales y escritas a través de medios de comunicación, entretenimiento y arte constituye un papel fundamental en la formación de las percepciones y actitudes hacia construcciones sociales como el género. Algunos estudios han mostrado cómo los discursos pueden perpetuar creencias erróneas y limitar las oportunidades para mujeres y hombres, profundizando las disparidades. Plataformas como YouTube, por su gran alcance, reflejan y contribuyen a moldear, amplificar y perpetuar tendencias culturales como los estereotipos. En consecuencia, se plantean preocupaciones sobre cómo la presencia de estereotipos de género en este medio fomentan situaciones de falta de equidad y representación.

El comportamiento no homogéneo de un sistema computacional con respecto a una variable de interés se conoce como sesgo, es por eso que hablamos de sesgo de género cuando los sistemas computacionales presentan un comportamiento diferente para diferentes géneros. En base a la premisa de que los estereotipos producen sesgos de género y éstos pueden perpetuar comportamientos nocivos para la sociedad, se plantea el siguiente interrogante: ¿Los videos de YouTube para niños evidencian sesgos de género?

La presente investigación aborda esta pregunta analizando si existen patrones y tendencias en los videos de YouTube para niños en inglés que reproduzcan estereotipos de género. A través de asociaciones no homogéneas entre palabras, se buscará identificar estereotipos presentes en el corpus. Es importante considerar que este trabajo se construye en un contexto cultural determinado, en particular, videos de entretenimiento para niños en inglés y las conclusiones a las que llega pueden no aplicarse a otros. Otra consideración importante es que, para adecuar el alcance a una tesis de maestría, se trató el género como binario: masculino y femenino.

Para el análisis se emplearon diversos métodos de procesamiento de lenguaje natural en un corpus curado compuesto de subtítulos de videos. En primer lugar, se trabajó con el Positive Pointwise Mutual Information (PPMI) y matrices de co-ocurrencias para obtener una medida cuantitativa de la fuerza de las asociaciones entre palabras. En segundo lugar, se infirió un modelo de Latent-Dirichlet-Allocation (LDA) con el objetivo de detectar temas prevalentes en el corpus y analizar si se descubren tendencias estereotipadas. Por último, se trabajó con el modelo de embeddings Word2Vec, la similitud coseno y el Word Embedding Association Test (WEAT) a fin de obtener representaciones vectoriales de los contextos de ocurrencias de las palabras con otras palabras y cuantificar la fuerza de asociación con cada género.

Los resultados obtenidos revelan la presencia de tendencias no homogéneas, es decir, desiguales, en la representación de los géneros masculino y femenino en los videos. Por un lado, se encontraron diferencias en la asociación de palabras que representan la expresión de emociones, con implicaciones que indicarían una disociación entre personas de género masculino y la manifestación de emociones. Por otro lado, se detectó una mayor asociación del género femenino con palabras que representan cuidados, lo cual supone una naturalización y por lo tanto un cierto mandato social en relación a las expectativas corporales hegemónicas, pero también un rol fuertemente ligado al cuidado de otros y del mantenimiento del hogar.

En conclusión, este estudio muestra que se pueden detectar representaciones diferenciadas para género masculino y femenino en videos de YouTube para niños en inglés. Estas representaciones pueden tener impactos sociales como invisibilización de opciones profesionales (hombres que se dedican a la enfermería, mujeres que se dedican a la cirugía) y personales (hombres a cargo de cuidados, mujeres que no expresan emociones). Es importante analizar críticamente la representación del género en estos medios con el fin de evitar la reproducción de patrones estereotipados que puedan influir negativamente en la percepción y desarrollo de las nuevas generaciones. Se recomienda adoptar estrategias interdisciplinarias, que integren enfoques de identificación de sesgos y perspectiva de género, con el objetivo de promover una representación más equitativa y diversa en el contenido audiovisual.

## Abstract

Since childhood, exposure to audiovisual and written productions through media, entertainment, and art plays a fundamental role in shaping individuals' perceptions and attitudes towards social constructs such as gender. Some studies have shown how discourses can perpetuate erroneous beliefs and limit opportunities for women and men, deepening disparities. Platforms like YouTube, due to their wide reach, reflect, shape, amplify and perpetuate cultural trends such as stereotypes. Consequently, concerns arise about how the presence of gender stereotypes in this media fosters situations of inequity and misrepresentation.

The non-uniform behavior of a computational system regarding a variable of interest is known as bias. This is why we talk about gender bias when computational systems exhibit different behaviors for different genders. Based upon the premise that stereotypes produce gender biases, perpetuating harmful behaviors for society, the following question is presented: **Do YouTube videos for children show gender biases?**

This research addresses the question by analyzing if there are patterns and trends in YouTube videos for children in english that reproduce gender stereotypes. Through non-uniform associations between words, the aim is to identify stereotypes present in the corpus. It is important to consider that this work is constructed within a specific cultural context, particularly entertainment videos for children in english, and the conclusions drawn may not apply in others. Another important consideration is that, to narrow the scope for a master's thesis, gender was treated as binary: male and female.

Various natural language processing methods were employed for the analysis on a curated corpus composed of video subtitles. Firstly, Positive Pointwise Mutual Information (PPMI) and co-occurrence matrices were used to obtain a quantitative measure of the strength of associations between words. Secondly, a Latent Dirichlet Allocation (LDA) model was inferred to detect prevalent themes in the corpus and analyze if stereotypical trends are discovered. Finally, the Word2Vec embedding model, cosine similarity, and Word Embedding Association Test (WEAT) were used to obtain vector representations of word occurrence contexts with other words and quantify the strength of association with each gender.

The results reveal the presence of non-uniform, i.e., unequal, trends in the representation of male and female genders in the videos. Differences were found in the association of words representing the expression of emotions, with implications suggesting a dissociation between individuals of the male gender and the expression of emotions. Additionally, a greater association of the female gender with words representing caregiving was detected, implying a naturalization and thus a certain social mandate regarding hegemonic bodily expectations, but also a role strongly linked to caring for others and doing homework.

In conclusion, this study shows that differentiated representations for male and female genders can be detected in YouTube videos for children in English. These representations may have social impacts such as invisibilization of professional options (men who dedicate themselves to nursing, women who dedicate themselves to surgery) and personal ones (men in charge of caregiving, women who do not express emotions). It is important to critically analyze gender representation in these media to avoid the reproduction of stereotypical patterns that may negatively influence the perception and development of new generations. Including interdisciplinary strategies that integrate bias identification approaches and gender perspectives is recommended to promote a more equitable and diverse representation in audiovisual content.

# Índice

|  |    |
|--|----|
| <b>1. Introducción</b>   | 7  |
| 1.1. Contexto  | 7  |
| 1.1.1. Estereotipos y niñez  | 7  |
| 1.1.2. Estereotipos y prejuicios en YouTube                            | 8  |
| 1.1.3. Análisis de sesgos y estereotipos con métodos de data science   | 9  |
| 1.2. Problema  | 10 |
| 1.3. Objetivo  | 11 |
| <b>2. Datos</b>  | 12 |
| 2.1. Estructura de los datos   | 12 |
| 2.2. Obtención del dataset   | 12 |
| 2.3. Curación y descripción de los datos                               | 13 |
| <b>3. Metodología</b>  | 15 |
| 3.1. Lista de palabras   | 15 |
| 3.1.1. EDIA (Estereotipos y discriminación en Inteligencia Artificial) | 17 |
| 3.2. Matriz de co-ocurrencias  | 20 |
| 3.3. Positive Pointwise Mutual Information (PPMI)                      | 21 |
| 3.3.1. PPMI con Laplace  | 22 |
| 3.3.2. PPMI en contexto  | 23 |
| 3.4. Latent-Dirichlet-Allocation (LDA)                                 | 23 |
| 3.4.1. Entrenamiento   | 25 |
| 3.4.1.1. Número de topics  | 25 |
| 3.4.1.2. Passes  | 27 |
| 3.4.1.3. Iterations  | 28 |
| 3.5. Embeddings  | 29 |
| 3.5.1. Word2Vec  | 29 |
| 3.5.2. Similitud Coseno  | 31 |
| 3.5.3. Word Embedding Association Test (WEAT)                          | 32 |
| <b>4. Análisis de Resultados</b>                                       | 34 |
| 4.1. Positive Pointwise Mutual Information (PPMI)                      | 34 |
| 4.1.1. Palabras asociadas a las emociones, sentimientos y cuidado      | 34 |
| 4.1.2. Palabras asociadas al trabajo y a lo doméstico                  | 36 |

|           |   |           |
|-----------|---|-----------|
| 4.1.3.    | PPMI con Laplace.....   | 39        |
| 4.1.4.    | Hipótesis EDIA .....  | 40        |
| 4.1.5.    | PPMI en contexto .....  | 42        |
| 4.2.      | Latent-Dirichlet-Allocation (LDA).....                          | 44        |
| 4.2.1.    | Resultados del modelo final .....                               | 44        |
| 4.2.2.    | Visualización de tópicos.....                                   | 45        |
| 4.3.      | Word2Vec.....   | 46        |
| 4.3.1.    | Similitud Coseno .....  | 46        |
| 4.3.1.1.  | Palabras asociadas a las emociones, sentimientos y cuidado..... | 47        |
| 4.3.1.2.  | Palabras asociadas al trabajo y a lo doméstico .....            | 48        |
| 4.3.2.    | Word Embedding Association Test (WEAT).....                     | 50        |
| <b>5.</b> | <b>Conclusiones .....</b>                                       | <b>51</b> |
| <b>6.</b> | <b>Futuras investigaciones .....</b>                            | <b>53</b> |
| <b>7.</b> | <b>Referencias .....</b>  | <b>55</b> |
| <b>8.</b> | <b>Apéndice.....</b>  | <b>58</b> |
| 8.1.      | Lista de canales consultados y número de suscriptores .....     | 58        |
| 8.2.      | Archivos de texto por canal consultado .....                    | 59        |
| 8.3.      | Palabras totales y tokens únicos.....                           | 60        |
| 8.4.      | 3 palabras más comunes en cada documento .....                  | 61        |
| 8.5.      | Frecuencia de tokens de la lista de palabras .....              | 101       |
| 8.6.      | Términos más frecuentes de los 4 tópicos más prevalentes .....  | 103       |

## Índice de Tablas

|          |  |    |
|----------|--|----|
| Tabla 1. | Definiciones de los conceptos incluidos en las listas de palabras .....                              | 16 |
| Tabla 2. | PPMIs para palabras asociadas a las emociones, sentimientos y cuidado .....                          | 35 |
| Tabla 3. | PPMIs para palabras asociadas al trabajo y a lo doméstico.....                                       | 37 |
| Tabla 4. | PPMIs con suavizado para palabras asociadas a las emociones, sentimientos y cuidado.....             | 39 |
| Tabla 5. | PPMIs con suavizado para palabras asociadas al trabajo y a lo doméstico .....                        | 40 |
| Tabla 6. | PPMIs promedio según género y grupo (palabras asociadas a las emociones sentimientos y cuidado)..... | 41 |
| Tabla 7. | PPMIs promedio según género y grupo (palabras asociadas al trabajo y a lo doméstico).....            | 41 |

## Índice de Figuras

|   |    |
|---|----|
| Figura 1. Proporción de ocurrencias de he y she .....   | 14 |
| Figura 2. Frecuencia de las 10 palabras más comunes en el corpus completo .....   | 14 |
| Figura 3. Visualización EDIA de palabras asociadas a las emociones, sentimientos y el cuidado .....                                 | 18 |
| Figura 4. Visualización EDIA de palabras asociadas al trabajo y a lo doméstico .....  | 19 |
| Figura 5. Nube de palabras para el pronombre he .....   | 21 |
| Figura 6. Nube de palabras para el pronombre she .....  | 21 |
| Figura 7. Valores de coherencia para el rango de 2 a 50 temas.....  | 26 |
| Figura 8. Valores de coherencia para el rango de 15 a 33 temas.....   | 27 |
| Figura 9. Valores de coherencia para el rango de 1 a 2,000 passes .....   | 28 |
| Figura 10. Valores de coherencia para el rango de 10 a 125 iterations .....   | 29 |
| Figura 11. Visualización de PPMIs para palabras asociadas a las emociones, sentimientos y cuidado .....                             | 35 |
| Figura 12. Diferencias de PPMIs para palabras asociadas a las emociones, sentimientos y cuidado ..                                  | 36 |
| Figura 13. Visualización de PPMIs para palabras asociadas al trabajo y a lo doméstico.....  | 38 |
| Figura 14. Diferencias de PPMIs para palabras asociadas al trabajo y a lo doméstico .....   | 38 |
| Figura 15. Porcentaje de negaciones en las co-ocurrencias para las palabras asociadas a las emociones, sentimientos y cuidado ..... | 43 |
| Figura 16. Porcentaje de negaciones en las co-ocurrencias para las palabras asociadas al trabajo y a lo doméstico.....              | 44 |
| Figura 17. Mapa de tópicos .....  | 45 |
| Figura 18. Top 30 de términos más relevantes .....  | 46 |
| Figura 19. Similitud coseno para palabras asociadas a las emociones, sentimientos y cuidado.....                                    | 47 |
| Figura 20. Diferencias en las similitudes promedio - Palabras asociadas a las emociones, sentimientos y cuidado .....               | 48 |
| Figura 21. Similitud coseno para palabras asociadas al trabajo y a lo doméstico.....  | 49 |
| Figura 22. Diferencias en las similitudes promedio - Palabras asociadas al trabajo y a lo doméstico ..                              | 50 |

# 1. Introducción

## 1.1. Contexto

### 1.1.1. Estereotipos y niñez

Los estereotipos de género son creencias ampliamente aceptadas acerca de cómo se supone que deberían ser los diferentes géneros de una cultura (Fiske et al. 2002). Algunos estudios han mostrado que estos, junto con los prejuicios, comienzan a formarse entre los 2 y los 4 años (Bigler y Liben 2007). La teoría de estos autores postula que los procesos cognitivos predisponen a los niños a adquirir y perpetuar estereotipos, llevando a la internalización de creencias arquetípicas sobre los roles de género en la sociedad. Esta investigación también explica que cuando los grupos sociales son etiquetados o tratados de manera diferente, los niños tienden a conceptualizar los segmentos como diferentes entre sí y a asimilar las expectativas estereotipadas asociadas a ellos. Es importante tener en cuenta que los estereotipos de género pueden influir en las acciones y decisiones de las personas incluso fuera de lo consciente (Bosson et al. 2019).

Existen trabajos que sugieren que las infancias adquieren comportamientos y actitudes a través de la observación de modelos simbólicos, como los reproducidos en medios populares (Bandura et al. 1963). La televisión y el cine, por ejemplo, tienen un papel importante en la perpetuación de estereotipos de género (Galvez, Tiffenberg y Altszyler 2019), ya que la forma en la que se representan las personas en los mismos puede influir en la percepción y las expectativas de la audiencia. En esta misma línea, los patrones culturales presentes en la producción lingüística escrita pueden fomentar desbalances de género (Lewis, M., & Lupyán, G., 2020). Los investigadores plantean que las creencias adquiridas en la niñez a partir de producciones lingüísticas propician menor representación de mujeres en campos de ciencia, tecnología, ingeniería y matemática (STEM). Postulan que los idiomas con asociaciones de género más fuertes tienden a reproducir relaciones más fuertes entre hombres y carrera y mujeres y familia. Esto a partir de una correlación positiva entre la fortaleza de las asociaciones de género en el lenguaje y las respuestas al Test de Asociación Implícita ( $r(25) = 0.48$  [0.12, 0.73],  $P = 0.01$ ). El problema de los estereotipos de género en los contenidos audiovisuales y en el lenguaje de estos constituye entonces un problema que se propaga en la sociedad reproduciendo y naturalizando roles de género desiguales.

En un estudio que analizó películas familiares, series del prime-time y series infantiles se encontró que los personajes femeninos tienden a estar menos asociados con el trabajo en comparación con los masculinos, por ejemplo, el porcentaje de personajes C-Level masculinos en películas familiares era del 97% mientras que la representación femenina era únicamente el 3% restante (Smith et al. 2012). Este estudio también demuestra que las figuras femeninas tienen menos probabilidad de aparecer en pantalla y, cuando lo hacen, a menudo reproducen modelos sociales estereotipados. Por lo general, relacionados principalmente con la apariencia física y que, en consecuencia, tienen potencialmente implicancias negativas en las expectativas de la imagen en mujeres. Grabe, Ward y Hyde (2008) en un meta-análisis de 77 estudios encuentran que existe una correlación entre la exposición a medios de

comunicación con cuerpos idealizados y la insatisfacción corporal, la internalización del cuerpo ideal delgado y comportamientos y creencias alimentarias ( $ds = -.28, -.39, \text{ and } -.30$ , respectivamente).

En la misma línea, en un análisis comparativo entre libros para adultos y libros infantiles se encontró una presencia más marcada de estereotipos de género en los segundos (Lewis et al. 2022). Estos contenidos asociaban, en su mayoría, a las mujeres con los roles familiares, a diferencia de los hombres quienes se encontraban más vinculados a carreras profesionales, reforzando así creencias relacionadas a los roles en la sociedad desde temprana edad.

Si bien se ha avanzado en cuanto a paridad de derechos en el mundo occidental, las representaciones en los medios audiovisuales reflejan aún representaciones de los géneros muy diferenciadas y con impactos en diferentes áreas de la vida, más allá de lo estrictamente identitario. Por ejemplo, la tendencia de asociación entre los pronombres masculinos y las capacidades cognitivas en películas se ha mantenido de manera constante en los últimos 50 años (Galvez, Tiffenberg y Altszyler 2019). Aunque se ha observado un aumento en la representación de las mujeres en los campos científicos del 24% entre 2016 y 2023 (STEM Women, 2023), los niños todavía tienen más probabilidad de dibujar científicos hombres que mujeres, particularmente, el estudio de Miller, et. al (2018) muestra que mujeres de 16 años dibujaron más científicos hombres en un ratio promedio de tres a uno. Casey, Novick y Lourenco, en un estudio realizado sobre libros infantiles en 2021, llegan a una conclusión similar respecto a las asociaciones de género en los niños. Si bien la proporción de personajes centrales masculinos disminuyó entre 1960 y 2020, lo que sugiere un movimiento hacia la paridad de género, aún hay una subrepresentación de protagonistas femeninas en la literatura infantil.

En conclusión, los contenidos consumidos por niños y el lenguaje de estos reproducen estereotipos y, a pesar de que han existido avances hacia la equidad de género y la eliminación de los mismos, estos continúan influenciando las narrativas y las representaciones en los medios de comunicación dirigidos a niños. Por ende, se mantienen como una fuente de perpetuación de histórica desigualdad.

### **1.1.2. Estereotipos y prejuicios en YouTube**

YouTube se ha consolidado como una de las plataformas de contenido audiovisual más influyentes y populares del mundo contemporáneo. Con aproximadamente 2.491 millones de usuarios activos según Statista (Dixon, 2023), esta red social ocupa el segundo lugar en términos de popularidad luego de Facebook. Además de disponer una amplia variedad de contenidos, esta plataforma desempeña un rol importante en la formación de percepciones y actitudes sociales.

Un estudio cuantitativo realizado sobre la representación de género en esta plataforma ha comprobado patrones de subrepresentación de producciones realizadas por mujeres (Döring y Mohseni, 2018). Esta investigación sugiere además que aquellas que se ajustan a las expectativas de género establecidas por la sociedad tienden a recibir menos comentarios negativos, en comparación con las que no lo hacen. Ahora bien, esto no significa que la experiencia de las mujeres que cumplen con las normas de género sea necesariamente positiva.

En un análisis posterior realizado por los mismos autores, se confirmó la persistencia de disparidades de género en los comentarios recibidos por las creadoras de contenido (Döring y Mohseni, 2020). Se



encontró que las mujeres recibían menos críticas positivas sobre su personalidad y el contenido de sus videos que los hombres, mientras que eran más propensas a recibirlas centradas en su apariencia física. Esto indica que, a pesar de que la plataforma es abierta y permite la participación de muchos sectores y disidencias, no está exenta de reproducir estereotipos de género y creencias nocivas para el desarrollo personal.

### **1.1.3. Análisis de sesgos y estereotipos con métodos de data science.**

El estudio de estereotipos de género en los medios populares utilizando técnicas de ciencia de datos abarca una amplia gama de enfoques y metodologías. Algunos investigadores utilizan técnicas del procesamiento del lenguaje natural (NLP por sus siglas en inglés) para analizar el contenido textual y extraer patrones. Otros emplean modelos predictivos de machine learning para identificar tendencias y cuantificarlas en los documentos. Estas metodologías permiten identificar y comprender los estereotipos de género en grandes conjuntos de texto. A continuación se describirán algunas que posteriormente se aplicarán en el presente trabajo.

En su trabajo, Galvez, Tiffenber y Altszyler (2019) emplearon matrices de co-ocurrencias y la información mutua puntual (Positive Pointwise Mutual Information, PPMI) para investigar las relaciones entre los pronombres de género y palabras que denotan habilidades cognitivas de alto nivel, como por ejemplo inteligente o genio. Esta técnica permite cuantificar la frecuencia con la que dos términos aparecen juntos en un corpus en relación a cuánto se esperaría que eso sucediera por simple azar (Jurafsky y Martin 2023). El análisis se llevó a cabo a partir de subtítulos de películas occidentales producidas entre 1967 y 2016. Para evaluar la significancia estadística de los resultados obtenidos, los investigadores utilizaron matrices de contingencia y ratios de probabilidad (Haddock et al. 1998; Olivier y Bell 2013). Los resultados confirmaron una asociación desigual de las palabras que denotaban habilidades cognitivas con pronombres de diferentes géneros. Esto se considera entonces como un indicador de la presencia de estereotipos en el contenido de las películas analizadas, lo que sugiere que las asociaciones entre los pronombres y las habilidades cognitivas reflejan y refuerzan las creencias culturales predominantes en la sociedad.

Por otro lado, el análisis de estereotipos utilizando embeddings es altamente popular en la actualidad. Estos son representaciones vectoriales de palabras en un espacio matemático que condensa información originalmente dispersa sobre el comportamiento de cada palabra en los textos, registrada como sus ocurrencias con otras palabras. Estas representaciones densas mejoran el rendimiento de las aplicaciones de machine learning, ya que están mejor conectadas con relaciones semánticas subyacentes (Mikolov, et al., 2013). Algunos trabajos han mostrado que los vectores aritméticos derivados de los embeddings pueden utilizarse para medir asociaciones culturales presentes en los datos lingüísticos (Bolukbasi et al, 2016). Ahora bien, los embeddings también pueden reproducir prejuicios sociales presentes en los datasets de entrenamiento, lo que puede llevar a sesgos en las aplicaciones finales del modelo (Bolukbasi et al, 2016). En este sentido, es crucial realizar evaluaciones exhaustivas de los modelos y considerar estrategias para mitigar este problema antes de su implementación en aplicaciones.

En su trabajo, Boutyline, Arseniev-Koehler y Cornell (2023), exploran los embeddings obtenidos de un corpus de medios impresos para estudiar cómo han evolucionado los arquetipos relacionados con la educación a medida que las mujeres han alcanzado niveles educativos equiparables a los de los hombres. Entrenando un modelo Word2Vec los investigadores analizan la asociación entre palabras que representan inteligencia y palabras que representan a los géneros para cada embedding generado. Los hallazgos revelan que los estereotipos asociados a la inteligencia experimentaron una polarización con el paso del tiempo. A medida que las mujeres alcanzaron un mayor nivel educativo, se observa una divergencia en la percepción de la inteligencia entre los géneros en los medios impresos analizados. Este estudio sugiere la persistencia de estereotipos arraigados en la sociedad, incluso en contextos con mayor igualdad educativa entre géneros.

Por último, Lewis y Lupyan (2020) utilizan embeddings para medir creencias culturales asociadas a carreras profesionales en 25 idiomas. Utilizan la distancia coseno promedio para estimar un puntaje de género entre “anclas” asociadas a ellos y las listas de palabras definidas. Los resultados comprueban una tendencia a asociar a los hombres con carreras profesionales y a las mujeres con la familia.

## 1.2. Problema

Desde la infancia, las personas están expuestas a una amplia variedad de contenidos (audiovisuales, literarios, impresos, entre otros) que desempeñan un rol crucial en la formación de sus percepciones y actitudes hacia el género. Los estereotipos presentes en estos pueden moldear las creencias y comportamientos desde una temprana edad, influyendo en cómo las personas perciben sus propias identidades y las de los demás.

En este sentido, plataformas como YouTube, con su inmenso alcance y popularidad, se convierten en una ventana a la cultura contemporánea y a las dinámicas sociales que la impulsan. El contenido generado y consumido en la plataforma no solo refleja las tendencias culturales existentes, sino que también contribuye a dar forma a las percepciones y valores de la sociedad en general, especialmente entre las generaciones más jóvenes que son las principales consumidoras de este tipo de contenido y suelen ser las más permeables.

La presencia de estereotipos de género en los medios de comunicación, incluido el mencionado en el párrafo anterior, plantea preocupaciones importantes en términos de equidad y representación. Los estudios consultados han mostrado cómo las líneas discursivas pueden reforzar y perpetuar creencias erróneas asociadas al género, limitando las oportunidades y roles disponibles para mujeres y hombres y en consecuencia prolongando disparidades en diversos ámbitos de la vida.

Las metodologías de ciencia de datos se constituyen como una herramienta fundamental en la identificación y análisis de sesgos, entendiendo a estos como el comportamiento no homogéneo de un sistema computacional con respecto a una variable de interés. Estos procedimientos permiten examinar grandes cantidades de datos para identificar patrones y tendencias en los mismos. Al comprender mejor la naturaleza y el alcance de estos estereotipos se pueden tomar medidas para abordarlos y promover una representación más equitativa y diversa en los medios. El análisis de los

contenidos de YouTube desde una perspectiva de género es fundamental para fomentar una cultura más inclusiva y representativa, promoviendo la igualdad y contribuyendo a una sociedad más equitativa.

### 1.3. Objetivo

Partiendo de la premisa de que los estereotipos de género refuerzan y perpetúan nociones y comportamientos sociales dañinos y discriminatorios, la detección de los mismos en los contenidos audiovisuales diseñados para niños puede proporcionar una información valiosa para el desarrollo de estrategias encaminadas a la eliminación o mitigación de los mismos, así como para la creación de contenido más inclusivo y equitativo.

El objetivo de este estudio es realizar un análisis sobre la presencia de estereotipos de género en los videos destinados al público infantil en YouTube en idioma inglés. Para ello, se emplearán los subtítulos como fuente primaria de datos y se aplicarán técnicas de procesamiento del lenguaje natural (NLP) con el objetivo de identificar patrones y tendencias en el lenguaje utilizado en estos videos. Se buscará responder a las preguntas ¿Son estos videos libres de estereotipos de género? ¿Se encuentran en ellos patrones y tendencias dispares en cuanto a género?

Para responder estas preguntas, se analizarán las posibles asociaciones entre palabras que representan a un género y palabras que representan características ampliamente difundidas de lo que socialmente “debería” corresponderse a un género o a otro. Estas asociaciones se basan en un determinado contexto cultural y pueden no aplicarse a otros. La lista de palabras seleccionada para este estudio surge de una decisión subjetiva pero se apoya en la literatura existente sobre la materia que fue consultada durante la investigación. Por ejemplo, el estudio realizado por Smith et. al (2012) en donde se detectó que los personajes femeninos de películas familiares estaban menos asociados al trabajo que los masculinos. También fueron consideradas las tendencias observadas en el corpus. Se destaca que la lista de palabras no está exenta de ser reduccionista o poco generalizable y será uno de los aspectos a priorizar en trabajo futuro.

Es importante considerar que este estudio trabajará sólo con dos géneros, masculino y femenino, y la lista de palabras que los representan se centrará en parte en dos pronombres de género: he y she. Si bien existen otros términos y expresiones que representan a estos géneros, se ha decidido limitar el alcance para garantizar la precisión de los resultados y hacer factible la complejidad computacional. En algunas metodologías aplicadas se usará una lista de palabras asociadas al género extendida.

En resumen, este trabajo representa un esfuerzo por comprender y abordar los estereotipos presentes en los contenidos audiovisuales dirigidos a niños en YouTube, con el objetivo de contribuir a la creación de un entorno digital más inclusivo e idealmente libre de prejuicios. Se diferencia de trabajos previos en la materia al utilizar como input el lenguaje de los contenidos audiovisuales de YouTube. Esto dado que si bien existen estudios sobre contenidos masivos, como libros o películas, hasta donde tengo conocimiento no se han realizado acerca de videos de YouTube empleando las herramientas metodológicas que se utilizarán en esta tesis (el Positive Pointwise Mutual Information o Embeddings por ejemplo).

Por último, los resultados de esta tesis son aplicables en diversos ámbitos empresariales. En primer lugar, en el desarrollo de contenidos audiovisuales, los hallazgos del estudio pueden utilizarse para lograr una representación más equitativa, promoviendo la igualdad de género y atrayendo a un público más amplio y diverso. En segundo lugar, la evaluación continua de las asociaciones de género y la presencia de estereotipos, permitiría a las empresas demostrar su compromiso con la igualdad de género, mejorando su reputación y reforzando su responsabilidad social corporativa. En tercer lugar, las compañías dedicadas a la educación y capacitación podrían desarrollar programas y recursos educativos basados en los insights de este trabajo. Por último, la metodología aplicada en este estudio es útil para empresas de consultoría y auditoría al ofrecer servicios especializados a otras organizaciones interesadas en entender y mitigar los sesgos de género en sus propios contenidos y comunicaciones.

## 2. Datos

### 2.1. Estructura de los datos

Los estereotipos de género se reproducen de múltiples formas en la sociedad. Desde actitudes hasta oportunidades disponibles. Una de las áreas donde pueden estar presentes, ser identificados y analizados es en el lenguaje. Los corpus de texto brindan una oportunidad para examinar la presencia de estos en el discurso.

En el contexto de contenidos audiovisuales, los subtítulos proporcionan una transcripción del lenguaje oral. En línea con esto, los subtítulos de los videos de YouTube para niños se constituirán como el corpus principal para este estudio. Para ello se utilizarán archivos de formato de texto (.txt) que contienen los diálogos y narrativas presentes en los videos. Cada archivo corresponderá a un video específico cuya identificación estará presente en el título del archivo.

### 2.2. Obtención del dataset

El corpus utilizado en este estudio se generó utilizando herramientas que permiten obtener información de los videos y descargar los subtítulos asociados a cada uno de ellos. En primer lugar, se utilizó la API de YouTube para recopilar los identificadores (IDs) de videos de 40 canales para niños. La selección de estos se basó en el ranking publicado por Social Blade *“Top 100 YouTubers made-for-kids channels sorted by sb rank”*. Este índice proporciona información sobre los contenidos para niños con el mayor número de suscriptores, asegurando así que los videos analizados sean representativos de los que los niños ven actualmente en la plataforma. El [apéndice 1](#) contiene la lista de los canales seleccionados junto con el número de suscriptores. Las categorías abarcan contenidos educativos, de entretenimiento y musicales. En base a estos canales se consultó a la API de Youtube para que devuelva los IDs de los videos alojados en los mismos. Cada consulta permitió obtener un máximo de 50 IDs por lo que se repitió este proceso 3 veces para cada canal y luego se filtraron duplicados. La

duración promedio de los videos cuyo ID se obtuvo es de 7 minutos. No se seteo un número de IDs a descargar sino que este volumen responde a los recursos disponibles para este trabajo.

Una vez obtenidos los IDs se procedió a obtener los subtítulos de cada uno de ellos. Esto se realizó utilizando la librería *YouTubeTranscript.Api*, que permite acceder a las transcripciones asociadas a un video a partir de su ID. Por último, estos textos fueron almacenados en archivos.

Se iteró sobre este proceso repetidas veces para obtener un número suficiente de IDs de videos y, en consecuencia, generar un conjunto amplio de archivos de texto. En total, se obtuvieron 3.855 txts que representan los subtítulos de los videos seleccionados.

El promedio de archivos obtenidos por canal de YouTube es de 96, siendo el rango entre 80 y 120 el volumen más frecuente de textos por canal ([apéndice 2](#)).

### 2.3. Curación y descripción de los datos

La plataforma elegida para el desarrollo del código fue Google Collaboratory. Por ello, como primer paso, se cargaron los archivos de texto a la misma.

En segundo lugar, se llevó a cabo la tokenización de todos los documentos dividiéndolos en unidades más pequeñas (palabras, símbolos o números). Para esto, se utilizó la función *word\_tokenize* de la librería NLTK. Durante este proceso, también se generó un diccionario que contiene el ID de cada documento, los tokens correspondientes y el recuento de ocurrencias de cada token dentro del documento.

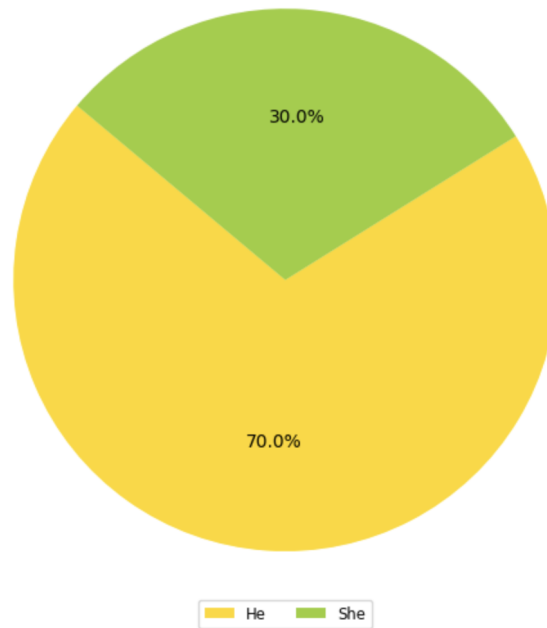
En tercer lugar, se trabajó en la limpieza y curación del corpus. Para ello se eliminaron los stopwords y las contracciones aplicando el paquete correspondiente de NLTK. Se adaptó la lista de stopwords de la librería para eliminar de ella palabras relacionadas con el género (ejemplo, *he* y *she*). Así también, se eliminaron tokens que consistían en una sola palabra, por ejemplo *b* o *z*.

Posteriormente, se llevó a cabo un proceso de curación adicional considerando las particularidades del dataset. Dado que los datos se obtuvieron a partir de subtítulos de YouTube, el conjunto de tokens incluía palabras mal escritas, las cuales fueron removidas. Por otro lado, para evitar un impacto en los análisis posteriores se eliminaron los términos que representan el 0.01% de mayor frecuencia. En este proceso se revisó no eliminar términos claves como *he* o *she*.

Después de este procesamiento, el volumen total de tokens en todo el corpus resultó en 1,446,526 de los cuales 343,566 son únicos. Esto significa que, en promedio, cada documento contiene alrededor de 375 y aproximadamente 89 son únicos ([apéndice 3](#)).

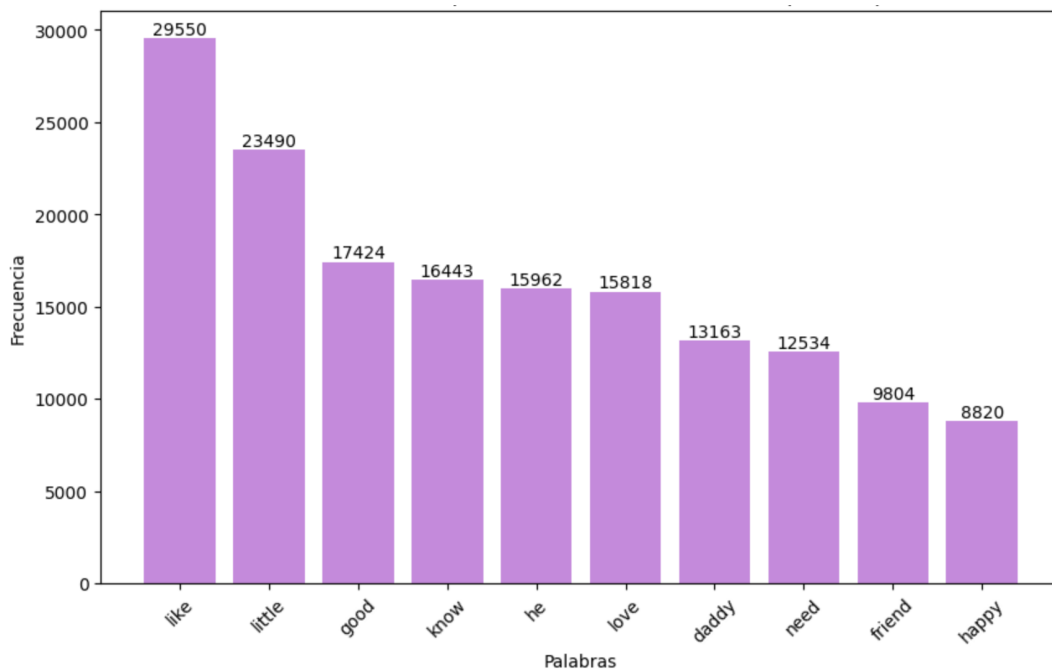
En base a la decisión de limitar parte del presente trabajo a los pronombres de género *he* y *she*, se contabilizaron las ocurrencias de ambos en todos los documentos del corpus. El total de ocurrencias de *he* es de 15,962 mientras que el de *she* es de 7,002. En principio ya se observa una sobre-representación del género masculino con un 128% más que el femenino. En términos de proporción sobre el total de tokens estos números representan el 4% para *he* y el 2% *she*.

Figura 1. Proporción de ocurrencias de "he" y "she".



Para analizar las palabras frecuentes se generaron histogramas de frecuencias. En todo el documento los 10 tokens únicos que más se repiten pueden visualizarse en la figura 2. Se destaca que *he* es el quinto término más frecuente del dataset curado.

Figura 2. Frecuencia de las 10 palabras más comunes en el corpus completo



Así también, se generaron los histogramas para cada uno de los documentos pero considerando solo los primeros 3 tokens únicos más frecuentes ([apéndice 4](#)).

### 3. Metodología

#### 3.1. Lista de palabras

El enfoque metodológico de este trabajo se centra en la selección y validación de una lista de palabras claves para el análisis de sesgos de género en el corpus. Dado que las mismas, utilizadas en los métodos de identificación de estereotipos, delimitan el espacio a explorar pueden influir significativamente en los resultados y potencialmente descubrir asociaciones culturales y cognitivas. Es crucial entonces establecer un proceso riguroso para su constitución y validación (Antoniak, M., y Mimno, D. 2021).

Existen múltiples métodos para la selección de la lista de términos, cada uno con sus ventajas y desventajas, pero ninguno está exento de reproducir sesgos. En consecuencia, todos estos procedimientos deben ser validados para asegurar su fiabilidad y relevancia para el análisis. En este contexto, resulta fundamental establecer los parámetros y el proceso de testeo utilizados para la constitución de la lista de palabras clave de este trabajo. Esto buscando garantizar la robustez y la validez de los resultados.

La metodología seleccionada para la obtención de la lista en este estudio es la curación a partir del corpus. Este enfoque implica analizar los tokens presentes en el conjunto de datos y seleccionar aquellos que sean más relevantes y representativos para los objetivos de la investigación. En base a ello, se determinaron las siguientes:

- Palabras que representan emociones, sentimientos y cuidado: *little, poor, sad, hurt, cry, cried, scare, feeling, violent, irrational, kind, rescue, hero, clean y healthy*.
- Palabras que representan el trabajo y lo doméstico: *job, responsibility, working, worked, professional, chief, school, teacher, teach, doctor, medicine, surgeon, healer, nurse, football, police, architecture, engineer, technician, technologist, science, house, vacuum, cooking, cooked y babysit*.

Otras formas flexivas de estas palabras fueron descartadas por su baja frecuencia en el corpus.

Como se mencionó previamente, es importante destacar que la selección de estas listas no está desvinculada del contexto cultural y demográfico en el que se inscriben. Igualmente, para mitigar el impacto de estos factores, se llevaron a cabo procesos de testeo y validación de las listas, siguiendo las recomendaciones de Antoniak, M., y Mimno, D. (2021):

1. Definiciones reductivas: Los términos incluidos en las listas de palabras pueden tener connotaciones reductivas que reflejan estereotipos de género arraigados en la sociedad. Para contrarrestar este problema, se llevó a cabo una revisión de la literatura en sociología y psicología, consultando múltiples estudios que abordan la asociación de ciertas profesiones y

habilidades con uno y otro género. Por ejemplo, investigaciones como la de Lewis et al. (2022) han mostrado que en libros para niños, los personajes femeninos suelen estar más asociados con roles familiares mientras que los masculinos se relacionaban más con carreras profesionales. Algo similar se evidencia en los resultados del trabajo de Galvez, Tiffenberg y Altszyler (2019) a partir de los cuales se descubrió la presencia de estereotipos de género relacionados a las habilidades cognitivas y los hombres en películas orientadas a niños. Así también, otros estudios, como el de Döring y Mohseni (2018), demuestran cómo los contenidos producidos por mujeres o con protagonistas femeninas suelen ser más propensos a recibir comentarios vinculados con la apariencia física de quienes lo generan o lo personalizan.

En este trabajo se realizan suposiciones similares: En primer lugar, que los hombres suelen estar asociados con más fuerza a carreras profesionales y por ende se seleccionaron las palabras *job, responsibility, working, worked, professional, chief, doctor, medicine, surgeon, architecture, engineer, technician, technologist* y *science*. En segundo lugar, que en contraposición, las mujeres ocupan un espacio de cuidado de otros y del hogar conectado con habilidades emocionales y sentimientos positivos. Para indagar sobre esto se seleccionaron palabras como *house, vacuum, cooking, cooked, babysit, sad, hurt, cry, cried, scare* y *feeling*. Por último, que las personas del género femenino suelen enfrentar expectativas más exigentes en cuanto a su salud y estándares de belleza en comparación con los hombres. Por ello se incluyeron las palabras *healthy* y *clean* en la investigación.

2. Definiciones imprecisas: Se trabajó en la definición clara de las dos ramas analizadas, emocionalidad, sentimientos y cuidado por un lado y trabajo profesional y doméstico por el otro. Esto implicó establecer criterios precisos y específicos para cada categoría, a fin de evitar ambigüedades y garantizar una interpretación coherente de los resultados obtenidos.

Tabla 1. Definiciones de los conceptos incluidos en las listas de palabras.

| Rama                                  | Definición   |
|---------------------------------------|--|
| Emocionalidad, sentimientos y cuidado | Esta rama busca entender las asociaciones de ambos géneros con habilidades emocionales y empáticas. Los sentimientos refieren a un estado de ánimo. La emocionalidad se entiende como la capacidad de percibir, comprender y expresar los sentimientos. Por último, el cuidado hace referencia a las acciones y actitudes dirigidas al bienestar propio y de los demás.                          |
| Trabajo profesional y doméstico       | Esta rama busca analizar las asociaciones de ambos géneros con habilidades y carreras profesionales. Se entiende al trabajo profesional como las actividades laborales que requieren un alto nivel de habilidad, conocimiento especializado y responsabilidad.<br><br>Por otro lado, incluye nociones del trabajo doméstico entendido como las actividades de mantenimiento y cuidado del hogar. |



3. Factores léxicos: La frecuencia y la categoría gramatical de los términos incluidos en las listas son factores léxicos que pueden influir en las mediciones de sesgos de género. Los investigadores sugieren que estos impactos son especialmente significativos en palabras poco comunes o raras. Para abordar esta consideración, se llevó a cabo un análisis de las ocurrencias de cada una con el objetivo de comprender mejor su distribución y su uso en el corpus de datos (apéndice 5).

En base a esta revisión se confirma la lista de palabras previamente definida y se acota su significado, limitando así el espacio semántico a inspeccionar en esta tesis.

### 3.1.1. EDIA (Estereotipos y discriminación en Inteligencia Artificial)

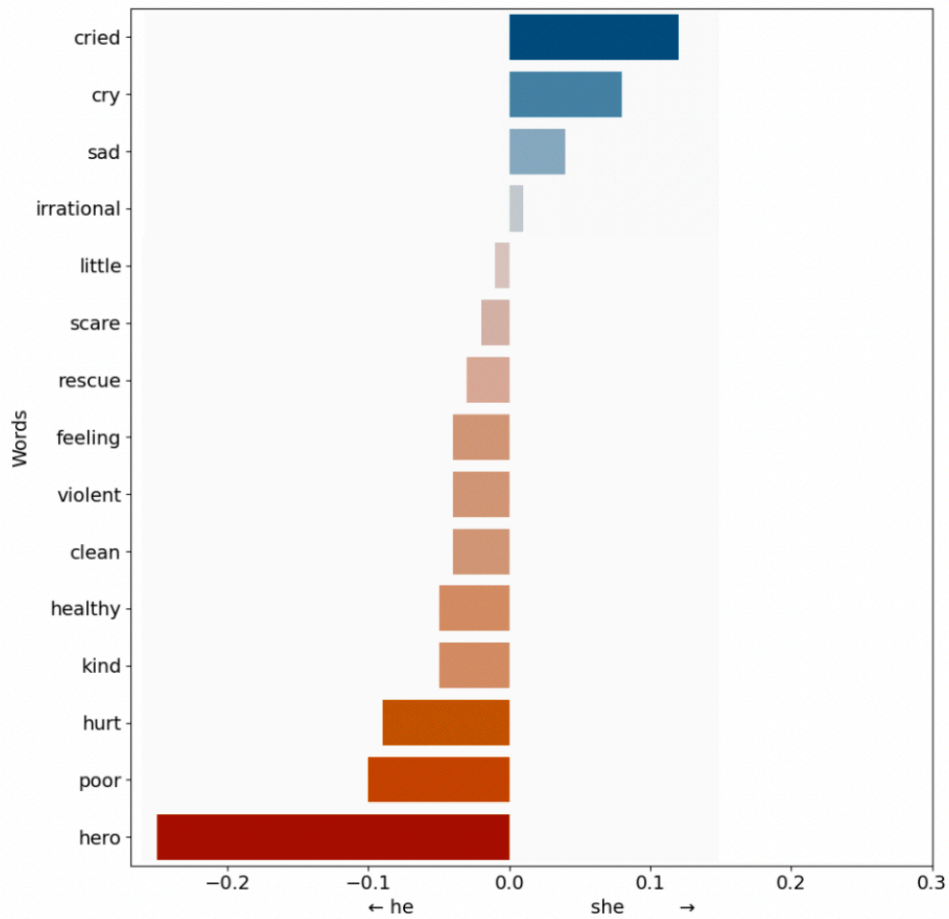
Como resulta evidente de lo expuesto hasta ahora, las listas de palabras son una componente determinante de cualquier estudio de sesgos basado en palabras. Por esa razón realizamos un estudio preliminar para validar la adecuación de las listas de palabras propuestas para representar los conceptos objetivo.

El uso de herramientas como EDIA (Estereotipos y discriminación en Inteligencia Artificial, Alonso Alemany, et. al, 2023) proporciona una manera efectiva de explorar los estereotipos presentes en los embeddings y modelos de lenguaje. Esto contribuye significativamente a la comprensión de las relaciones semánticas y contextuales en corpus de texto. El modelo calcula similitudes entre palabras a partir de la distancia euclídea en un espacio de embedding neuronal donde cada palabra está representada por su valor en 300 dimensiones. El embedding utilizado es FastText entrenado con el corpus Spanish Billion Word Corpus compuesto de 1.4 billones de palabras descrito en <https://github.com/dccuchile/spanish-word-embeddings>.

En este estudio, se aprovechará la funcionalidad de exploración de sesgos disponible en [Edia Full En - a Hugging Face Space by vialibre](#) para validar las suposiciones y decisiones realizadas en la selección de la lista de palabras. La herramienta permite observar listas de términos en un espacio bidimensional utilizando una proyección PCA (análisis de componentes principales), facilitando así la evaluación de la proximidad entre los mismos. Esta cercanía se determina a partir de los contextos de ocurrencia en los datos de entrenamiento del modelo, brindando información sobre las relaciones en función de su uso en el lenguaje. Con el objetivo de validar las asunciones realizadas en la selección de la lista de palabras se utilizó EDIA para visualizar cada palabra y entender su relación con los respectivos pronombres de género, he y she.

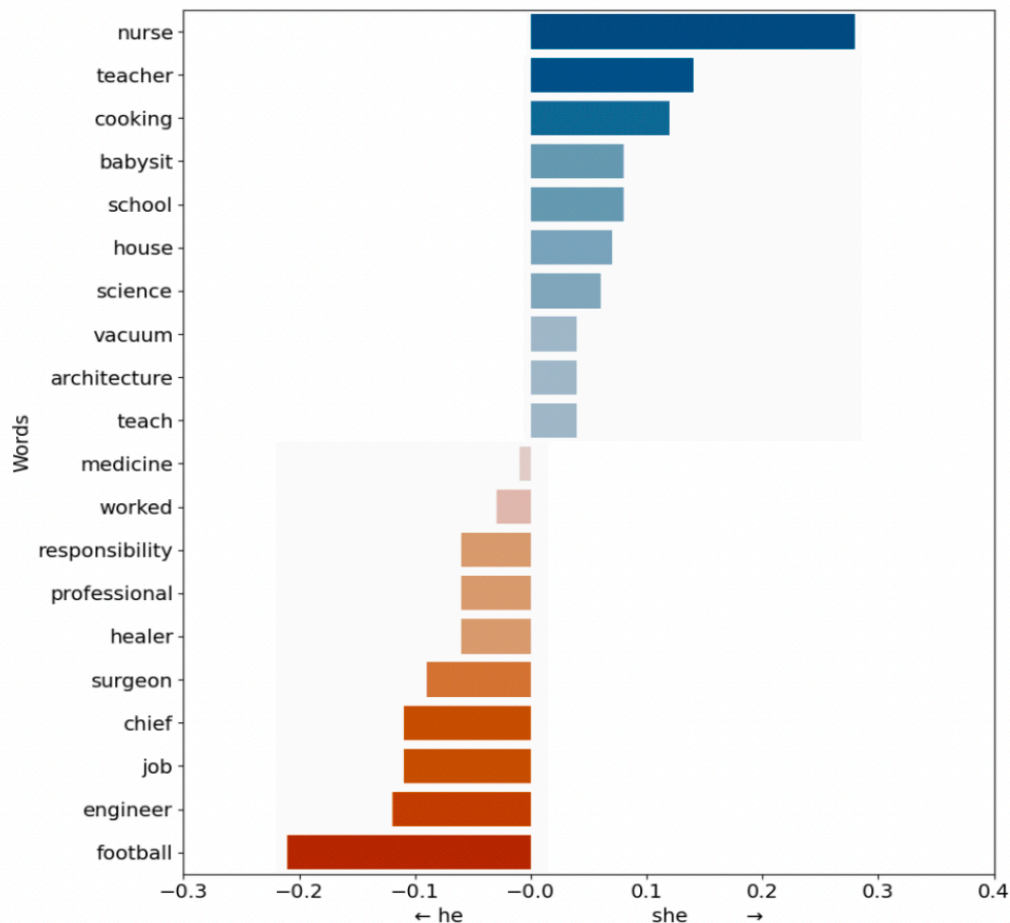
Por un lado, en el conjunto asociado a las emociones, sentimientos y cuidado es notable la distancia en algunos términos de la lista. Por ejemplo, *cried* muestra una mayor proximidad con *she* que con *he*. Al contrario, *hero* tiene una mayor cercanía mucho más fuerte con *he*. La herramienta permite descubrir que para las palabras de esta categoría hay diferencias de asociaciones de género.

Figura 3. Visualización EDIA de palabras asociadas a las emociones, sentimientos y el cuidado.



Por otro lado, en el conjunto vinculado al trabajo y a lo doméstico también hay algunas diferencias interesantes en las proximidades. Por ejemplo, *nurse* y *teacher* se visualizan como más cercanas a *she*. En el otro extremo, *football* y *engineer* parecen ser más asociadas a *he*. Al igual que con el grupo de palabras anterior, la visualización confirma diferencias de asociación entre los términos lo que puede ser un indicador de estereotipos de género.

Figura 4. Visualización EDIA de palabras asociadas al trabajo y a lo doméstico.



En conclusión, según los resultados obtenidos en la exploración con la herramienta EDIA, las listas de palabras seleccionadas se presentan como válidas para el análisis de sesgos y estereotipos. Esto debido a que en la visualización bidimensional se observan, para muchas de ellas, asociaciones más fuertes con uno u otro género. A partir ello se da lugar a la siguiente hipótesis con la que se trabajará posteriormente: “Los términos de las listas definidas tienen una connotación de género que se puede comprobar a partir de la cercanía o la fuerza de asociación de cada una de ellas con los pronombres de género”. Para el futuro análisis se desagregará esta hipótesis de la siguiente manera:

1. En las palabras que representan a las emociones, sentimientos y cuidado la asociación es mayor a:
  - a. he: *little, poor, healthy, rescue, feeling, kind, clean, hurt, scare, hero y violent.*
  - b. she: *sad, cry, cried e irrational.*
2. En las palabras que representan al trabajo y a lo doméstico la asociación es mayor a:
  - a. he: *job, working, doctor, medicine, healer, worked, professional, chief, football, engineer, surgeon y police.*

- b. *she: school, teacher, house, vacuum, cooking, cooked, responsibility, teach, nurse, babysit, architecture, technician, science y technologist.*

### 3.2. Matriz de co-ocurrencias

Considerando todo el corpus se generó una matriz de co-ocurrencias de ventana 5. Estas matrices permiten entender las relaciones entre palabras al asociar términos con términos contabilizando la ocurrencia conjunta de ambos en todos los documentos de acuerdo a la ventana definida. De esta manera, se puede ver el contexto para cada palabra y permiten analizar qué términos ocurren con más frecuencia con cada uno de los pronombres de género.

La elección de una ventana de tamaño 5 implica que se están considerando las palabras que aparecen hasta 5 palabras antes o después de cada término *x*. La elección de la amplitud define el contexto y depende del objetivo del análisis y de la naturaleza del corpus. Siguiendo lo expuesto por Church y Hanks (1990) ventanas más chicas captarán asociaciones fijas mientras que ventanas más amplias pueden capturar relaciones léxico-semánticas y otras relaciones a mayor escala.

Una vez generada la matriz de co-ocurrencias, se pueden realizar diferentes análisis para comprender mejor las relaciones entre las palabras, por ejemplo calcular el Positive Pointwise Mutual Information que se explicará posteriormente. Una forma de visualización rápida para entender qué términos están cercanos a los pronombres de género y sus frecuencias en el corpus son las nubes de palabras. Estas visualizaciones pueden ayudar a identificar tendencias en el uso del lenguaje a partir de la simple observación. Utilizando entonces la matriz de co-ocurrencias como input y los pronombres “*he*” y “*she*” como palabras objetivo se construyeron nubes que mostraron las palabras más frecuentes que co-ocurrían con “*he*” o “*she*” respectivamente.

El proceso para obtener las visualizaciones se puede resumir en dos etapas. En primer lugar, con la función *get()* se crea un diccionario con los tokens que co-ocurren con la palabra objetivo y sus frecuencias. En segundo lugar, a partir de la biblioteca *WordCloud* se genera la nube de palabras considerando las frecuencias contenidas en el diccionario.

Figura 5. Nube de las palabras más frecuentes que co-ocurren con el pronombre "he"



Figura 6. Nube de las palabras más frecuentes que co-ocurren con el pronombre "she"



### 3.3. Positive Pointwise Mutual Information (PPMI)

El Positive Pointwise Mutual Information (PPMI) es una métrica que proporciona información sobre la asociación entre términos en un corpus de texto. Al calcular el PPMI para cada una de las palabras definidas en las listas y ambos pronombres de género se puede obtener una medida cuantitativa de la vinculación entre ellas.

La métrica compara la probabilidad de que dos palabras ocurran juntas en un contexto dado, con la probabilidad esperada de que ocurran independientemente una de la otra, es decir por azar (Church y Hanks, 1990). Si el valor obtenido de PPMI para dos términos es mayor a cero se puede asumir que existe un vínculo entre ambas más allá de lo que se esperaría por casualidad. Un PPMI más alto sugiere una conexión más fuerte entre las palabras.

Este indicador se construye a partir del PMI (Pointwise-Mutual-Information) que se obtiene dividiendo la probabilidad conjunta de una palabra 1 y una palabra 2 en el corpus, es decir la probabilidad de observarlas juntas, por la probabilidad de cada una de ellas independientemente. Si realmente existe una asociación entre los términos, se espera que el numerador sea mayor al denominador y por ende esta operación devuelva un valor positivo. A este resultado se le aplica el logaritmo en base 2. Ahora bien, el PMI puede arrojar resultados negativos, los cuales según Jurafsky y Martin (2023) son poco confiables e interpretables. Para solucionar este inconveniente, se define el PPMI como el valor mayor entre el PMI y cero.

$$\text{PPMI}(\text{word}_1, \text{word}_2) = \max\left(\log_2 \frac{P(\text{word}_1, \text{word}_2)}{P(\text{word}_1)P(\text{word}_2)}, 0\right)$$

Por otro lado, Church y Hanks (1990) postulan que es importante tener en cuenta que el PMI puede volverse poco estable cuando los conteos de ocurrencias son muy bajos. Para abordar este factor y garantizar la robustez de los resultados se decidió limitar el corpus a los tokens que aparecen más de 10 veces en cada documento para el cálculo del PPMI. Como resultado, el volumen total de tokens se redujo a 1,406,650 y el volumen de tokens únicos a 310,991, representando una reducción del 3% y del 9% de los valores iniciales respectivamente.

Los resultados obtenidos con esta técnica serán utilizados luego para comparar las asociaciones con cada uno de los pronombres de género y evaluar la hipótesis planteada a partir de la herramienta EDIA.

### 3.3.1. PPMI con Laplace

El ratio PPMI que se utiliza en este trabajo suele representar de forma inadecuada eventos poco frecuentes como se explicó anteriormente (Jurafsky y Martin 2023). Esto hace que palabras con bajo volumen de ocurrencias en el corpus puedan presentar un PPMI más alto de lo que corresponde a su distribución real, afectando la interpretación de las asociaciones entre palabras.

El suavizado de Laplace es una técnica comúnmente utilizada para abordar este problema de eventos poco frecuentes. La idea detrás de este método es agregar una constante a todos los conteos en la matriz de co-ocurrencias antes de calcular el PPMI. Esto debería ayudar a reducir el impacto de las palabras poco frecuentes, al aumentar ligeramente los conteos. En este sentido, se generó una nueva



matriz de co-ocurrencias aplicando el suavizado de Laplace sumando una constante de valor 3 a todos los valores. A partir de la misma, se calculó el PPMI y posteriormente se compararán los resultados con los valores obtenidos sin el suavizado. Es conocido que, en una distribución exponencial como la de las frecuencias de las palabras en lenguaje natural, el suavizado de Laplace puede reservar una masa de probabilidad exageradamente grande para los eventos poco frecuentes, es decir, para las palabras poco frecuentes (Manning y Schütze, 1999). Tendremos en cuenta esta consideración en nuestra aproximación.

### 3.3.2. PPMI en contexto

Por otro lado, el PPMI se calcula a partir de la matriz de co-ocurrencias que está construida considerando una ventana de palabras. Esta métrica mide asociaciones de acuerdo a cuanto más ocurren dos términos conjuntamente de lo que deberían ocurrir por simple azar, pero no considera si dos palabras aparecen juntas frecuentemente de manera positiva o negativa. Es decir, para el PPMI resulta indistinto si la frase es “*boys cry*” que si es “*boys don’t cry*”, lo que por supuesto modifica la interpretación de los resultados. En base a este inconveniente se plantea la siguiente pregunta: ¿Los PPMIs obtenidos en este estudio están ignorando negaciones?

Para abordar el interrogante se revisaron los contextos de ocurrencias dentro del corpus. La metodología se explica a continuación:

1. Se definió la función *find\_documents\_with\_cooccurrence* que itera sobre todo el corpus y devuelve el conteo de los documentos en donde co-ocurren el pronombre y los términos de la lista de palabras.
2. Se definió la función *check\_negations\_in\_docs* que, a partir de una lista de tokens que representan negaciones, busca las co-ocurrencias entre los pronombres y los términos de la lista de palabras y analiza si existe una negación considerando una ventana de tamaño 5. La función devuelve el conteo de documentos con negaciones para cada uno de los pares.
3. A partir de estas dos funciones se obtiene la proporción de documentos con la co-ocurrencia que presentan una negación sobre el total de documentos con la co-ocurrencia.
4. Finalmente se graficaron estos resultados para cada una de las listas de palabras y los pronombres de género considerados en este trabajo.

Este proceso se realizó para ambos pronombres y ambas listas de términos.

## 3.4. Latent-Dirichlet-Allocation (LDA)

Latent-Dirichlet-Allocation (LDA) es un modelo probabilístico que permite representar los documentos de un corpus como mezclas aleatorias de temas latentes (Blei, D. M., Ng, A. Y., & Jordan, M. I., 2003). Cada tema se caracteriza por una distribución de palabras y los textos pueden estar asociados a

múltiples temas. Esta técnica de aprendizaje no supervisado posibilita descubrir tópicos, entendidos con un grupo de términos que representan un concepto amplio (Rao, P., & Taboada, M., 2021). En este sentido, resulta interesante analizar qué temas son detectados a partir del dataset y si estos se pueden asociar a estereotipos de género.

Para este trabajo se utilizó la librería *gensim* que facilita inferir el modelo LDA. Previo a ello se construyó un *pandas data frame* conteniendo todos los subtítulos y su *id* de documento. En la misma, se realizaron las tareas de curación previamente descritas. Esto es la tokenización de los documentos y remoción de *stop words*, términos muy frecuentes y sin significado léxico.

Con el objetivo de mejorar la calidad se crearon bigramas compuestos de dos palabras que ocurren frecuentemente juntas en el corpus. Esta técnica ayuda al algoritmo a capturar frases con mayor sentido que pueden no ser visibles en términos aislados. Para esto se utilizó también el paquete de *gensim* con las funciones *phrases* y *phraser*. Los parámetros a establecer fueron dos:

- *“min\_count”*: define el recuento mínimo de frecuencia de los bigramas. Todos aquellos que ocurren con menor frecuencia a *“min\_count”* no son considerados. Esto permite enfocar el entrenamiento solo en los que sean significativos y controlar el tamaño del vocabulario, potencialmente mejorando la calidad. Se estableció en 10, es decir, no se considerarán los que ocurran menos de 10 veces en todos los documentos.
- *“threshold”*: controla el umbral de puntuación para formar bigramas, la fuerza de asociación entre dos palabras para ser consideradas como una. De la misma manera que el parámetro anterior utilizar un umbral implica que solo se consideren aquellos que estén fuertemente asociados. Se definió en 150.

Una vez definidos los bigramas se trabajó en la construcción del diccionario y el corpus necesarios para el entrenamiento del modelo. El mismo requiere como input una representación numérica compuesta de dos partes:

1. *id2word* (diccionario): mapea cada término único del dataset a un identificador.
2. *corpus*: la lista de documentos preprocesados se transforma en una matriz término-documento utilizando el método *doc2bow*. Cada documento representa un vector de frecuencia de palabras y cada elemento del vector corresponde al conteo de ocurrencias de una palabra en ese documento.

Para el entrenamiento se pueden ajustar los hiper parámetros de acuerdo a las particularidades del análisis y buscando optimizar la calidad del modelo. Se destacan los siguientes:

- *“num\_topics”*: define el número de tópicos que se van a generar. Un valor menor implica temas más amplios y generales que pueden omitir particularidades de los documentos. Por otro lado, un valor más alto produce resultados más específicos y captura mayor variabilidad, pero complejiza la interpretabilidad e incrementa el tiempo de entrenamiento.
- *“passes”*: este parámetro indica cuántas veces se debe pasar por todo el corpus en la generación de los temas. Cuanto mayor sea más puede el algoritmo ajustarse a los datos, pero incrementa el tiempo de entrenamiento y puede llegar a ocasionar *over-fitting*.



- “*iterations*”: determina el número de iteraciones a realizar ajustando los parámetros. Cada iteración actualiza los pesos de los tópicos y las distribuciones de palabras. Al igual que en los dos casos anteriores, un número mayor permite mejorar la precisión pero aumenta los tiempos de entrenamiento.

Por último, se definió la coherencia como la métrica de validación. Esta se refiere a cuán interpretables y coherentes son los temas generados. Mide cuánto se relacionan los términos de cada tópico y en este sentido, constituye una medida de calidad del modelo. Cuanto más alto sea este valor más precisos se considerarán los resultados.

### 3.4.1. Entrenamiento

Como se mencionó anteriormente, la construcción y calidad del modelo LDA depende de la elección de los valores para los hiper parámetros. Con el objetivo de descubrir qué mix de valores construyen el modelo con mayor coherencia se realizó una iteración para los tres hiper parámetros principales. El foco de este proceso fue obtener tópicos interpretables que arrojen insights sobre el corpus.

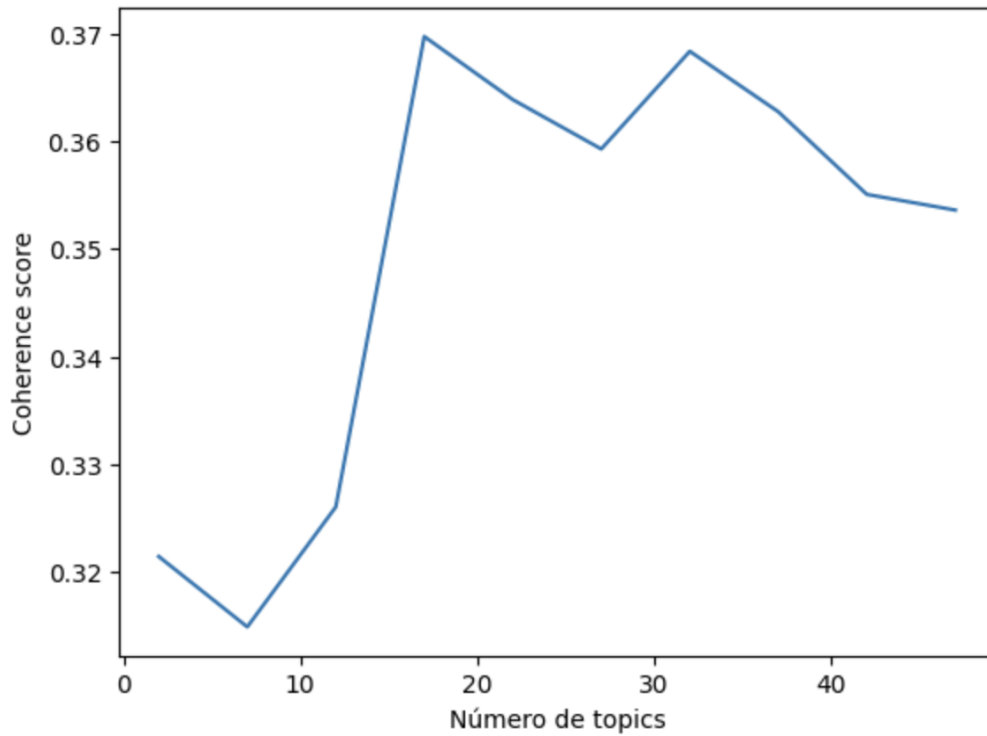
Es importante destacar, que existen limitaciones computacionales en algunos casos. Por ejemplo, un número demasiado alto de passes podría generar un modelo con un tiempo de entrenamiento elevado, imposible de terminar con la capacidad computacional disponible para este trabajo. Considerando el objetivo de maximizar la coherencia cómo así también las limitaciones computacionales se iteró sobre distintos valores para el número de temas, los passes y las iterations.

#### 3.4.1.1. Número de topics

En primer lugar, se trabajó con el número de temas óptimo para maximizar la coherencia. Como se mencionó, este determina el número de tópicos que el modelo va a entregar. Cuanto más alto más específicos serán los tópicos, lo que puede complejizar demasiado la interpretabilidad, además de incrementar el tiempo de entrenamiento.

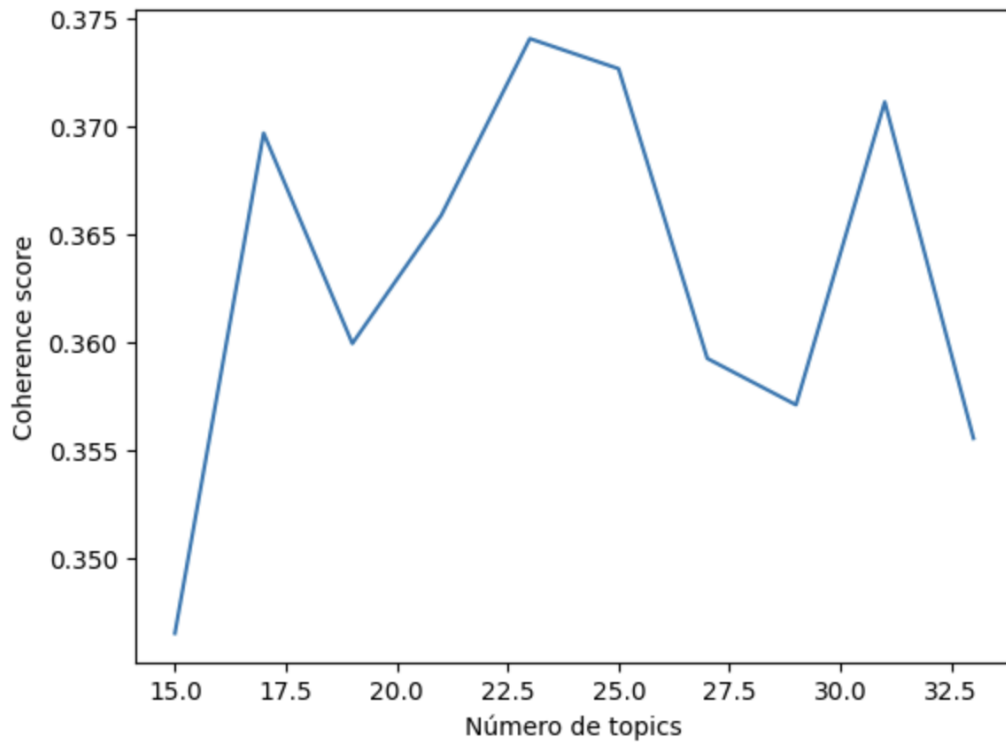
Se iteró en un rango de 2 a 50 temas, avanzando de 5 en 5. Para cada uno de estos valores se entrenó el modelo y se calculó la medida de validación. Como se observa en la figura 7, el rango con mayores valores de coherencia está entre los 15 y los 35 tópicos. Específicamente, para los valores 17 y 32 la coherencia fue muy similar, de 0.369 y 0.368 respectivamente.

Figura 7. Valores de coherencia para el rango de 2 a 50 temas.



En consecuencia, se realizó una nueva iteración entre el rango de 15 y 33 temas (figura 7), avanzando de 2 en 2 para obtener mayor visibilidad sobre la coherencia para estas cantidades de tópicos. El valor que obtuvo la mayor métrica fue 23, con un indicador de 0.374. Se definió entonces que el modelo final se entrenará para generar 23 temas a partir del corpus.

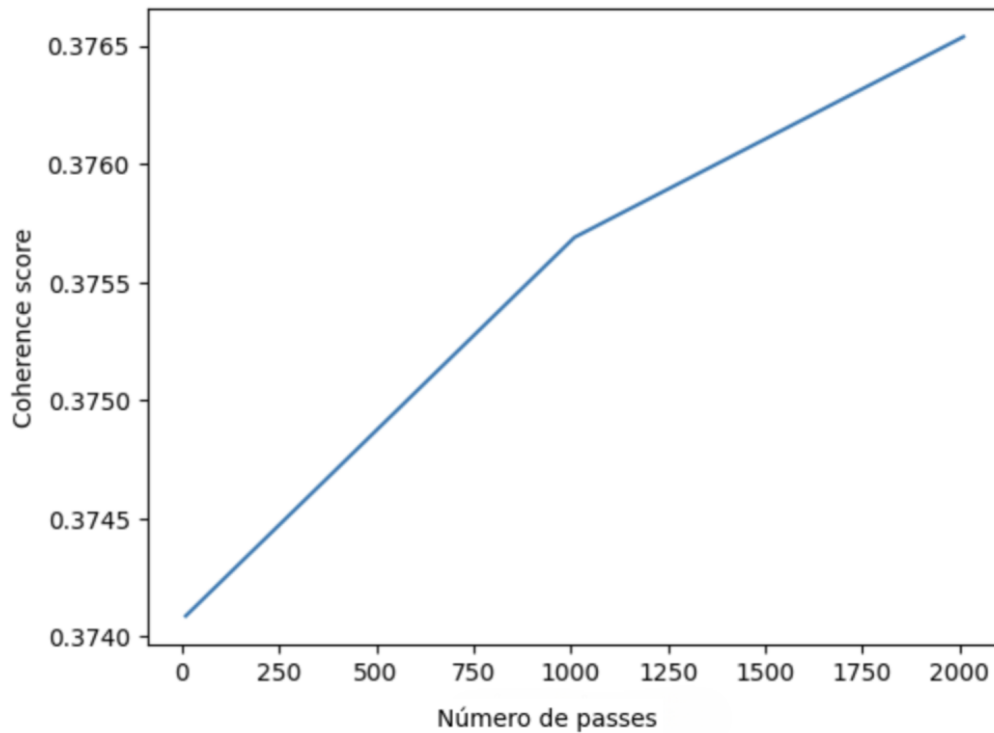
Figura 8. Valores de coherencia para el rango de 15 a 33 temas.



#### 3.4.1.2. Passes

En segundo lugar, el mismo procedimiento se realizó para el parámetro *passes*. Este define cuántas veces en entrenamiento debe pasar el modelo por todo el corpus antes de generar los temas. La limitante computacional es crucial para definir el número de passes ya que cuanto mayor se defina el parámetro, más tiempo de entrenamiento y capacidad de procesamiento se requiere.

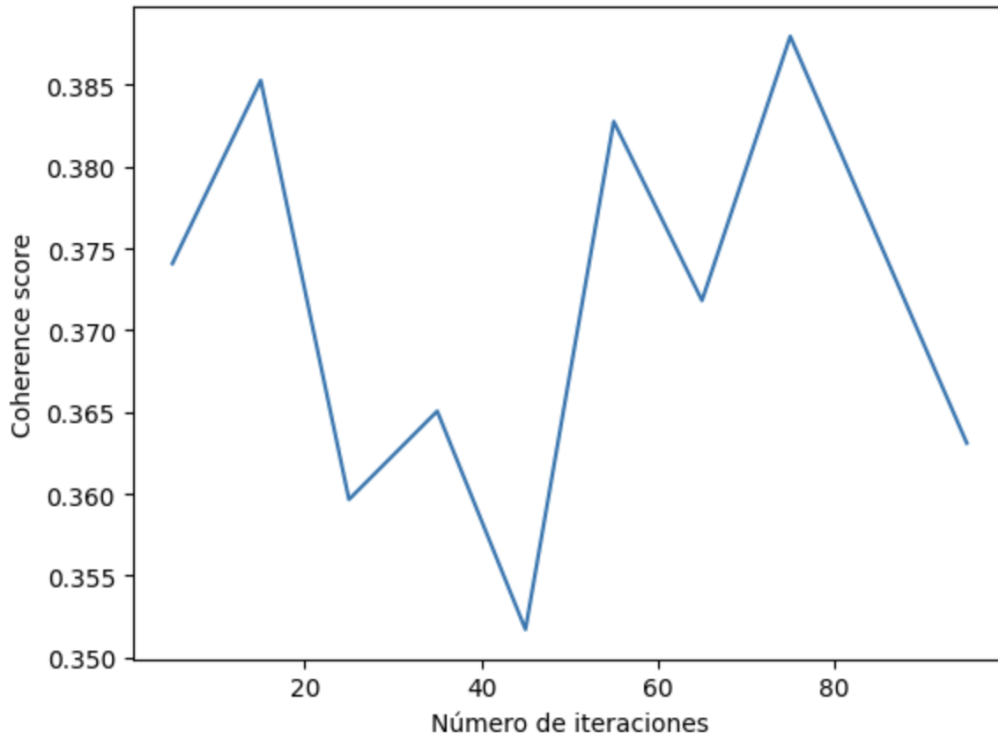
Se entrenaron modelos en un rango de 1 y 2,000 passes y para cada uno de ellos se calculó la coherencia. Como se puede observar en la figura 8, cuanto mayor el volumen de *passes* mayor el valor de coherencia.

Figura 9. Valores de coherencia para el rango de 1 a 2,000 *passes*.


Considerando la capacidad computacional y en línea con los resultados obtenidos en la primera iteración, se hicieron dos entrenamientos más. El primero con 4,000 *passes* y el segundo con 4,500 *passes* y se obtuvieron los siguientes valores de coherencia: 0.381 para 4,000 y 0.377 para 4,500. Dado que el modelo con 4,500 *passes* no genera un incremento en la coherencia, se definió entrenar el modelo con 4,000 *passes*.

### 3.4.1.3. Iterations

Por último, se aplicó el mismo esquema a *iterations* que determina cuántas veces el modelo ajusta los pesos de los temas y las distribuciones de palabras. Se entrenaron modelos considerando valores entre el rango de 10 a 125. Según los resultados (figura 9), el número de iteraciones que generó una mayor coherencia fue 75 con un valor de 0.388. Este será entonces el valor con el que se entrenará el modelo final.

Figura 10. Valores de coherencia para el rango de 10 a 125 *iterations*.


### 3.5. Embeddings

Los embeddings son representaciones vectoriales de los contextos de ocurrencia de las palabras con otras palabras (Jurafsky y Martin 2023). Las palabras se representan como un punto en un espacio matemático. Esta metodología fue desarrollada para convertir en más densas las representaciones vectoriales de las palabras (Garg, N., Schiebinger, L., Jurafsky, D. y Zou, J., 2018).

Como explican los autores, usar embeddings y listas de palabras permite medir la fuerza de asociación entre las segundas y un grupo, por ejemplo grupos de género. Esto basándose en la premisa de que términos cercanos tendrán vectores que apunten en la misma dirección o sean similares. En este trabajo se utilizaron embeddings para analizar qué representaciones obtienen las palabras de las listas definidas, buscando comprender si en el corpus se reproducen estereotipos midiendo de la fuerza de asociación entre el género y cada uno de los términos.

#### 3.5.1. Word2Vec

Word2Vec es un modelo de embeddings estáticos, es decir, que aprende un embedding fijo para cada palabra del corpus. Siguiendo a Jurafsky y Martin (2023) la intuición detrás se fundamenta en entrenar un clasificador con la tarea de predecir si es probable que una término  $x$  aparezca cerca del término  $y$ . La importancia del modelo radica en los pesos aprendidos tomándolos como los embeddings, no la

predicción en sí, y lo revolucionario, según los autores, es la posibilidad de usar el corpus como datos de entrenamiento a partir de la auto-supervisión, evitando la necesidad de señales etiquetadas a mano.

El modelo se basa en la premisa de que si dos palabras comparten contextos también son similares en cuanto a su significado y por ende tendrán una representación vectorial semejante. En consecuencia, Word2Vec puede descubrir relaciones entre términos dentro del corpus y calcular la similitud entre ellas.

El algoritmo skip-gram es uno de los dos posibles a utilizar al entrenar Word2Vec y fue introducido por Mikolov, et al. (2013). Este es un método eficiente para aprender representaciones vectoriales de palabras en un corpus no estructurado ya que se construye en la no necesidad de multiplicaciones de matrices densas. En este trabajo se utilizará skip-gram con muestreo negativo, popularmente llamado SGNS.

A diferencia de otros modelos basados en redes neuronales, Word2Vec simplifica el procedimiento de detectar las asociaciones en dos aspectos (Jurafsky y Martin 2023). Por un lado, realiza una clasificación binaria en lugar de una predicción de términos y, por otro lado, simplifica la arquitectura al entrenar un clasificador de regresión logística en lugar de una red neuronal multicapa.

En conclusión, el modelo de Word2Vec con el algoritmo de skip-gram ejecuta los siguientes pasos en la construcción de embeddings:

1. Considera el término objetivo y una palabra de contexto vecina como ejemplos positivos.
2. Muestra aleatoriamente otras palabras en el léxico a fin de obtener ejemplos negativos.
3. Usa la regresión logística para entrenar un clasificador con la tarea de distinguir los casos positivos de los negativos.
4. Establece los pesos aprendidos como los embeddings.

Un punto importante a considerar respecto a esta metodología es el volumen de datos con el que se entrenará el modelo. Para este trabajo se cuenta con un corpus de 3,855 documentos que resultan en 1,446,526 tokens de los cuales 343,566 son únicos. Los modelos de embeddings populares suelen entrenarse con volúmenes de datos que exceden el disponible en esta tesis por lo que es posible que el modelo no logré encontrar diferencias significativas.

En este trabajo se entrenó Word2Vec utilizando la librería Gensim (Radim Řehůřek, 2024). Los parámetros considerados fueron:

- *“vector\_size”*: es el número de dimensiones (N) del espacio N-dimensional al que mapea gensim Word2Vec los términos. Configurarlos en valores más altos requiere más datos de entrenamiento pero mejora la precisión.
- *“window”*: establece la distancia máxima entre la palabra actual y la predicha en la oración.
- *“min\_count”*: define el mínimo de frecuencia e ignora todas las palabras que ocurren menos veces de lo seteado. Esto permite no considerar en el entrenamiento tokens que ocurran muy poco frecuentemente en el corpus.

- *“negative”*: si se configura mayor a 0 el modelo utilizará muestreo negativo. El valor seteado indica cuántos términos de ruido se deben seleccionar, usualmente se utilizan valores entre 5 y 20. El conjunto de palabras que el modelo seleccionará en este punto varía en cada entrenamiento, cómo consecuencia los resultados pueden cambiar de acuerdo a la muestra.
- *“sample”*: es el umbral para configurar que palabras de frecuencia más alta se reducen aleatoriamente. Comúnmente se configura entre 0 y  $1e-5$ .
- *“sg”*: Esto indica que algoritmo de entrenamiento se usará y como se mencionó se utilizó skip-gram.
- *“epochs”*: define el número de iteraciones sobre todo el corpus. Es importante destacar en este punto la complejidad computacional de establecer un valor de epochs demasiado alto.

Los parámetros establecidos para el entrenamiento fueron los siguientes:

- *“vector\_size”*: se estableció en 200 buscando incrementar la precisión.
- *“window”*: al igual que en la matriz de co-ocurrencias, se definió una ventana de 5.
- *“min\_count”*: el conteo mínimo de frecuencias en el entrenamiento será de 10. Es decir, todos los términos que ocurran menos de 10 veces en el corpus no serán considerados.
- *“negative”*: con el propósito mejorar la precisión se definió el muestreo negativo en 20. El modelo seleccionará 20 palabras de ruido en el entrenamiento.
- *“sample”*: se entrenará con un sample de 0.001. De esta manera se buscará minimizar el impacto de palabras muy frecuentes.
- *“epochs”*: por último, se harán 150 iteraciones en el corpus.

### 3.5.2. Similitud Coseno

La métrica más utilizada con el objetivo de medir asociaciones en embeddings es la similitud coseno (Jurafsky y Martin, 2023). Este indicador calcula el coseno del ángulo entre los vectores y tiende a ser alto cuando los dos vectores tienen valores grandes en la misma dirección.

En esta tesis se replicó lo propuesto por Garg, N. et al. (2018) en su trabajo realizado para cuantificar sesgos de género y raciales en embeddings. El procedimiento consiste de los siguientes pasos:

1. Calcular la similitud promedio de las palabras que representan al género femenino y una lista de palabras.
2. Aplicar el mismo cálculo a los términos que representan al masculino.
3. Computar la diferencia entre ambos resultados, es decir, la similitud promedio para las palabras del género femenino menos la distancia promedio para el masculino.

A partir de esta diferencia se pueden detectar sesgos en los embeddings, entendiendo que si el valor es negativo los términos considerados tienen una mayor asociación con el género masculino.

Siguiendo lo propuesto por Lewis y Lupyan (2020) en este análisis se amplió la lista de términos asociados al género, buscando extender el alcance de los resultados y mejorar la precisión de las conclusiones. La similitud coseno se calculó para las listas de términos previamente definidas y las siguientes palabras:

- Masculino: *he, him, man, boy, brother, father, son.*
- Femenino: *she, her, woman, girl, sister, mother, daughter.*

### 3.5.3. Word Embedding Association Test (WEAT)

Por último, en este trabajo también se utilizó el Word Association Test (WEAT) (Schröder, S. et. al., 2021). Este compara dos conjuntos de palabras objetivos con dos grupos de atributos de género de igual tamaño. Se basa en la hipótesis de que uno de los dos está más asociado a un género que al otro y viceversa. Con la finalidad de cuantificar el sesgo en los conjuntos utiliza el effect sizes, una medida normalizada para la diferencia de asociación. Un valor positivo de effect sizes confirma la hipótesis mientras que si es negativo indica estereotipos en la dirección opuesta.

El cálculo de esta métrica se compone de tres pasos:

1. La asociación de un término  $W$  con los dos segmentos de atributos ( $A$  y  $B$ ) se calcula utilizando la similitud coseno entre  $W$  y cada uno de los grupos. El resultado se obtiene sumando las similitudes coseno entre la palabra y cada atributo en  $A$  y restando las similitudes coseno entre la palabra y cada atributo en  $B$ .

$$s(\mathbf{w}, A, B) = \frac{1}{n} \sum_{\mathbf{a} \in A} \cos(\mathbf{w}, \mathbf{a}) - \frac{1}{n} \sum_{\mathbf{b} \in B} \cos(\mathbf{w}, \mathbf{b}).$$

2. El effect sizes mide la diferencia de asociación promedio entre los conjuntos de palabras ( $X$  e  $Y$ ) con los conjuntos de atributos. Se obtiene a partir de la diferencia entre la media de las asociaciones de los términos en  $X$  y en  $Y$ , dividida por la desviación estándar de todas las relaciones de las palabras en  $X$  y en  $Y$ .

$$d = \frac{\text{mean}_{\mathbf{x} \in X} s(\mathbf{x}, A, B) - \text{mean}_{\mathbf{y} \in Y} s(\mathbf{y}, A, B)}{\text{stddev}_{\mathbf{w} \in X \cup Y} s(\mathbf{w}, A, B)},$$



- Por último, para medir la significancia estadística de los sesgos se utiliza un test de permutación. A partir de la generación de particiones  $(X_i, Y_i)$  se calcula la probabilidad de observar un valor de  $s(X_i, Y_i, A, B)$  mayor que  $s(X, Y, A, B)$  por azar (p-value). La intuición es simular un muestreo aleatorio bajo la hipótesis nula de que no hay una asociación entre los conjuntos  $X$  e  $Y$  y los atributos  $A$  y  $B$ .

$$s(X, Y, A, B) = \sum_{x \in X} s(x, A, B) - \sum_{y \in Y} s(y, A, B)$$

El p-value se obtiene observando cuántas veces el estadístico de prueba, obtenido a partir de permutar los conjuntos de palabras  $X$  e  $Y$  ( $X_i, Y_i$ ), es más alto que el estadístico de prueba observado con los conjuntos originales. Si el p-value es chico, usualmente menos de 0.05, se rechazaría la hipótesis nula y se confirma una asociación significativa entre las variables analizadas.

$$p = P_r[s(X_i, Y_i, A, B) > s(X, Y, A, B)].$$

En la aplicación de esta metodología se trabajó únicamente con los términos relacionados al trabajo y a lo doméstico, buscando entender si el corpus utilizado evidencia sesgos de género vinculados a las ocupaciones. Esta decisión se fundamenta en que es factible construir segmentos comparables para testear la hipótesis planteada, a diferencia de lo que sucede con la lista de términos relacionados a las emociones, sentimientos y cuidado. La selección de cada grupo de palabras a contrastar surge de la literatura consultada en donde se evidencia que existen estereotipos acerca de que los hombres desarrollan con más frecuencia carreras profesionales de alto nivel cognitivo en los campos STEM mientras que las mujeres se dedican a las tareas de la casa o al cuidado y la enseñanza. Los grupos se definen entonces como:

- Grupo de términos asociados al género masculino: *doctor, medicine, surgeon, science, engineer, technologist, technician, police y football*.
- Grupo de términos asociados al género femenino: *house, cooking, cooked, vacuum, babysit, nurse, healer, tech, teacher*.

## 4. Análisis de Resultados

### 4.1. Positive Pointwise Mutual Information (PPMI)

#### 4.1.1. Palabras asociadas a las emociones, sentimientos y cuidado

En este grupo de palabras pareciera haber una mayor asociación general con el género femenino (tabla 2). Del total de la lista en un 53% los PPMIs fueron mayores entre el término y *she*. En cuanto al pronombre *he*, un 20% obtuvo valores más elevados y hubo un 27% iguales a 0.

Las palabras que muestran una asociación más fuerte al pronombre femenino que al masculino, de acuerdo a las diferencias entre ambos PPMIs, son: *little, poor, hurt, cry, cried, feeling, rescue* y *healthy*. En cambio, los términos más asociados al masculino son: *sad, scare* y *hero*. De ello se desprende la idea de que, a pesar que *sad* y *scare* obtuvieron valores mayores para *he*, la expresión de estos sentimientos como *cry* o *cried* están más asociados a *she*. Algo similar sucede con *feeling*, término que tuvo un valor de PPMI más elevado con *she*.

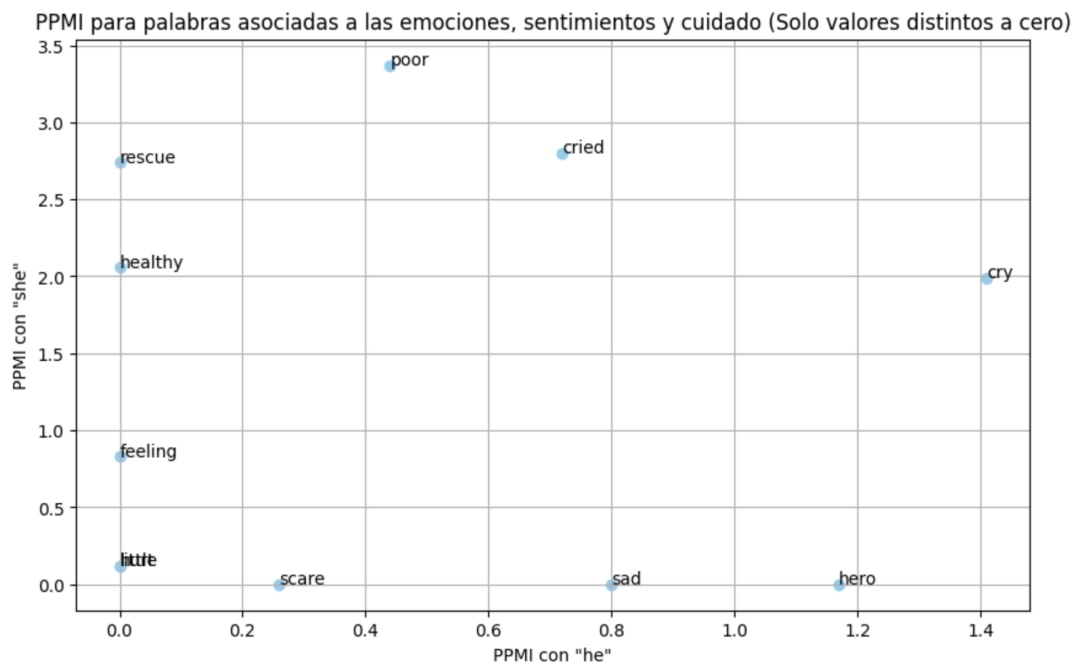
Por otro lado, aunque *rescue* mostró más ocurrencia conjunta con *she*, la palabra *hero* se asocia más con *he*, lo que nos lleva a pensar que la noción de héroe está estereotipada hacia el género masculino. Para entender mejor esta tendencia se recuperaron ejemplos de contextos con las co-ocurrencias. Respecto a la asociación del género femenino con *rescue* se encontraron casos positivos, en donde el personaje mujer se presenta como el sujeto que rescata a otro, por ejemplo "*she is going to rescue daddy*". No obstante, no existen co-ocurrencias con *hero*. A diferencia, los contextos que contienen la co-ocurrencia de *he* con *hero* lo hacen presentando al personaje masculino como fuerte y virtuoso, por ejemplo "*it was tough on our hero, but he never backed down*".

Por último, la palabra *healthy* obtuvo uno de los PPMIs más altos asociado con el pronombre femenino, pero su valor con *he* fue cero. Esta mayor asociación podría responder a estereotipos asociados al cuidado personal atado a los estándares hegemónicos de belleza de los cuales las mujeres son muchas veces víctimas pero también a la idea de que son las mujeres las que están encargadas del cuidado de los otros. Para comprender el resultado se analizaron algunos contextos con la co-ocurrencia. Se encontraron frases como "*she'll make sure you're healthy in every way*" relacionadas al cuidado hacia un otro pero también asociadas al cuidado personal "*she's so yummy eating sweets, but is not healthy, take care of your health*".

Tabla 2. PPMIs para palabras asociadas a las emociones, sentimientos y cuidado.

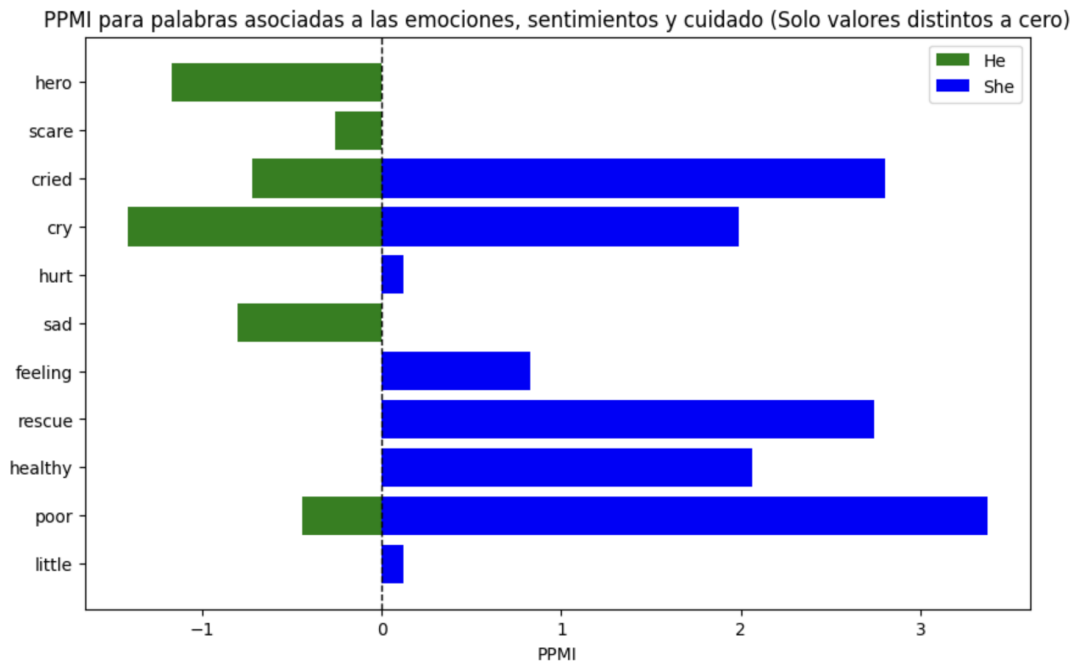
| palabra    | he   | she  | dif   |
|------------|------|------|-------|
| little     | 0    | 0.12 | -0.12 |
| poor       | 0.44 | 3.37 | -2.93 |
| sad        | 0.8  | 0    | 0.8   |
| hurt       | 0    | 0.12 | -0.12 |
| cry        | 1.41 | 1.99 | -0.58 |
| cried      | 0.72 | 2.8  | -2.08 |
| scare      | 0.26 | 0    | 0.26  |
| feeling    | 0    | 0.83 | -0.83 |
| violent    | 0    | 0    | 0     |
| irrational | 0    | 0    | 0     |
| kind       | 0    | 0    | 0     |
| rescue     | 0    | 2.74 | -2.74 |
| hero       | 1.17 | 0    | 1.17  |
| clean      | 0    | 0    | 0     |
| healthy    | 0    | 2.06 | -2.06 |

Figura 11. Visualización de PPMIs para palabras asociadas a las emociones, sentimientos y cuidado.



Entre la lista de palabras analizada, se destaca la distancia en los términos *poor*, *healthy*, *rescue* y *cried*. En todas ellas la diferencia de PPMI entre *he* y *she* fue mayor a 2 (figura 11).

Figura 12. Diferencias de PPMIs para palabras asociadas a las emociones, sentimientos y cuidado.



#### 4.1.2. Palabras asociadas al trabajo y a lo doméstico

En la mayoría de las palabras los valores de PPMI son más altos para el género femenino, solo en dos ocasiones el PPMI fue mayor con *he* (tabla 3). En primer lugar, no se observa una asociación más fuerte de profesiones médicas con el pronombre masculino, sino que tanto para *doctor*, *nurse* y *medicine* el PPMI fue mayor con *she*. Esto indicaría que no hay un estereotipo similar a los mencionado en la revisión de literatura presente en el corpus. Ahora bien, la palabra *science* sí obtuvo un PPMI más elevado con el pronombre masculino, lo que lleva a pensar que a pesar de que hay una relación de las mujeres con campos médicos no lo hay respecto a campos científicos. No parece existir un estereotipo asociado al trabajo dado que las palabras *working*, *worked*, *professional* y *job* tuvieron valores de PPMI mayores para *she*.

En segundo lugar, al igual que con los términos asociados al cuidado, en este grupo se observa una mayor asociación del género femenino con el mantenimiento del hogar. Esto a partir de PPMIs más altos para *she* y palabras como *cooked*, *vacuum* y *babysit*. En línea similar, las expresiones relacionadas a la enseñanza (*teach* y *teacher*) también parecieran estar más vinculadas al género femenino.

Por último, la palabra *football* obtuvo un PPMI de 0.92 con *he* y de 0 con *she*. En base a estos resultados se podría concluir que este deporte tiene un vínculo más fuerte con el género masculino, lo cual podría desprender estereotipos relacionados.

Tabla 3. PPMIs para palabras asociadas al trabajo y a lo doméstico.

| <b>palabra</b> | <b>he</b> | <b>she</b> | <b>dif</b> |
|----------------|-----------|------------|------------|
| job            | 0         | 0.35       | -0.35      |
| responsibility | 0.85      | 3.39       | -2.54      |
| working        | 0         | 0.64       | -0.64      |
| worked         | 0.6       | 1.21       | -0.61      |
| professional   | 0.95      | 1.97       | -1.02      |
| chief          | 0.12      | 0.92       | -0.8       |
| school         | 0         | 0.16       | -0.16      |
| teacher        | 0         | 0.66       | -0.66      |
| teach          | 0         | 1.35       | -1.35      |
| doctor         | 1.17      | 1.87       | -0.7       |
| medicine       | 1.32      | 2.31       | -0.99      |
| surgeon        | 0         | 0          | 0          |
| healer         | 1.12      | 2.18       | -1.06      |
| nurse          | 1.7       | 2.1        | -0.4       |
| football       | 0.92      | 0          | 0.92       |
| police         | 0         | 0          | 0          |
| architecture   | 0         | 0          | 0          |
| engineer       | 0         | 0          | 0          |
| technician     | 0         | 0          | 0          |
| technologist   | 0         | 0          | 0          |
| science        | 2.71      | 0          | 2.71       |
| house          | 0         | 0          | 0          |
| vacuum         | 0         | 2.74       | -2.74      |
| cooking        | 0         | 0          | 0          |
| cooked         | 1.97      | 2.92       | -0.95      |
| babysit        | 1.43      | 2.32       | -0.89      |

Figura 13. Visualización de PPMIs para palabras asociadas al trabajo y a lo doméstico.

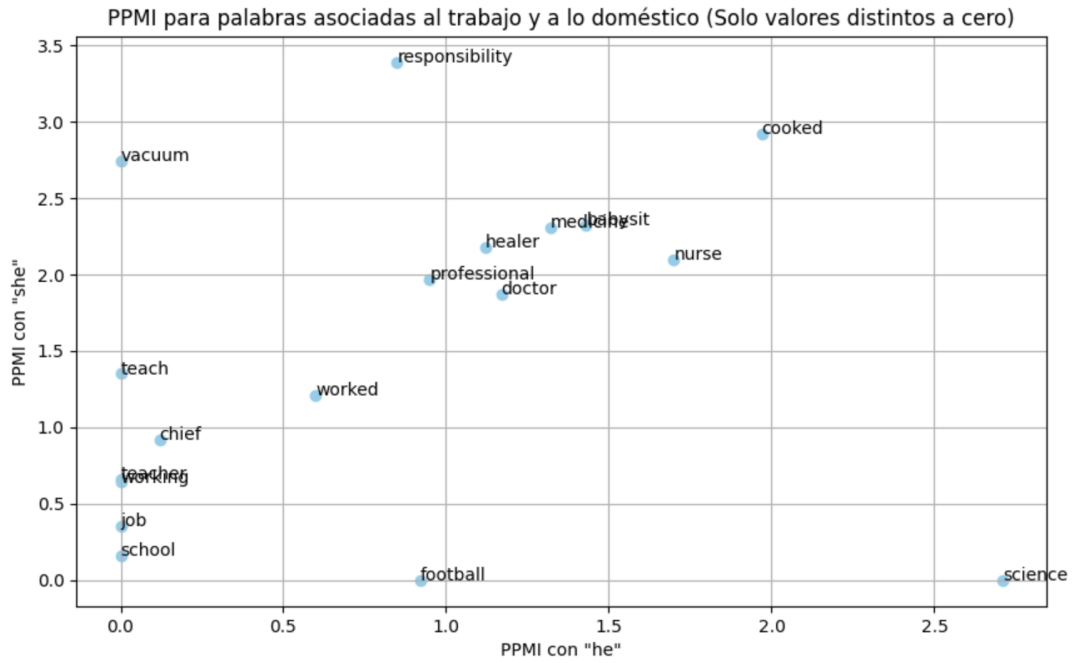
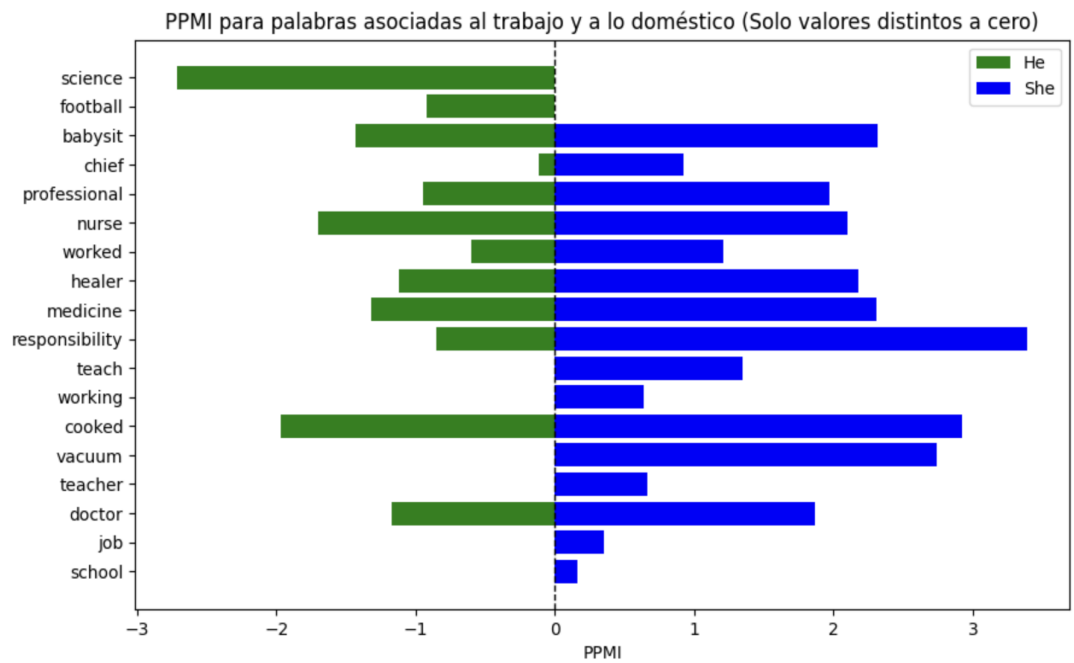


Figura 14. Diferencias de PPMIs para palabras asociadas al trabajo y a lo doméstico.



#### 4.1.3. PPMI con Laplace

La modificación de la matriz de co-ocurrencias con el suavizado de Laplace no tuvo gran impacto en los cálculos del PPMI. No se observan diferencias significativas en los valores respecto a los obtenidos sin el suavizado y no hay variaciones en los descubrimientos (tablas 4 y 5).

El suavizado de Laplace parece no ser adecuado para mitigar el impacto de palabras poco frecuentes dado que este reserva una masa de probabilidad desproporcionada para los mismos (Manning y Schütze, 1999).

Tabla 4. PPMIs con suavizado para palabras asociadas a las emociones, sentimientos y cuidado.

| palabra    | sin-laplace |      |       | laplace |      |       |
|------------|-------------|------|-------|---------|------|-------|
|            | he          | she  | dif   | he      | she  | dif   |
| little     | 0           | 0.12 | -0.12 | 0       | 0.27 | -0.27 |
| poor       | 0.44        | 3.37 | -2.93 | 0.32    | 3.14 | -2.82 |
| sad        | 0.8         | 0    | 0.8   | 0.78    | 0    | 0.78  |
| hurt       | 0           | 0.12 | -0.12 | 0       | 0    | 0     |
| cry        | 1.41        | 1.99 | -0.58 | 1.09    | 1.61 | -0.52 |
| cried      | 0.72        | 2.8  | -2.08 | 0.62    | 2.53 | -1.91 |
| scare      | 0.26        | 0    | 0.26  | 0.03    | 0    | 0.03  |
| feeling    | 0           | 0.83 | -0.83 | 0       | 0.71 | -0.71 |
| violent    | 0           | 0    | 0     | 0       | 0    | 0     |
| irrational | 0           | 0    | 0     | 0       | 0    | 0     |
| kind       | 0           | 0    | 0     | 0       | 0    | 0     |
| rescue     | 0           | 2.74 | -2.74 | 0       | 2.61 | -2.61 |
| hero       | 1.17        | 0    | 1.17  | 0.81    | 0    | 0.81  |
| clean      | 0           | 0    | 0     | 0       | 0    | 0     |
| healthy    | 0           | 2.06 | -2.06 | 0       | 1.94 | -1.94 |

Tabla 5. PPMIs con suavizado para palabras asociadas al trabajo y a lo doméstico.

| palabra        | sin-laplace |      |       | laplace |      |       |
|----------------|-------------|------|-------|---------|------|-------|
|                | he          | she  | dif   | he      | she  | dif   |
| job            | 0           | 0.35 | -0.35 | 0       | 0.37 | -0.37 |
| responsibility | 0.85        | 3.39 | -2.54 | 0.67    | 2.69 | -2.02 |
| working        | 0           | 0.64 | -0.64 | 0       | 0.46 | -0.46 |
| worked         | 0.6         | 1.21 | -0.61 | 0       | 0.55 | -0.55 |
| professional   | 0.95        | 1.97 | -1.02 | 0.75    | 1.74 | -0.99 |
| chief          | 0.12        | 0.92 | -0.8  | 0       | 0.49 | -0.49 |
| school         | 0           | 0.16 | -0.16 | 0       | 0.08 | -0.08 |
| teacher        | 0           | 0.66 | -0.66 | 0       | 0.44 | -0.44 |
| teach          | 0           | 1.35 | -1.35 | 0       | 0.87 | -0.87 |
| doctor         | 1.17        | 1.87 | -0.7  | 1.26    | 1.88 | -0.62 |
| medicine       | 1.32        | 2.31 | -0.99 | 0.86    | 1.8  | -0.94 |
| surgeon        | 0           | 0    | 0     | 0       | 0    | 0     |
| healer         | 1.12        | 2.18 | -1.06 | 0.85    | 1.86 | -1.01 |
| nurse          | 1.7         | 2.1  | -0.4  | 1       | 1.5  | -0.5  |
| football       | 0.92        | 0    | 0.92  | 0.76    | 0    | 0.76  |
| police         | 0           | 0    | 0     | 0       | 0    | 0     |
| architecture   | 0           | 0    | 0     | 0       | 0    | 0     |
| engineer       | 0           | 0    | 0     | 0       | 0    | 0     |
| technician     | 0           | 0    | 0     | 0       | 0    | 0     |
| technologist   | 0           | 0    | 0     | 0       | 0    | 0     |
| science        | 2.71        | 0    | 2.71  | 2.64    | 0    | 2.64  |
| house          | 0           | 0    | 0     | 0       | 0    | 0     |
| vacuum         | 0           | 2.74 | -2.74 | 0       | 2.55 | -2.55 |
| cooking        | 0           | 0    | 0     | 0       | 0    | 0     |
| cooked         | 1.97        | 2.92 | -0.95 | 1.73    | 2.64 | -0.91 |
| babysit        | 1.43        | 2.32 | -0.89 | 1.49    | 2.42 | -0.93 |

#### 4.1.4. Hipótesis EDIA

Retomando las hipótesis sobre las asociaciones de las listas de palabras planteadas a partir de los resultados obtenidos en la exploración inicial con EDIA, se calcularon los PPMI promedios para cada género y cada grupo de términos. Estos cálculos se hicieron considerando el PPMI sin el suavizado laplaciano.

Las hipótesis eran las siguientes:



1. En las palabras asociadas a las emociones, sentimientos y cuidado la cercanía es mayor a:
  - a. he: *little, poor, healthy, rescue, feeling, kind, clean, hurt, scare, hero y violent.*
  - b. she: *sad, cry, cried e irrational.*
  
2. En las palabras relacionadas al trabajo y a lo doméstico la proximidad es más alta en:
  - a. he: *job, working, doctor, medicine, healer, worked, professional, chief, football, engineer, surgeon y police.*
  - b. she: *school, teacher, house, vacuum, cooking, cooked, responsibility, teach, nurse, babysit, architecture, technician, science y technologist.*

Hipótesis 1: palabras vinculadas a las emociones, sentimientos y cuidado: en este caso la suposición no se comprueba. Si bien, como se esperaba, el PPMI promedio de *she* es más elevado con el grupo b, también lo es con el segmento a (tabla 6). Los resultados no demuestran una mayor asociación de los términos del grupo a con el género masculino.

Tabla 6. PPMIs promedio según género y grupo (palabras asociadas a las emociones sentimientos y cuidado).

| <b>pronombre</b> | <b>grupo a</b> | <b>grupo b</b> |
|------------------|----------------|----------------|
| he               | 0.17           | 0.73           |
| she              | 0.84           | 1.20           |

Hipótesis 2: palabras asociadas al trabajo y a lo doméstico: de la misma manera que con en el análisis anterior, ambos segmentos tuvieron un PPMI mayor para *she* (tabla 7). Esto indica que no existe una vinculación más fuerte de los términos del conjunto a con el género masculino. Ahora bien, en este caso las diferencias entre los promedios son mayores para el segmento b, lo que demuestra una distancia más elevada de las palabras del grupo con el género masculino.

Tabla 7. PPMIs promedio según género y grupo (palabras asociadas al trabajo y a lo doméstico).

| <b>pronombre</b> | <b>grupo a</b> | <b>grupo b</b> |
|------------------|----------------|----------------|
| he               | 0.52           | 0.62           |
| she              | 0.95           | 1.12           |

En conclusión, en base a los resultados obtenidos a partir del cálculo del PPMI, el corpus de este trabajo no parece representar las mismas asociaciones que EDIA identifica en la lista de palabras. Esto puede estar relacionado a que cada una de las aproximaciones, la basada en PPMI y la realizada con

EDIA, representan diferentes lenguajes. Efectivamente, en este trabajo hemos construido las matrices de coocurrencia, los embeddings y todos los cálculos derivados sobre el corpus de subtítulos de YouTube. En cambio, las exploraciones que se realizan a través de EDIA son sobre un embedding obtenido del Spanish Billion Word Corpus. Además, en cada una de estas aproximaciones el lenguaje se representa de una manera distinta. El PPMI considera cuánto co-ocurren las palabras y computa su asociación en base a ello. A diferencia, EDIA, que explora un modelo de embeddings, encuentra relaciones entre palabras a partir de establecer analogías entre sus contextos de ocurrencia y la probabilidad de que dos palabras ocurran en el mismo contexto.

#### 4.1.5. PPMI en contexto

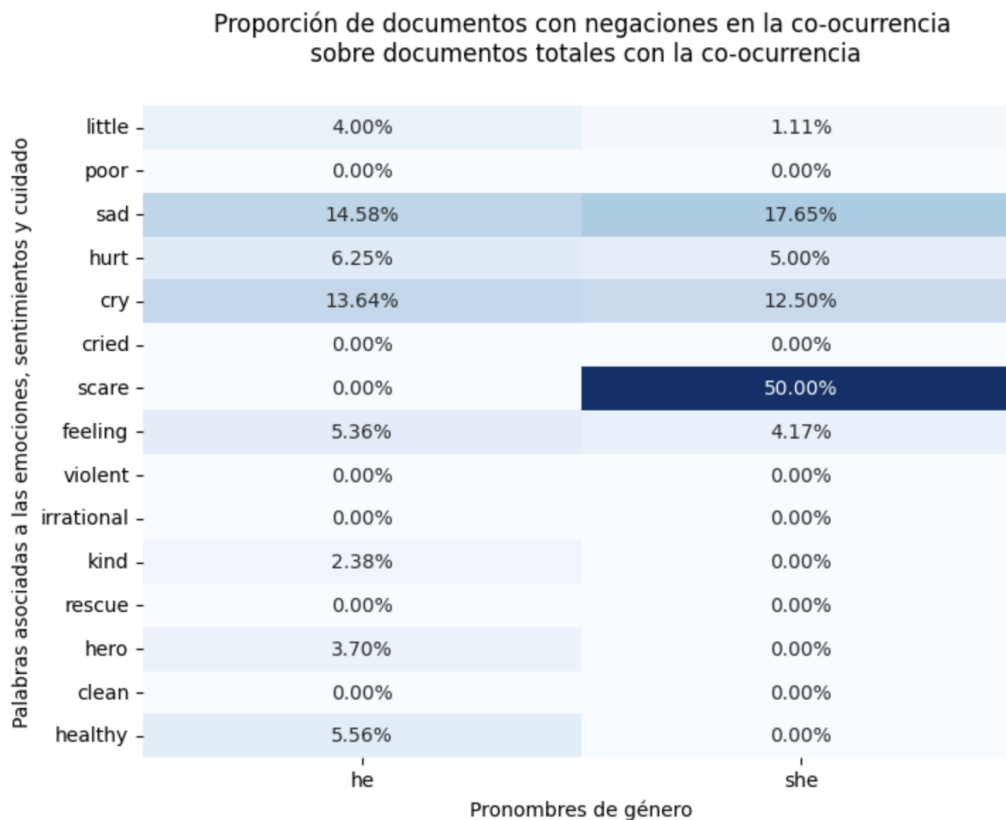
Entendiendo que el PPMI no considera si en las ventanas existen expresiones negativas y siguiendo el procedimiento previamente explicado, se analizaron los contextos de las co-ocurrencias para cada una de las palabras de la lista definida y los pronombres de género.

Por un lado, en la lista de términos asociados a las emociones, sentimientos y cuidado (figura 14), se destaca que en la ventana de las co-ocurrencias entre *she* y *rescue* un 50% contiene negaciones. No obstante, este resultado no es relevante dado que su PPMI fue 0.

En expresiones como *hurt*, *cry* y *feeling* se detectaron negaciones en los contextos pero el porcentaje es bastante similar entre ambos géneros, por lo que se entiende que las conclusiones no estarán sesgadas.

Ahora bien, existen algunos términos en donde las negaciones si podrían haber afectado los resultados obtenidos. Por ejemplo, en palabras como *little* y *healthy*, ambas con un PPMI más elevado con *she*, el porcentaje de negaciones fue mayor para los contextos con he: 2.9 y 5.56 puntos porcentuales respectivamente. Si el cálculo del PPMI considerase las mismas se esperaría que la diferencia entre los valores de asociación sea aún más amplia.

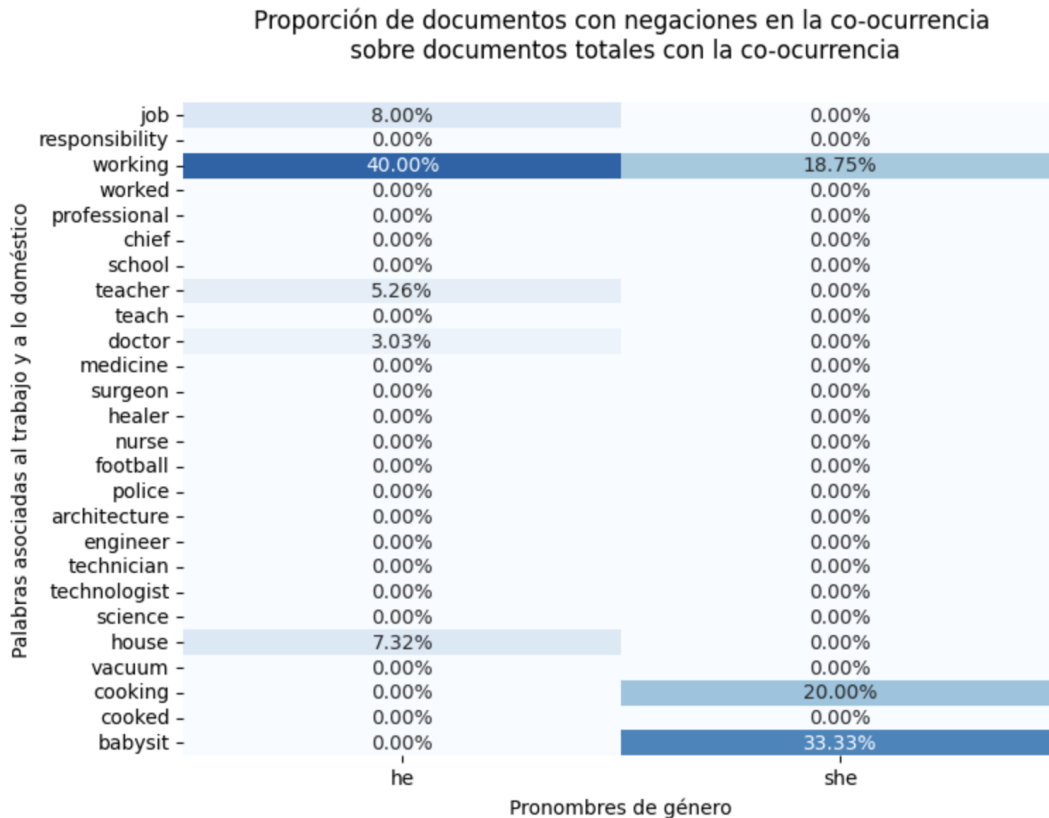
Figura 15. Porcentaje de negaciones en las co-ocurrencias para las palabras asociadas a las emociones, sentimientos y cuidado.



Por otro lado, en la lista de palabras asociadas al trabajo y a lo doméstico se evidencian 3 términos con contextos con negaciones considerables (figura 15). En primer lugar, en *working*, con un PPMI mayor para *she*, se encontró un 40% de negaciones en el contexto con *he*. Si bien existen también expresiones negativas en las co-ocurrencias con *she*, dada la diferencia entre los porcentajes se esperaría una distancia aún más elevada en los PPMIs con cada género.

En segundo lugar, se observan negaciones en un 20% de las co-ocurrencias de *she* y *cooking*. No obstante, con ambos pronombres el PPMI fue cero, por lo que no se espera una variación en los resultados. Por último, *babysit* demuestra negaciones en la concurrencia con *she*. Es con este pronombre con el cual el PPMI fue mayor, si el cálculo considerase las expresiones negativas debería reducirse el valor obtenido achicando la diferencia con el pronombre masculino.

Figura 16. Porcentaje de negaciones en las co-ocurrencias para las palabras asociadas al trabajo y a lo doméstico.



## 4.2. Latent-Dirichlet-Allocation (LDA)

### 4.2.1. Resultados del modelo final

A partir del procedimiento realizado para iterar sobre los hiper parámetros y encontrar la configuración que maximiza la coherencia, se infirió el modelo LDA final considerando los siguientes valores:

- "num\_topics": 23
- "passes": 4,000
- "iterations": 75

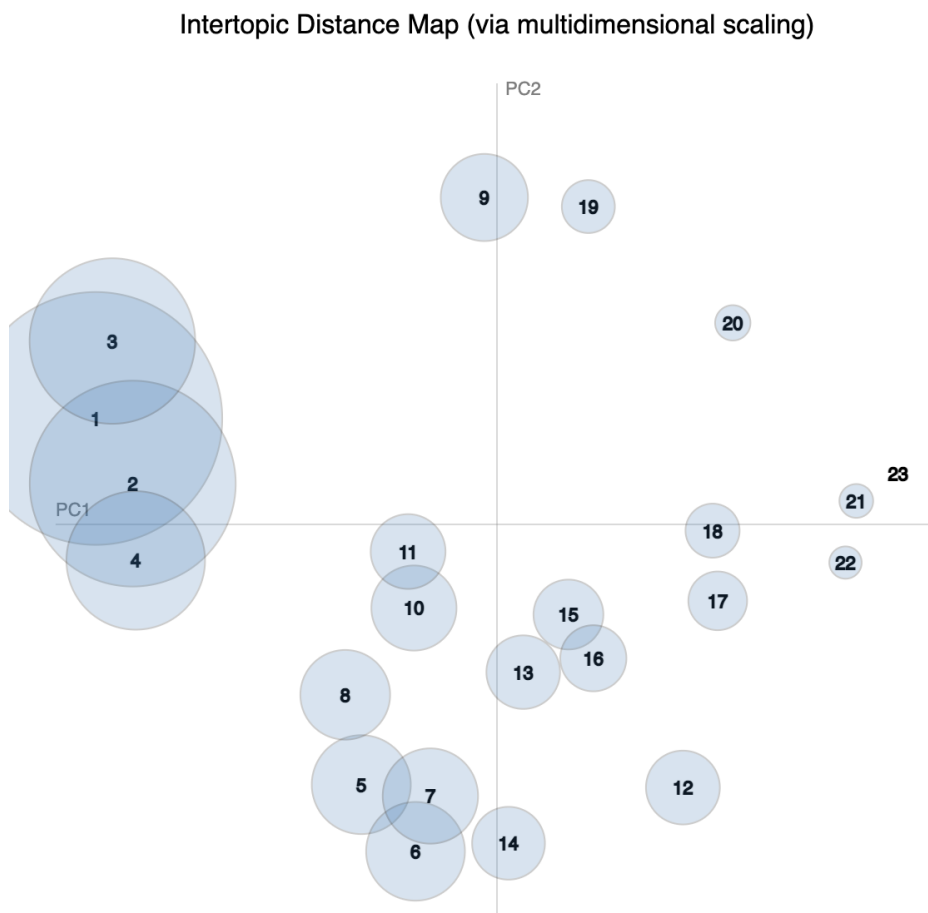
Con estos parámetros la coherencia del modelo fue de 0.39.

#### 4.2.2. Visualización de tópicos

A continuación se generó una visualización usando el paquete interactivo pyLDAvis. Cada burbuja representa un tema y cuanto mayor el tamaño, más prevalente es el tópico en el corpus. Un buen modelo LDA debería tener burbujas grandes y no superpuestas.

En el modelo entrenado en este trabajo, la visualización permite observar que las burbujas de los cuatro temas principales están bastante solapadas en el extremo izquierdo. El resto de las burbujas si bien son más dispersas tienen tamaños bastante más pequeños lo que indica que su representación del corpus es menor.

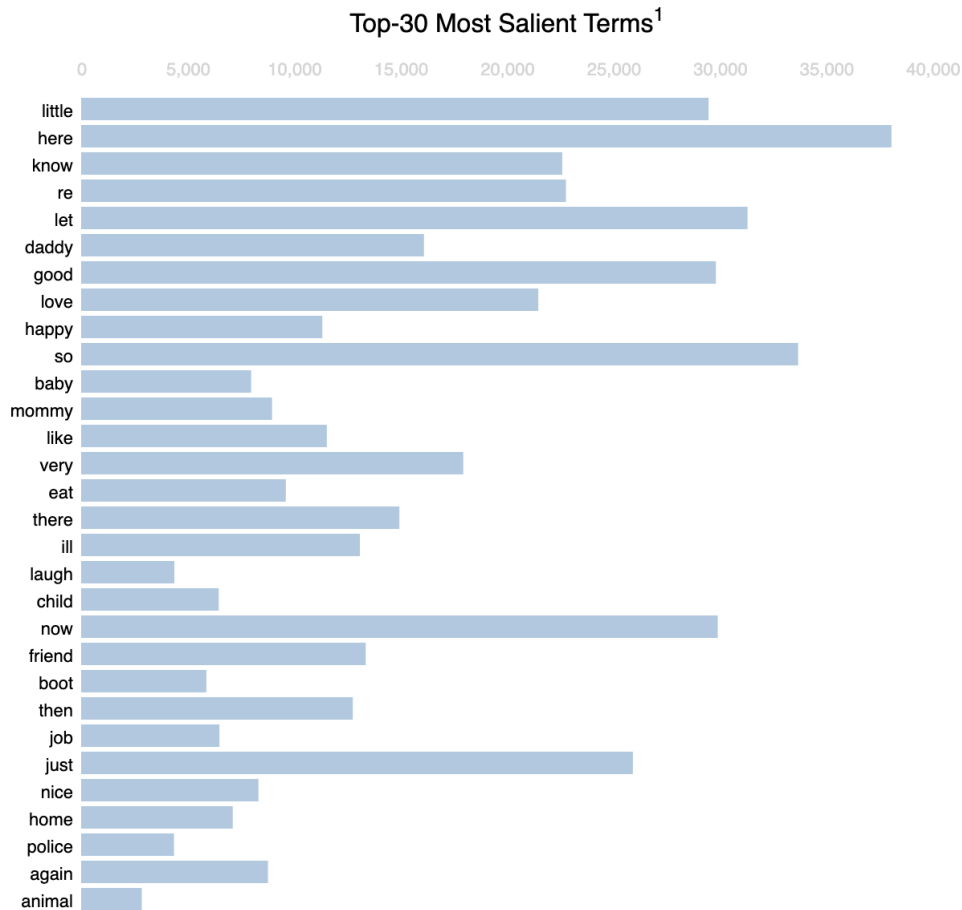
Figura 17. Mapa de tópicos.



La visualización también permite observar los términos más relevantes del modelo. La relevancia refiere al peso, un valor numérico asociado a las palabras que indica cuán central es cada una en los tópicos, en comparación con otras tanto del mismo tópico como de otros. Si bien se puede diferenciar una temática relacionada a la familia y los sentimientos positivos con palabras como *daddy*, *baby*, *mommy*, *child*, *good*, *love* y *happy*, no es evidente una tendencia estereotipada. Los términos más

relevantes de los 4 tópicos más prevalentes ([apéndice 6](#)) tampoco presentan claros patrones que permitan descubrir sesgos de género en el corpus.

Figura 18. Top 30 de términos más relevantes



### 4.3. Word2Vec

En base a los parámetros definidos en la metodología se entrenó un modelo Word2Vec a partir del corpus compuesto de subtítulos de videos de YouTube para niños en inglés. El tamaño del modelo final entrenado fue de 66,341,176 palabras en el vocabulario y cada palabra fue representada como un vector de 200,772,900 dimensiones.

#### 4.3.1. Similitud Coseno

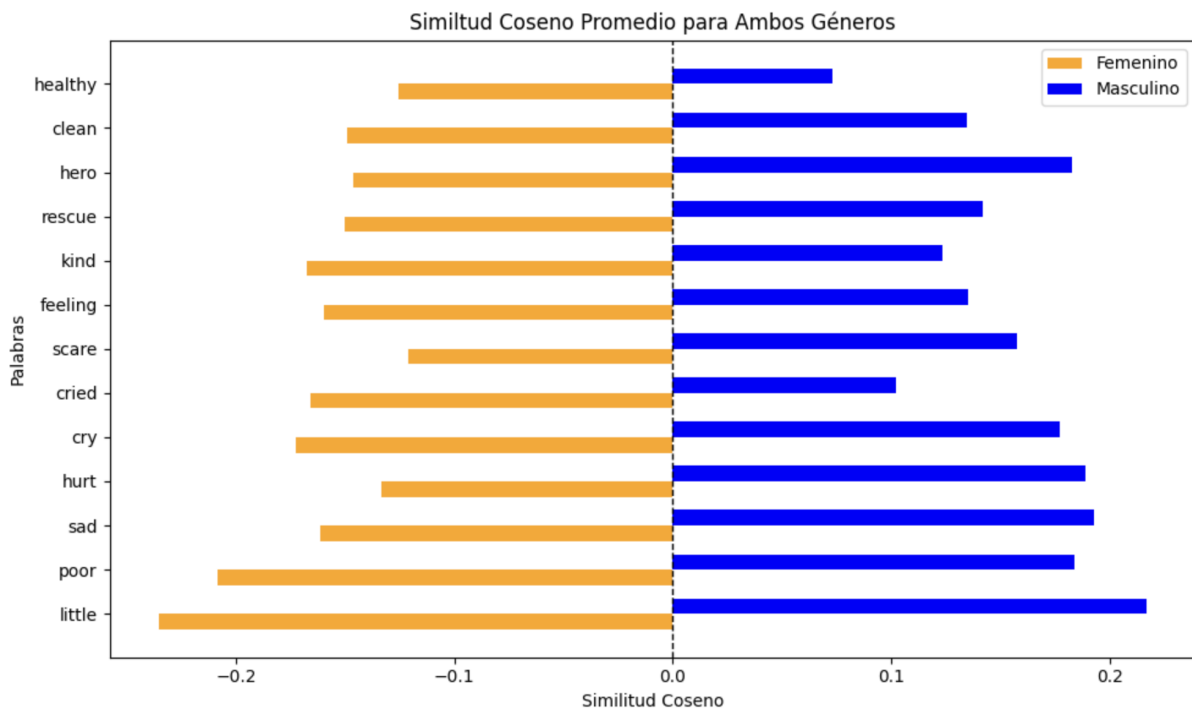
Como se explicó en la sección metodología, la similitud coseno será utilizada con el objetivo detectar asociaciones entre palabras que representan género y palabras que representan temas en los que se sospecha que puede haber sesgos de género en los embeddings generados por el modelo. En primer lugar se calcularon las similitudes promedio de cada una de los términos que representan a ambos

géneros y las dos listas de palabras previamente definidas. A partir de estos valores se computaron las diferencias, entendidas como el promedio entre la distancia con las palabras que representan el género femenino menos el promedio con las palabras que representan el género masculino. De esta manera, indicadores negativos reflejarán asociaciones más fuertes con los hombres.

#### 4.3.1.1. Palabras asociadas a las emociones, sentimientos y cuidado

Por un lado, en la lista de términos asociados a las emociones, sentimiento y al cuidado, las medias de la similitud coseno entre las palabras que representan a ambos géneros no muestran grandes diferencias. Esto está en línea con lo explicado anteriormente acerca del volumen de datos disponibles en este trabajo. Como se observa en la figura 17 los promedios para cada uno de los segmentos de género suelen tener valores similares. El modelo no parece estar diferenciando significativamente las asociaciones entre los términos.

Figura 19. Similitud coseno para palabras asociadas a las emociones, sentimientos y cuidado.



Ahora bien, cuando se computan las diferencias es posible observar algunas tendencias de asociaciones de género (figura 18). Un punto interesante a considerar es que gráficos como el de diferencias suelen ser tendenciosos y los resultados se observan como más significativos que lo real, es fundamental tener presente que las diferencias no son elevadas.

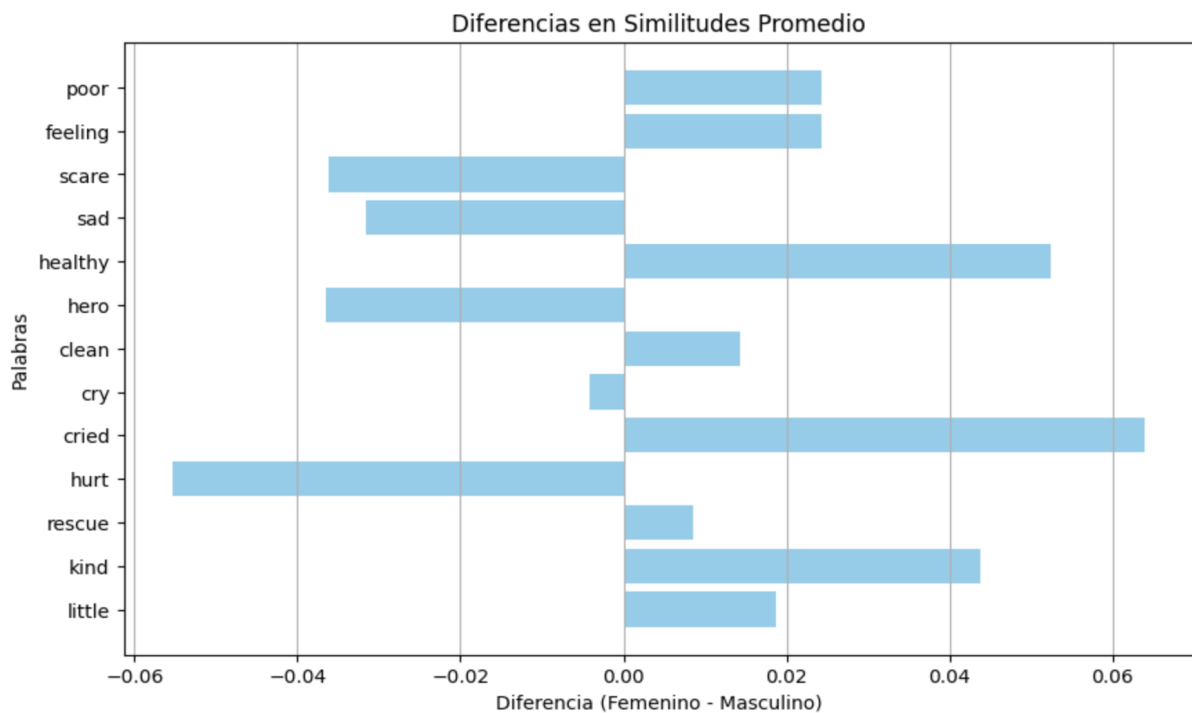
Como se mencionó anteriormente, valores negativos refieren a palabras más asociadas a los pronombres masculinos mientras que los indicadores positivos se presentan en términos más

relacionados al género femenino. Las palabras que obtuvieron una diferencia negativa fueron *hurt*, *cry*, *hero*, *sad* y *scare*; y que las que tuvieron un valor positivo fueron *little*, *rescue*, *feeling*, *healthy*, *kind*, *clean*, *poor* y *cried*.

De la misma manera que los resultados obtenidos en el cálculo de PPMI, se observa que *sad* y *hurt* obtuvieron una diferencia negativa lo que refleja una mayor asociación con los pronombres masculinos. No obstante, en los embeddings *cry* también parece estar más cercano a los atributos masculinos. El resto de las palabras que representan emociones que podrían estar vinculadas a estos sentimientos (*cried* y *feeling*) tuvieron diferencias positivas, es decir, en el corpus su tendencia es mayormente femenina.

Por último, la palabra *healthy* también mostró estar asociada con más fuerza al género femenino en los embeddings generados por el modelo.

Figura 20. Diferencias en las similitudes promedio - Palabras asociadas a las emociones, sentimientos y cuidado.

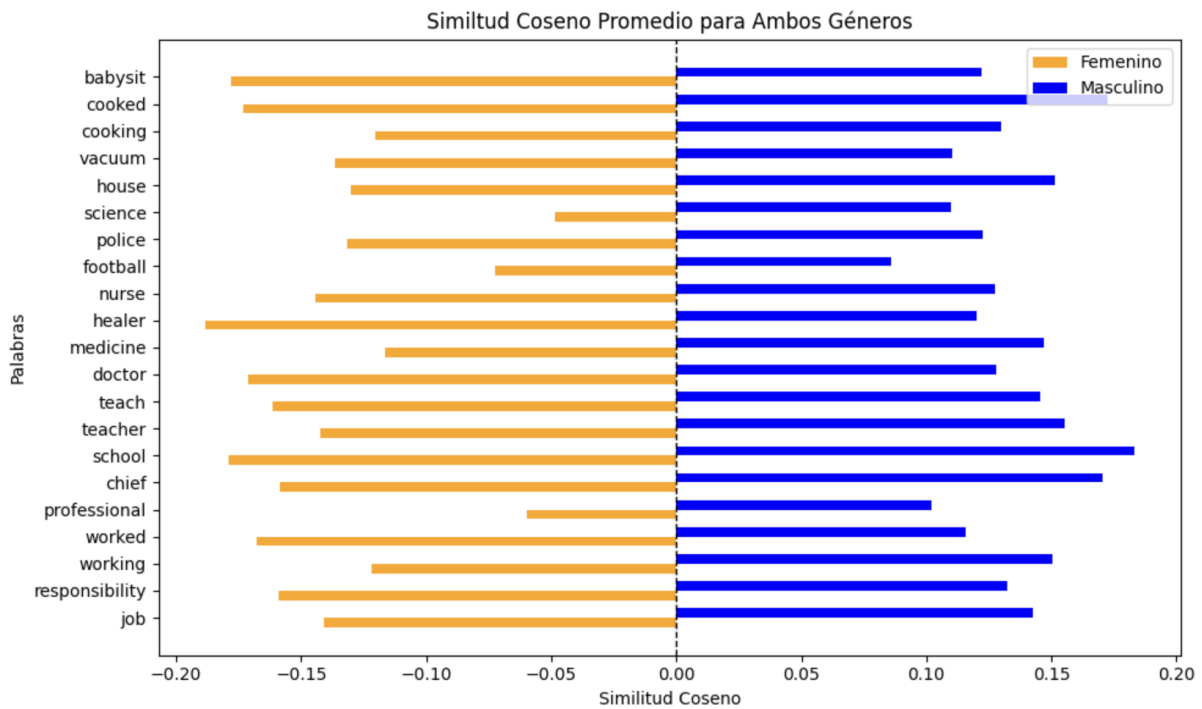


#### 4.3.1.2. Palabras asociadas al trabajo y a lo doméstico

Por otro lado, en la lista de palabras asociadas al trabajo y a lo doméstico, los promedios de la similitud coseno para ambos géneros son más dispares. La figura 19 muestra estos cálculos para cada uno de los segmentos.



Figura 21. Similitud coseno para palabras asociadas al trabajo y a lo doméstico.



En cuanto a las diferencias en las similitudes coseno (figura 20), los embeddings que se asociaron más al género masculino fueron *professional*, *football*, *chief*, *house*, *teacher*, *school*, *job*, *working*, *medicine*, *cooking* y *science*. Por otro lado, con el género femenino los términos fueron *vacuum*, *police*, *nurse*, *worked*, *cooked*, *teach*, *healer*, *doctor*, *responsibility* y *babysit*.

En primer lugar, al igual que lo observado en los resultados del PPMI, es interesante que no hay una vinculación más fuerte de profesiones médicas con los pronombres masculinos. Tanto las palabras *doctor*, como *nurse*, tuvieron una diferencia positiva. No obstante, la palabra *science* tuvo una diferencia negativa.

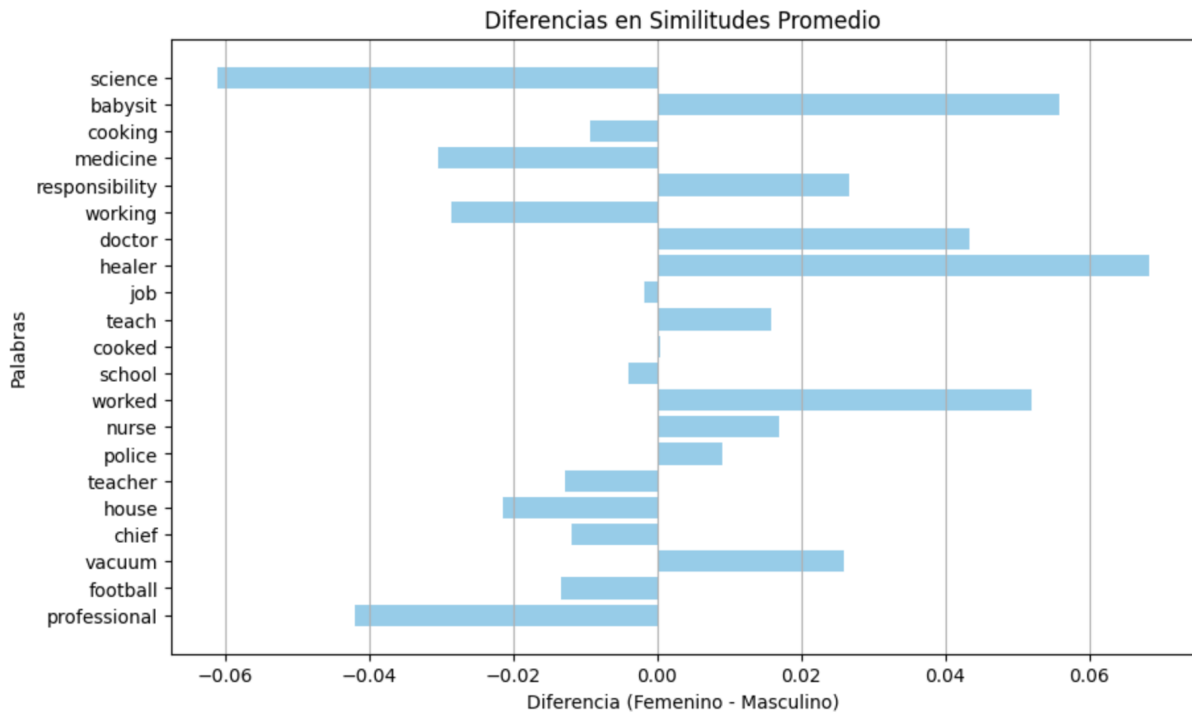
En segundo lugar, no hay una clara diferenciación en la asociación a trabajar entre los atributos para hombres y mujeres. Los términos *working*, *job* y *professional* tuvieron diferencias negativas pero la palabra *worked* positiva. No pareciera existir un sesgo acerca de que las mujeres trabajan menos.

En tercer lugar, se vuelve a hacer presente una mayor asociación de las tareas de cuidado y mantenimiento del hogar con el género femenino dado que *babysit*, *vacuum* y *cooked* obtuvieron diferencias positivas. Ahora bien, el embedding *cooking* pareciera estar más asociado a los atributos masculinos.

En cuarto lugar, a diferencia de los resultados obtenidos en el cálculo del PPMI, las palabras vinculadas a la enseñanza (*teach* y *teacher*) no muestran una relación más robusta con ninguno de los dos géneros.

Por último, la palabra *football* tuvo una diferencia negativa, se repite la misma tendencia que lo analizado a partir de PPMI, la noción de que los hombres están más relacionados con este deporte.

Figura 22. Diferencias en las similitudes promedio - Palabras asociadas al trabajo y a lo doméstico.



#### 4.3.2. Word Embedding Association Test (WEAT)

En última instancia, se aplicó el el Word Association Test (WEAT) (Schröder, S. et. al., 2021). Los conjuntos a contrastar se definieron como:

- A. Grupo de términos asociados al género masculino: *doctor, medicine, surgeon, science, engineer, technologist, technician, police y football.*
- B. Grupo de palabras vinculadas al género femenino: *house, cooking, cooked, vacuum, babysit, nurse, healer, teach y teacher.*

Se entiende que un valor positivo de effect sizes confirma que el segmento A está más asociado con los pronombres masculinos. Un indicador negativo indicaría sesgos en el sentido contrario.

El valor obtenido de effect sizes fue de 0.77. Según este resultado se comprobaría la hipótesis acerca de la relación más cercana de los hombres con carreras profesionales en el campo de la ciencia, medicina, deporte y seguridad mientras que las mujeres están más asociadas al trabajo doméstico, el cuidado y la enseñanza. No obstante, el indicador no resulta estadísticamente significativo con un p-value de 0.08. En consecuencia, no hay suficiente evidencia para rechazar la hipótesis nula y no se confirma una asociación significativa entre las variables analizadas.

## 5. Conclusiones

En esta tesis se plantea la pregunta ¿Son estos videos libres de estereotipos de género? ¿Se encuentran en ellos patrones y tendencias dispares en cuanto a género? Se parte de la premisa que los hombres suelen estar asociados con más fuerza a carreras profesionales mientras que las mujeres ocupan un espacio de cuidado, conectado con habilidades emocionales y sentimientos positivos. También que las últimas enfrentan expectativas más exigentes en cuanto a su salud y estándares de belleza.

Con el objetivo de responder estas preguntas se ha recopilado un corpus y se lo ha curado de manera de poder tratarlo de forma adecuada. Después se han aplicado diferentes métodos de exploración y se han obtenido resultados que indican la presencia de comportamientos no uniformes en relación a cómo son representados ambos géneros.

En cuanto al corpus, este se compone de archivos de texto que contienen los subtítulos de los videos a estudiar. A partir de estos se puede analizar el lenguaje oral representado como una transcripción. Se obtuvo mediante la API de YouTube y la librería YouTubeTranscript.Api. La primera se utilizó con el fin de recopilar identificadores de videos (IDs) de 40 canales populares para niños según Social Blade. La segunda herramienta permitió obtener los subtítulos de cada uno de los contenidos, que luego fueron almacenados en formato .txt, manteniendo en el nombre el canal y el id del video. El dataset se compuso entonces de 3.855 transcripciones.

En este dataset se realizaron tareas de curación para poder utilizar la información como input de los distintos métodos de exploración. En primer lugar, se tokenizaron los documentos dividiéndolos en unidades más pequeñas con la función *word\_tokenize* de la librería NLTK. En segundo lugar, se filtraron stopwords y contracciones aplicando el paquete de NLTK. Se adaptó la lista para eliminar de ella palabras relacionadas con el género. De la misma manera y se removieron tokens que consistían en una sola palabra, por ejemplo b o z. Por último, respondiendo a las particularidades del dataset se removieron palabras mal escritas o sin significado léxico y se eliminaron los términos que representan el 0.01% de mayor frecuencia. El corpus curado tuvo un volumen total de 1,446,526 tokens de los cuales 343,566 son únicos.

Se trabajó con distintos métodos de exploración en esta tesis, cada uno con asunciones diferentes pero todos con el objetivo de analizar la presencia de sesgos y tendencias asociadas. Se utilizó el Positive Pointwise Mutual Information (PPMI) que proporciona información sobre la asociación entre palabras. La métrica permite obtener una medida cuantitativa de cuán vinculados están dos términos al comparar la probabilidad de que ocurran juntos en un contexto dado, con la probabilidad esperada de que ocurran independientemente. Se asume que si el PPMI es mayor a cero existe una relación más fuerte de lo que se esperaría por azar. Para mejorar la robustez de los resultados, y considerando que el indicador suele tener problemas con conteos de ocurrencias bajos, solo se consideraron los tokens que aparecen más de 10 veces en cada documento en la aplicación de este método. Así también, con el mismo propósito, se aplicó con el suavizado de Laplace. A partir del PPMI se analizaron las asociaciones entre las listas de palabras definidas y los pronombres de género y se evaluó la hipótesis planteada en base a los resultados de EDIA. Adicionalmente, se revisó la presencia de negaciones en los contextos para estudiar si generaban un impacto en los insights obtenidos.

Se infirió un modelo de Latent-Dirichlet-Allocation (LDA) que permite representar los documentos de un corpus en forma de mezclas aleatorias de temas latentes, entendiendo cada uno como una distribución de términos. Esta técnica de aprendizaje no supervisado posibilita descubrir tópicos y analizar si estos se pueden asociar a estereotipos de género.

Se trabajó con el modelo de embeddings Word2Vec. Este método infiere representaciones de palabras como vectores mediante el entrenamiento de un clasificador con la tarea de predecir si es probable que un término  $x$  aparezca cerca del término  $y$ . Asume que dos palabras tendrán una representación vectorial semejante si son similares en significado y por ende permite descubrir relaciones entre términos dentro del corpus y calcular la similitud. Con este fin se utilizaron dos métricas. Por un lado, la similitud coseno que calcula el coseno del ángulo de los vectores y tiende a ser alta cuando apuntan en la misma dirección. En este trabajo se computó la métrica entre las listas de palabras y atributos de género para luego estudiar las diferencias de ambos resultados y detectar tendencias. Por otro lado, se utilizó el Word Association Test (WEAT) que compara dos conjuntos de palabras objetivo con dos grupos de atributos de género de igual tamaño y se basa en la hipótesis de que uno de los dos está más asociado a un género que al otro. A partir de ello, se computaron los effect sizes para medir las diferencias de asociación y se calculó la significancia estadística del resultado.

A continuación se describen los resultados obtenidos. En líneas generales, cómo se mencionó previamente, se ha identificado la presencia de tendencias heterogéneas en relación a cómo son representados ambos géneros. En primer lugar, se detectaron diferencias de género en cuanto a los sentimientos y la expresión de los mismos en el corpus estudiado. Las palabras que representan los sentimientos, *sad* y *scare*, tuvieron mayores valores de PPMI con la palabra que representa el género masculino, no obstante la expresión de estos sentimientos, palabras como *cry*, *cried* y *feeling*, tuvieron mayores valores de PPMI con la palabra que representa al género femenino. Esta misma tendencia se repite de cierta manera en los resultados de la similitud coseno con una similitud más alta entre *sad* y los atributos masculinos pero una relación más fuerte entre los femeninos y *cried* o *feeling*. Relacionado con este punto, la noción de héroe pareciera estar más asociada al género masculino mientras que *rescue* dió un valor de PPMI más elevado para *she* y se encontraron contextos en el corpus en los que las mujeres se representan como valientes y virtuosas. Esto indica que en principio el rol heroico está reservado para los hombres a pesar de que las mujeres también tienen representaciones ligadas a las características propias de un héroe.

En segundo lugar, tanto en el PPMI como en el análisis a partir de la similitud coseno la palabra *healthy* tuvo mayor asociación con los atributos femeninos. Para comprender mejor la tendencia se analizaron los contextos de ocurrencias conjuntas permitiendo descubrir dos aristas de la relación. Por un lado, los personajes femeninos en los videos analizados son representadas en un lugar de cuidado hacia otros con más frecuencia que los hombres. Por otro lado, la presencia de presiones sobre el cuidado físico propio y las expectativas hegemónicas también está presente en el corpus de esta tesis.

En tercer lugar, respecto a las palabras asociadas al trabajo y a lo doméstico se detectaron diferencias interesantes. Si bien no parecen existir sesgos acerca de las mujeres y el trabajo o los campos médicos, tanto en el PPMI como en la similitud coseno la palabra *ciencia* presentó una mayor relación con el género masculino. Ahora bien, las tareas de la casa están fuertemente relacionadas con las mujeres. Todas las palabras asociadas a ello dieron resultados mayores de PPMI con *she* y en la similitud coseno

hubo una tendencia similar pero más equilibrada dado que el término *cooking* tuvo una mayor similitud con los atributos masculinos.

Por último, los tópicos generados por el modelo Latent-Dirichlet-Allocation no evidenciaron tendencias estereotipadas y los resultados del Word Association Test no resultaron estadísticamente significativos por lo que no fue posible obtener conclusiones de estas dos metodologías. Por otro lado, es importante destacar que en la mayor parte de las metodologías las diferencias fueron chicas. Esto probablemente responde al tamaño del corpus disponible para este trabajo y en consecuencia la imposibilidad de las herramientas para diferenciar ampliamente tendencias.

En conclusión, las metodologías utilizadas y sus resultados permiten mostrar que las representaciones de los géneros en los videos de YouTube para niños en inglés analizados en esta tesis no son homogéneas sino que en algunos casos reproducen patrones estereotipados. Es fundamental entonces trabajar en minimizar la presencia de estereotipos con el objetivo de no reproducir creencias nocivas que moldean el pensar y accionar de las nuevas generaciones. Para lograrlo se recomienda abordar un enfoque multidisciplinario. Por un lado, apoyar la generación de contenidos en técnicas de identificación de sesgos, similares a las utilizadas en este trabajo. Esto permitirá entender qué tendencias se evidencian y mantienen a lo largo del tiempo y comprender entonces las áreas susceptibles de mejora. Por otro lado, incorporar perspectiva de género en las producciones, fomentando equipos de trabajos diversos y participación de mujeres. Además es fundamental contar con el respaldo de expertos en áreas de sociología y psicología que permitan guiar las líneas discursivas de acuerdo a lo adecuado para las edades pero con el foco de no reproducir estereotipos considerando las particularidades del contexto.

En el ámbito de negocios los hallazgos de esta tesis pueden utilizarse como disparador para generar contenidos que no reproduzcan narrativas sesgadas sobre los géneros, sino que incorporen una representación más equitativa. Esto permitiría a las empresas alcanzar potencialmente un público más diverso y reforzar su responsabilidad social corporativa. Así también, la metodología propuesta en este estudio podría ser empleada por empresas de consultoría y auditoría para ofrecer servicios de análisis de los contenidos en medios audiovisuales.

## 6. Futuras investigaciones

El análisis de estereotipos de género en el lenguaje utilizado en videos de Youtube es aún un camino extenso por recorrer. En este trabajo las metodologías aplicadas se limitaron para circunscribir el alcance y de acuerdo al corpus disponible pero existen diversos análisis que serían interesantes realizar.

En primer lugar, los distintos idiomas varían en su composición gramatical. En el español, a diferencia del inglés, las palabras suelen estar asociadas al género con mucha más frecuencia. Por ejemplo, la palabra *teacher* en inglés es utilizada indistintamente para hombres y mujeres. A diferencia, en español se utiliza “maestra” y “maestro” u “profesora”/“profesor”. Esta distinción de género en la gramática representa un desafío interesante de abordar, tanto por las diferencias culturales y de asociación implícitas en el lenguaje como desde el punto de vista metodológico y computacional.

Entonces se plantea la pregunta ¿Cómo varían los resultados obtenidos en este trabajo si como input se utilizan videos en español? ¿Se mantienen las mismas tendencias?

En segundo lugar, en la aplicación de metodologías como el Positive Pointwise Mutual Information se consideraron solo dos atributos de género ¿Los resultados mantendrían la misma tendencia de calcular la métrica a partir de un set de atributos de género ampliado? ¿Cómo varían los porcentajes de negaciones al ampliar las muestras?

En tercer lugar, en base a los resultados obtenidos en este trabajo hay algunas relaciones entre palabras que podrían ser investigadas con mayor detalle. Por ejemplo, se observó que la palabra *cooking* tuvo una mayor similitud coseno con las palabras que representan al género masculino, mientras que la mayor parte de los términos asociados a lo doméstico tuvieron tanto un PPMI como una similitud coseno mayor con las palabras que representan al género femenino ¿Qué sucede entonces con *cooking*? ¿Cómo son y qué información aportan los contextos de ocurrencias conjuntas con cada género?

En cuarto lugar, se propone ampliar el corpus incorporando más canales infantiles y estudiando las variaciones en los resultados entre ellos y su metadata ¿Qué relaciones se detectan entre los creadores y los resultados? ¿Hay alguna tendencia de acuerdo al volumen de suscriptores? En la misma línea, se sugiere expandir la lista de palabras con el objetivo de analizar otras asociaciones y tendencias culturales. Así también, sería interesante comprender si los resultados de este trabajo son similares a aquellos que puedan obtenerse construyendo el corpus a partir de videos para adultos.

En quinto lugar, existen varios estudios cuya finalidad es entender las variaciones y la evolución en los sesgos contenidos en el lenguaje a través del tiempo. Sería interesante contrastar los resultados obtenidos en este trabajo con aquellos que se generen a partir de contenidos más antiguos. Por ejemplo, comparar los contenidos para niños de los 80s y 90s con contenidos actuales y entender qué diferencias se presentan y si existió o no una evolución hacia la paridad de género.

En sexto lugar, el universo de embeddings es amplio. Actualmente existen grandes modelos entrenados a partir de diversos corpus que tienen en ellos diferencias culturales y metodológicas. Resulta interesante en consecuencia plantear la interrogante ¿Cuáles son las similitudes y diferencias entre los resultados obtenidos en este trabajo y otros modelos? En esta línea es importante comprender que la construcción de modelos de embeddings utiliza muestreos aleatorios. Este factor ocasiona que cada modelo resulte único en sus resultados. Para robustecer la confianza en los descubrimientos se podrían entrenar distintos modelos utilizando subsets de muestras y luego calcular la desviación estándar de los resultados.

Por último, para entender el impacto de los estereotipos de género en las métricas de visualizaciones y engagement (likes, suscripciones al canal, comentarios) se podría generar un estudio que compare estos KPIs con un scoring de neutralidad en los videos. Este puede construirse a partir de diversos puntajes ponderados para cada una de las metodologías aplicadas en esta tesis y asignarlo a cada video. A partir de ello, se puede analizar si existe una correlación entre mayor neutralidad y mayor volumen de visualizaciones y engagement.

## 7. Referencias

1. Alonso Alemany, L., Benotti, L., Maina, H., Gonzalez, L., Martínez, L., Busaniche, B., Halvorsen, A., Rojo, A., & Rajngewerc, M. (2023). *Bias assessment for experts in discrimination, not in computer science*. In L. Alonso Alemany, L. Benotti, H. Maina, L. Gonzalez, L. Martínez, B. Busaniche, A. Halvorsen, A. Rojo, & M. Rajngewerc (Eds.), *Proceedings of the First Workshop on Cross-Cultural Considerations in NLP (C3NLP)* (pp. 91–106). Dubrovnik, Croatia: Association for Computational Linguistics.
2. Alonso Alemany, L., Benotti, L., Maina, H., González, L., Rajngewerc, M., Martínez, L., Sánchez, J. Schilman, M., Ivetta, G., Halvorse, A., Mata Rojo, A., Bordone, M. y Busaniche, B. (2023). *A methodology to characterize bias and harmful stereotypes in natural language processing in Latin America*. <https://doi.org/10.48550/arXiv.2207.06591>
3. Antoniak, M., & Mimno, D. (2021). *Bad Seeds: Evaluating Lexical Methods for Bias Measurement*. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)* (pp. 1889–1904). Online: Association for Computational Linguistics.
4. Bandura, A., Ross, D. y Ross, S.A. (1963). *Imitation of film-mediated aggressive models*. *The Journal of Abnormal and Social Psychology*, 66(1), 3-11.
5. Bigler, R. S., & Liben, L. S. (2007). *Developmental Intergroup Theory: Explaining and Reducing Children's Social Stereotyping and Prejudice*. *Current Directions in Psychological Science*, 16(3), 162–166. <http://www.jstor.org/stable/20183186>
6. Bird S., Klein E., Loper E. (2009). *Natural Language Processing with Python*. O'Reilly Media. <https://www.nltk.org/book/>
7. Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). *Latent Dirichlet Allocation*. *Journal of Machine Learning Research*, 3, 993-1022.
8. Bolukbasi, T., Kai-Wei C., James Z., Venkatesh S. y Adam K. (2016). *Man Is to Computer Programmer as Woman Is to Homemaker? Debiasing Word Embeddings*. *Adv. Neural Inf. Process. Syst.* 4349–57.
9. Bosson, J. K, Vandello, J. A. y Buckner, C. E. (2019). *The psychology of sex and gender*. Thousand Oaks, CA: SAGE Publications.
10. Boutyline, A., Arseniev-Koehler, A., & Cornell, D. J. (2023). *School, Studying, and Smarts: Gender Stereotypes and Education Across 80 Years of American Print Media, 1930–2009*. *Social Forces*, 102(1). Oxford University Press. Recuperado el 27 de noviembre de 2023 de <https://muse-jhu-edu.ezproxy.utdt.edu/article/906534>
11. Casey, K., Novick, K. y Lourenco, S. F. (2021). *Sixty years of gender representation in children's books: Conditions associated with overrepresentation of male versus female protagonists*. *PLOS ONE*. <https://doi.org/10.1371/journal.pone.0260566>
12. Church, K. W., & Hanks, P. (1990). *Word Association Norms, Mutual Information, and Lexicography*. *Computational Linguistics*, 16(1), 22–29.



13. Dixon, S. J. (2023). *Most popular social networks worldwide as of October 2023, ranked by number of monthly active users*. Statista. Recuperado el 17 de noviembre 2023 de <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>
14. Döring, N., & Mohseni, M. R. (2020). *Gendered hate speech in YouTube and YouNow comments: Results of two content analyses*. *SCM Studies in Communication and Media*, 9(1), 62-88. Recuperado el 24 de noviembre de 2023 de <https://www.nomos-elibrary.de/10.5771/2192-4007-2020-1-62.pdf>
15. Fiske, S. T., Amy, J. C. Cuddy, Glick, P. y Jun, X. (2002). *A Model of (Often Mixed) Stereotype Content: Competence and Warmth Respectively Follow From Perceive Status and Competition*. *Journal of Personality and Social Psychology*, 82(6): 878-902. Recuperado el día 26 de noviembre 2023 de [https://cos.gatech.edu/facultyres/Diversity\\_Studies/Fiske\\_StereotypeContent.pdf](https://cos.gatech.edu/facultyres/Diversity_Studies/Fiske_StereotypeContent.pdf)
16. Galvez R., Tiffenberg V., Altszyler E. (2019). *Half a Century of Stereotyping Associations Between Gender and Intellectual Ability in Films*. *Sex Roles* 81, 643–654 (2019). <https://doi.org/10.1007/s11199-019-01019-x>
17. Garg, N., Schiebinger, L., Jurafsky, D. y Zou, J. (2018). *Word embeddings quantify 100 years of gender and ethnic stereotypes*. *Proceedings of the National Academy of Sciences*, 115(16), E3635-E3644. <https://doi.org/10.1073/pnas.1720347115>
18. Geena Davis Institute on Gender in Media. (2015). *The reel truth: Women aren't seen or heard. An automated analysis of gender representation in popular films*. <https://seejane.org/wp-content/uploads/gdiq-reel-truth-women-arent-seen-or-heard-automated-analysis.pdf>. Recuperado el día 24 de octubre de 2023.
19. Grabe, S., Ward, L. M., & Hyde, J. S. (2008). *The role of the media in body image concerns among women: A meta-analysis of experimental and correlational studies*. *Psychological Bulletin*.
20. Haddock, C., Rindskopf, D. y Shadish, W. (1998). *Using Odds Ratios as Effect Sizes for Meta-Analysis of Dichotomous Data: A primer on methods and issues*. *Psychological Methods*. 3. 339-353. Recuperado el 26 de noviembre de 2023 de [https://www.researchgate.net/publication/232562571\\_Using\\_Odds\\_Ratios\\_as\\_Effect\\_Sizes\\_for\\_Meta-Analysis\\_of\\_Dichotomous\\_Data\\_A\\_Primer\\_on\\_Methods\\_and\\_Issues](https://www.researchgate.net/publication/232562571_Using_Odds_Ratios_as_Effect_Sizes_for_Meta-Analysis_of_Dichotomous_Data_A_Primer_on_Methods_and_Issues)
21. Jurafsky D. y Martin J. H. (2023). *Speech and language processing: An introduction to natural language processing, computational linguistics and speech recognition (3rd ed.)*. Upper Saddle River: Prentice Hall. Recuperado el día 26 de noviembre 2023 de <https://web.stanford.edu/~jurafsky/slp3/ed3book.pdf>
22. Lewis, M., & Lupyan, G. (2020). *Gender stereotypes are reflected in the distributional structure of 25 languages*. *Nature Human Behaviour*, 4(10), 1021–1028. <https://www.nature.com/articles/s41562-020-0918-6>
23. Lewis, M., Cooper Borkenhagen, M., Converse, E., Lupyan, G., & Seidenberg, M. S. (2022). *What Might Books Be Teaching Young Children About Gender?* *Psychological Science*, 33(1), 33-47. <https://doi.org/10.1177/09567976211024643>



24. Lopez Yse D. (2021). *Text Normalization for Natural Language Processing (NLP)*. Recuperado el día 26 de noviembre 2023 de <https://towardsdatascience.com/text-normalization-for-natural-language-processing-nlp-70a314bfa646>
25. Manning, C. D., & Schütze, H. (1999). *Foundations of statistical natural language processing*. Massachusetts Institute of Technology.
26. Mikolov, T., Sutskever, I., Chen, K., Corrado, G., y Dean, J. (2013). *Distributed Representations of Words and Phrases and their Compositionality*. <https://doi.org/10.48550/arXiv.1310.4546>
27. Miller, D. I., Nolla, K. M., Eagly, A. G., y Uttal, D. H. (2018). *The Development of Children's Gender-Science Stereotypes: A Meta-analysis of 5 Decades of U.S. Draw-A-Scientist Studies*. *Child Development*, 89(6), 1943-1955. <https://doi.org/10.1111/cdev.13039>
28. Nicola Döring & M. Rohangis Mohseni (2018): *Male dominance and sexism on YouTube: results of three content analyses*. *Feminist Media Studies*, DOI: 10.1080/14680777.2018.1467945. Recuperado el 24 de noviembre de 2023 de [https://www.nicola-doering.de/wp-content/uploads/2018/06/D%C3%B6ring\\_Mohseni\\_Male\\_Dominance\\_Sexism\\_YouTube\\_2018.pdf](https://www.nicola-doering.de/wp-content/uploads/2018/06/D%C3%B6ring_Mohseni_Male_Dominance_Sexism_YouTube_2018.pdf)
29. Olivier J. y Bell M. L. (2013). *Effect Sizes for 2x2 Contingency Tables*. *PLoS ONE* 8(3): e58777. Recuperado el 26 de noviembre de 2023 de <https://doi.org/10.1371/journal.pone.0058777>
30. Python Software Foundation. (2023). *Regular expressions operations*. Python Documentation. Recuperado el día 26 de noviembre 2023 de <https://docs.python.org/3/library/re.html>
31. Radim Řehůřek. (27 de marzo de 2024). *Tutorials - gensim documentation*. Recuperado de [https://radimrehurek.com/gensim/auto\\_examples/tutorials/run\\_word2vec.html](https://radimrehurek.com/gensim/auto_examples/tutorials/run_word2vec.html)
32. Rao, P., & Taboada, M. (2021). *Gender Bias in the News: A Scalable Topic Modelling and Visualization Framework*. *Frontiers in Artificial Intelligence, Section Language and Computation*, 4(2021). <https://doi.org/10.3389/frai.2021.664737>
33. Regueira U., Alonso Ferreiro A., Da-Vila S. (2020). *Women on YouTube: Representation and participation through the Web Scraping technique*. *Comunicar*. 63.
34. Scharrer, E., Warren, S. (2022). *Adolescents' Modern Media Use and Beliefs About Masculine Gender Roles and Norms*. *Journalism & Mass Communication Quarterly*, 99(1), 289–315.
35. Schröder, S., Schulz, A., Kenneweg, P., Feldhans, R., Hinder, F., & Hammer, B. (2021). *Evaluating Metrics for Bias in Word Embeddings*. Recuperado de <https://doi.org/10.48550/arXiv.2111.07864>
36. Smith, S. L., Choueiti, M., Prescott, A., y Pieper, K. (2012). *Gender roles & occupations: A look at character attributes and job-related aspirations in film and television*. *Geena Davis Institute on Gender in Media*, 1-46.
37. STEM Women. (2023). *Women in STEM: Statistics, progress, and challenges*. Recuperado el día 1 de mayo de 2024 de <https://www.stemwomen.com/women-in-stem-statistics-progress-and-challenges>
38. *Top 100 YouTubers Made-For-Kids Channels*. Recuperado el día 25 de noviembre 2023 de <https://socialblade.com/youtube/top/category/made-for-kids>

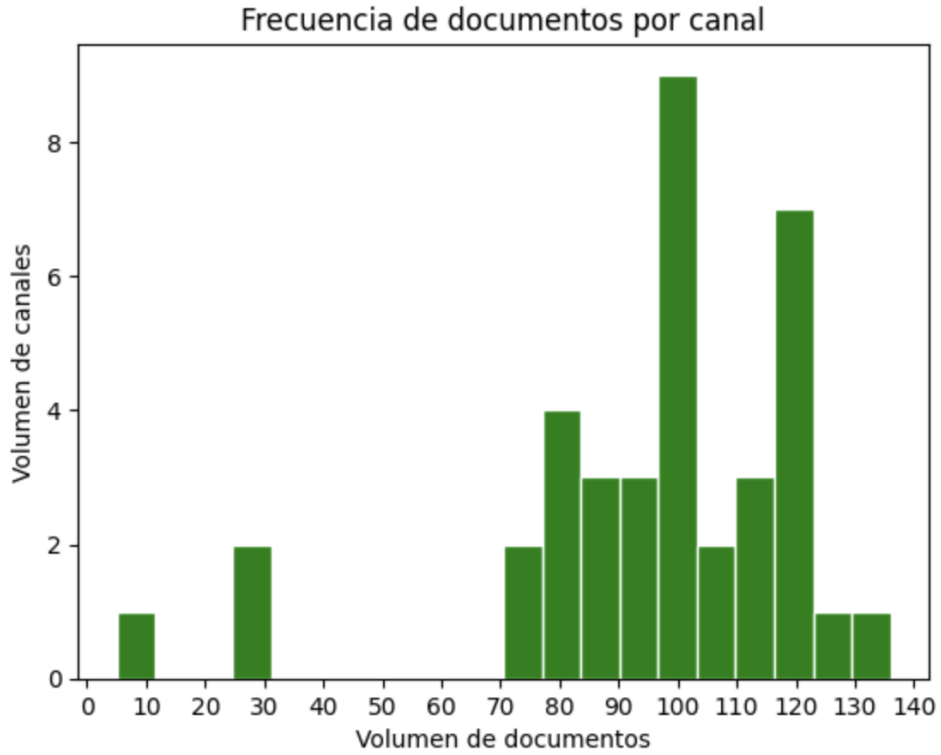
## 8. Apéndice

### 8.1. Lista de canales consultados y número de suscriptores

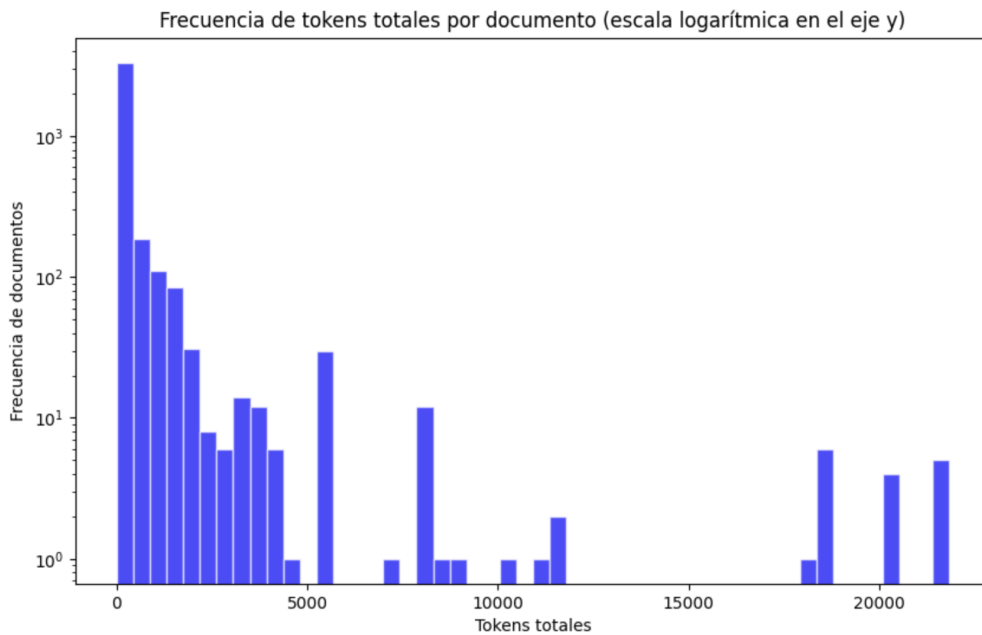
| Canal |                      | Suscriptores | Canal |                      | Suscriptores |
|-------|----------------------|--------------|-------|----------------------|--------------|
| 1     | KidsDianaShow        | 114 M        | 21    | Bebefinn             | 13.8 M       |
| 2     | LikeNastyaofficial   | 108 M        | 22    | LittleBabyBum        | 12.5 M       |
| 3     | ChuChuTV             | 68 M         | 23    | Pinkfong             | 11.7 M       |
| 4     | LooLooKids           | 54.9 M       | 24    | Supercrazykids       | 11.5 M       |
| 5     | Toysandcolors        | 44.5 M       | 25    | Mother Goose Club    | 9.29 M       |
| 6     | Masha and The Bear   | 44.2 M       | 26    | Videogyan Kids Shows | 9.28 M       |
| 7     | GenevievesPlayhouse  | 35.9 M       | 27    | CoComelonAnimalTime  | 9.17 M       |
| 8     | CyberVillageSolution | 34.2 M       | 28    | KedooToonsTV         | 7.12 M       |
| 9     | PeppaPigOfficial     | 32.5 M       | 29    | WildBrainBananas     | 7.1 M        |
| 10    | LittleAngel          | 30.5 M       | 30    | KidsChannel          | 6.67 M       |
| 11    | Bouncepatrol         | 28.1 M       | 31    | Msrachel             | 6.61 M       |
| 12    | Nickjr               | 27.8 M       | 32    | MyLittlePonyOfficial | 5.01 M       |
| 13    | Dbillions            | 27.4 M       | 33    | CoComelonCodyTime    | 4.29 M       |
| 14    | SesameStreet         | 23.5 M       | 34    | ChuChuTVStorytime    | 4.2 M        |
| 15    | VaniaManiaKids       | 23.4 M       | 35    | BlueyOfficialChannel | 3.96 M       |
| 16    | VladandNiki          | 23.1 M       | 36    | Ryan's World         | 3.61 M       |
| 17    | Blippi               | 18.7 M       | 37    | Graciescorner        | 2.54 M       |
| 18    | Disneyjunior         | 17 M         | 38    | PBSKIDS              | 2.15 M       |
| 19    | EliKids              | 15.7 M       | 39    | WildBrainZoo         | 2.11 M       |
| 20    | NetflixJr            | 15.2 M       | 40    | NinasFamilia         | 745 k        |

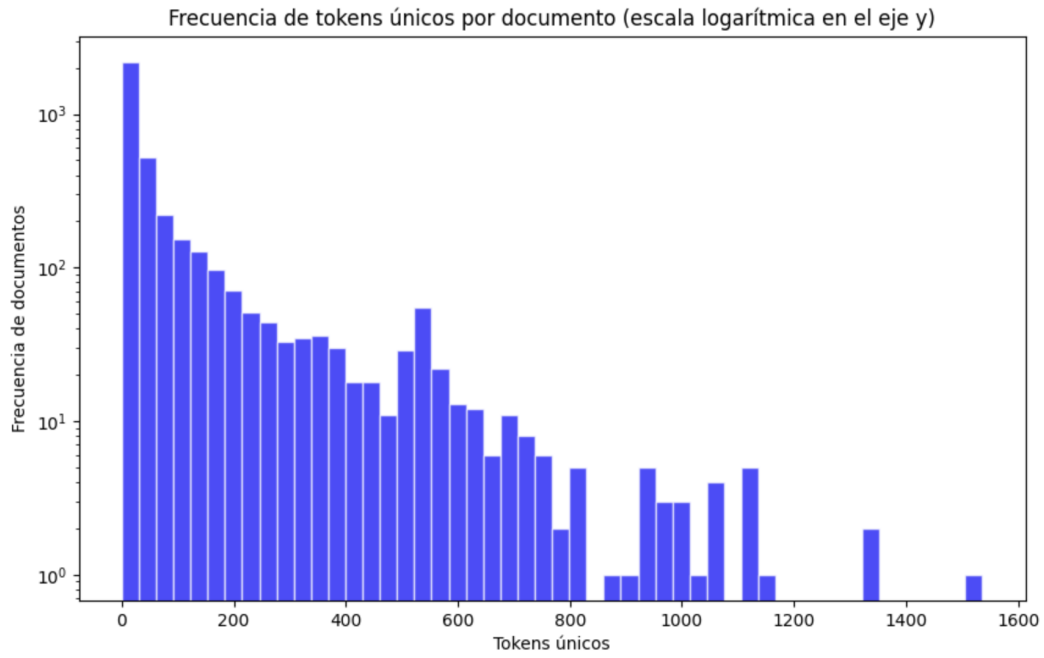
## 8.2. Archivos de texto por canal consultado

| Canal |                      | Archivos txt | Canal |                      | Archivos txt |
|-------|----------------------|--------------|-------|----------------------|--------------|
| 1     | KidsDianaShow        | 101          | 21    | Bebefinn             | 77           |
| 2     | LikeNastyaofficial   | 93           | 22    | LittleBabyBum        | 81           |
| 3     | ChuChuTV             | 27           | 23    | Pinkfong             | 100          |
| 4     | LooLooKids           | 82           | 24    | Supercrazykids       | 82           |
| 5     | Toysandcolors        | 98           | 25    | Mother Goose Club    | 108          |
| 6     | Masha and The Bear   | 117          | 26    | Videogyan Kids Shows | 102          |
| 7     | GenevievesPlayhouse  | 119          | 27    | CoComelonAnimalTime  | 76           |
| 8     | CyberVillageSolution | 85           | 28    | KedooToonsTV         | 90           |
| 9     | PeppaPigOfficial     | 103          | 29    | WildBrainBananas     | 93           |
| 10    | LittleAngel          | 100          | 30    | KidsChannel          | 5            |
| 11    | Bouncepatrol         | 124          | 31    | MsRachel             | 106          |
| 12    | Nickjr               | 110          | 32    | MyLittlePonyOfficial | 113          |
| 13    | Dbillions            | 29           | 33    | CoComelonCodyTime    | 117          |
| 14    | SesameStreet         | 112          | 34    | ChuChuTVStorytime    | 98           |
| 15    | VaniaManiaKids       | 136          | 35    | BlueyOfficialChannel | 108          |
| 16    | VladandNiki          | 122          | 36    | Ryan's World         | 117          |
| 17    | Blippi               | 102          | 37    | Graciescorner        | 95           |
| 18    | Disneyjunior         | 103          | 38    | PBSKIDS              | 119          |
| 19    | EliKids              | 81           | 39    | WildBrainZoo         | 119          |
| 20    | NetflixJr            | 88           | 40    | NinasFamilia         | 117          |



### 8.3. Palabras totales y tokens únicos





#### 8.4. 3 palabras más comunes en cada documento



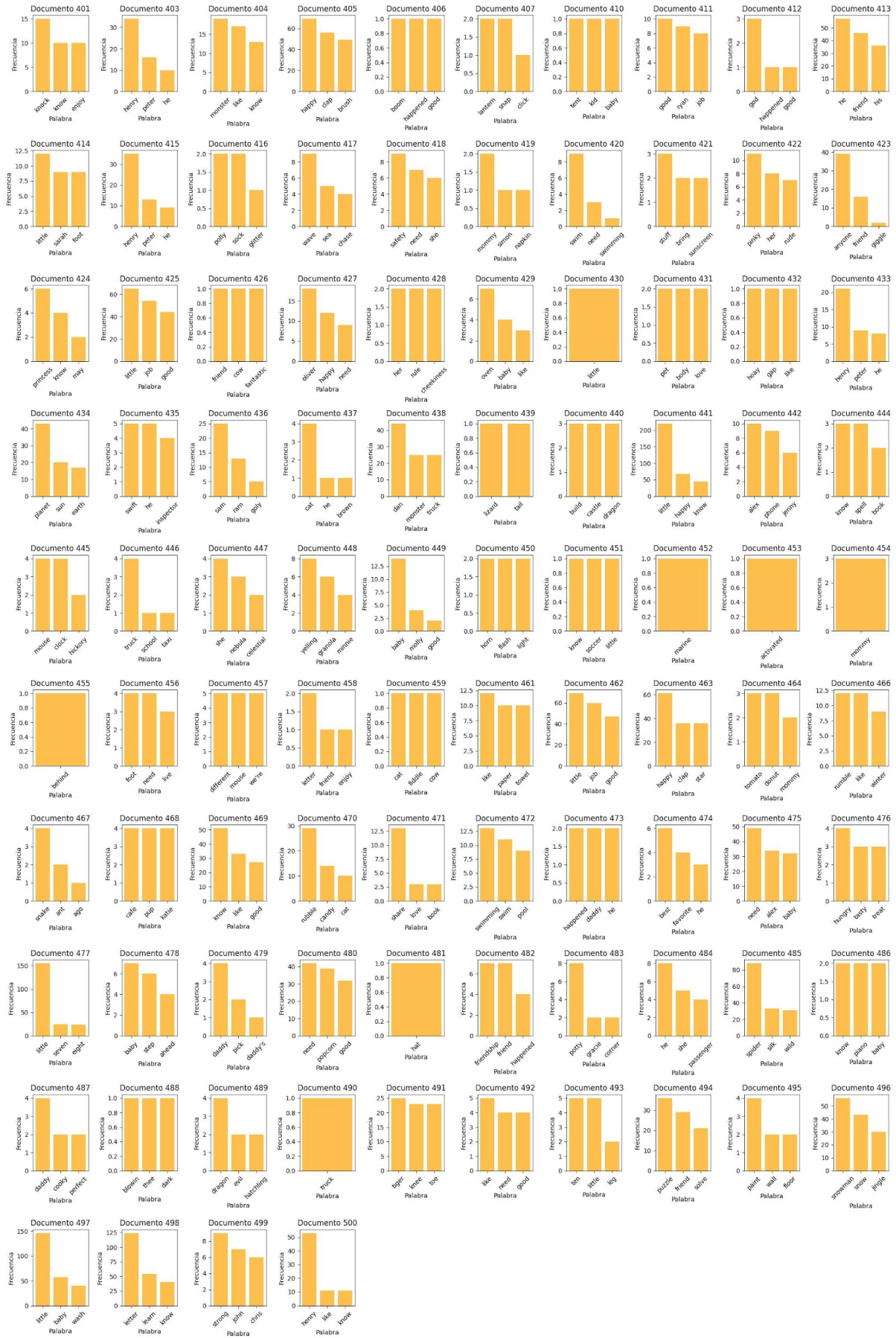














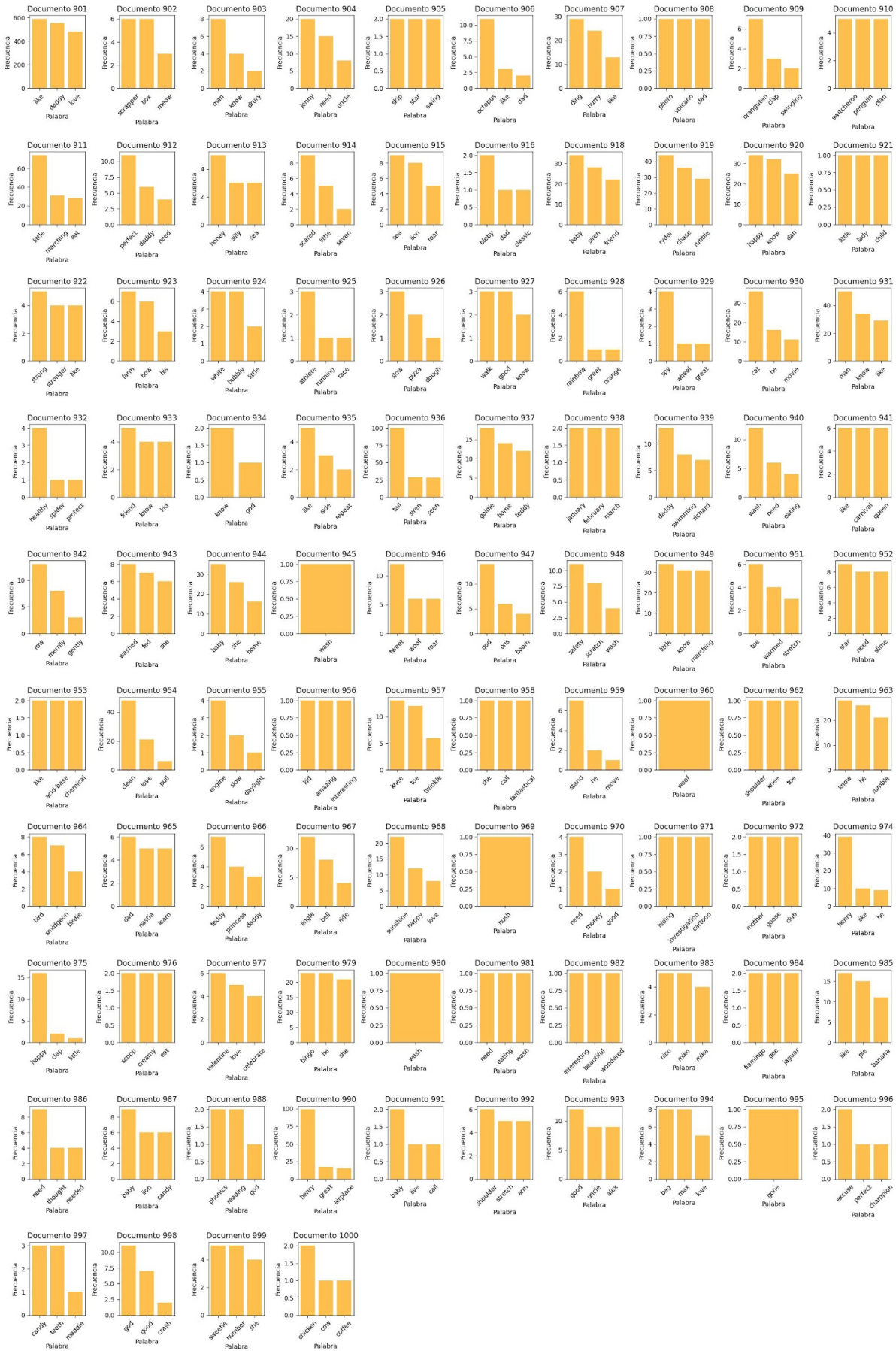








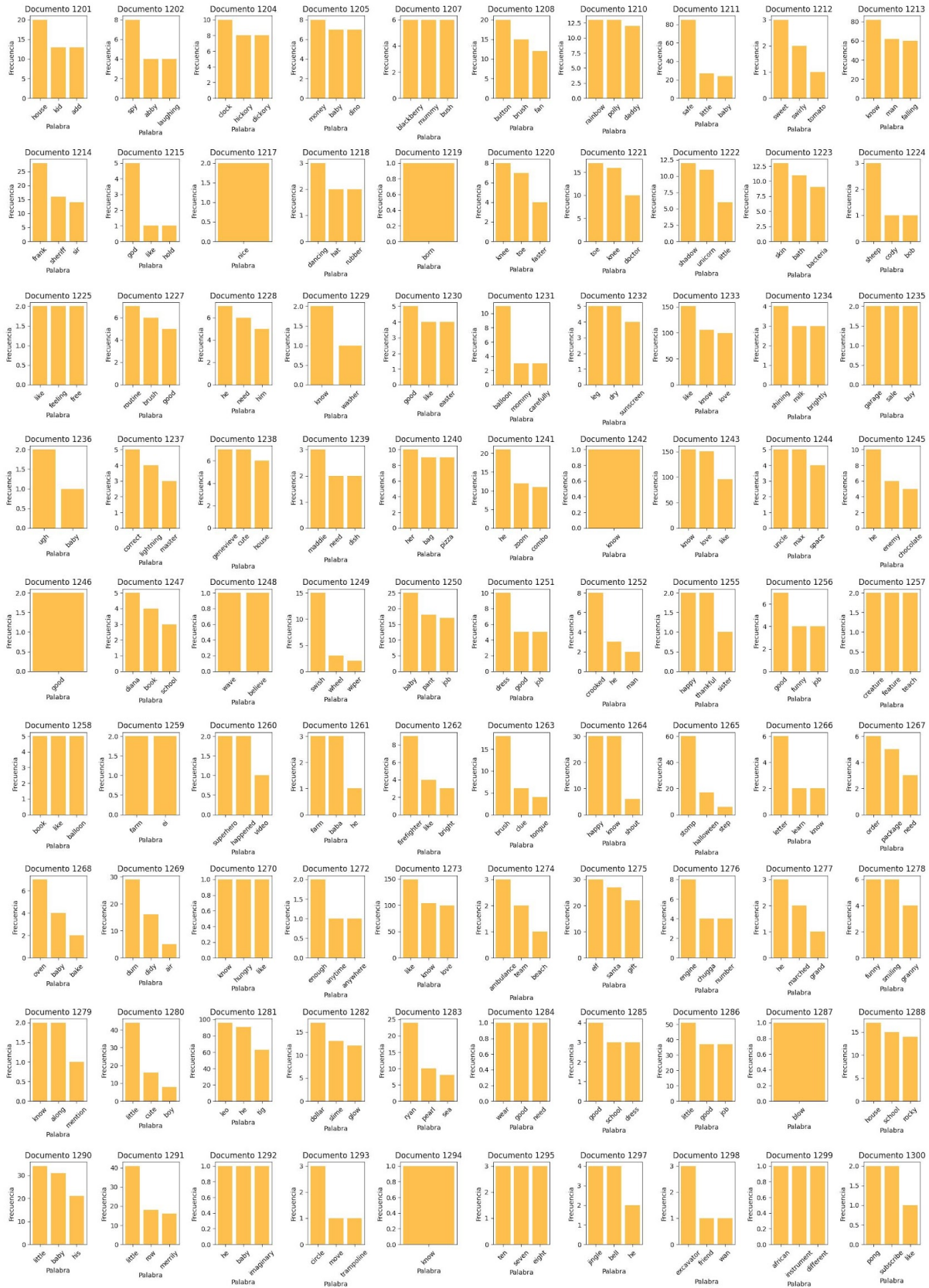


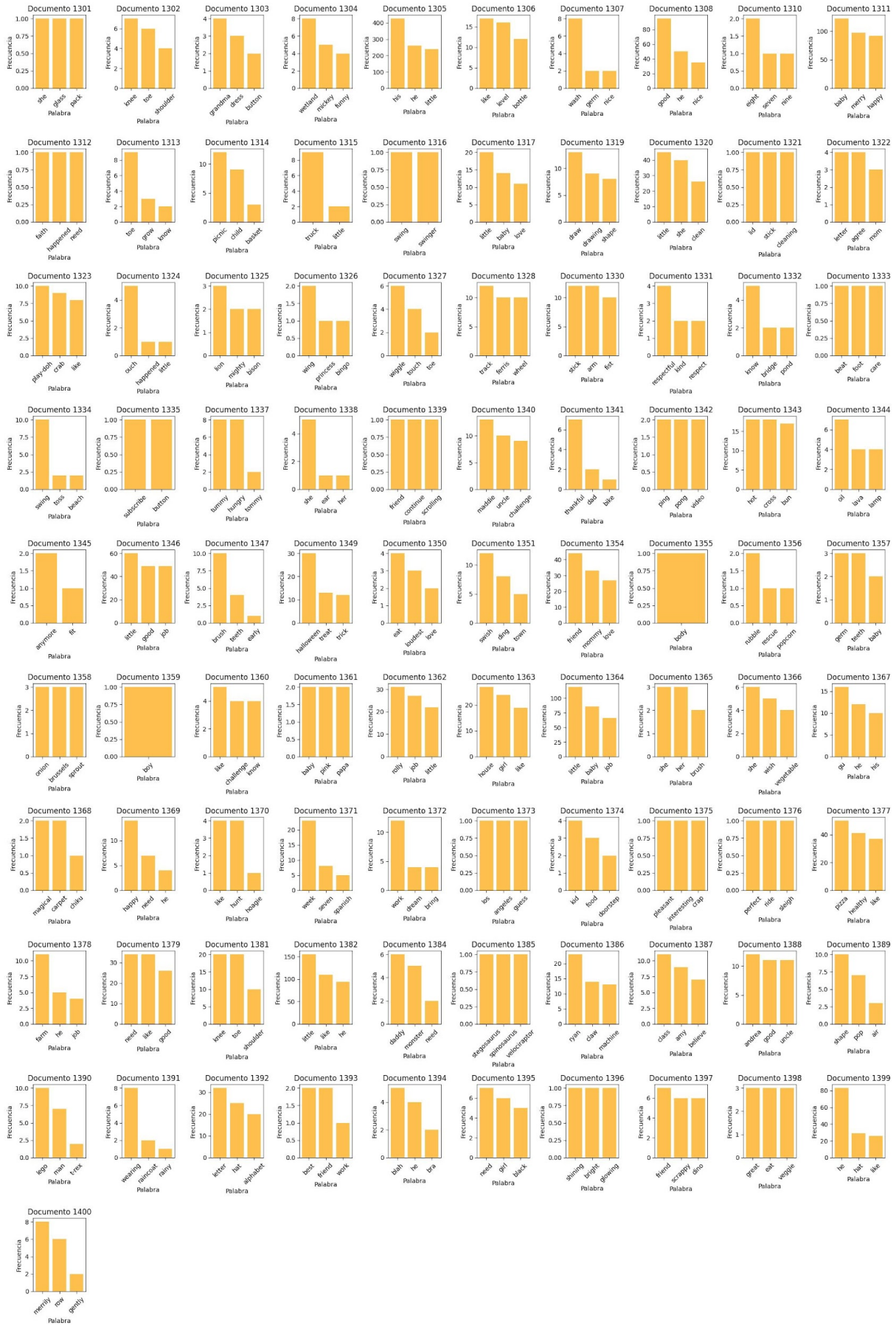












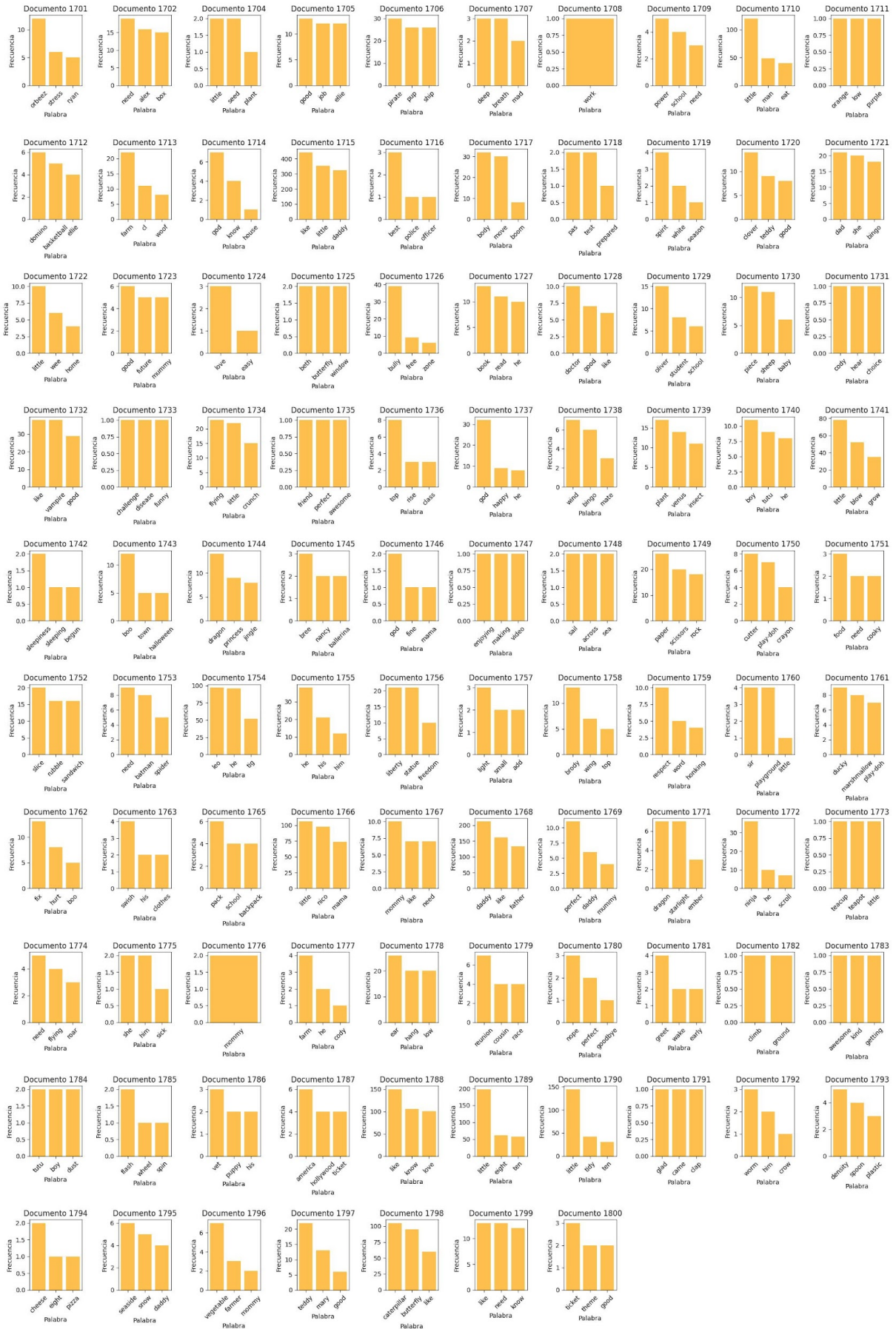
















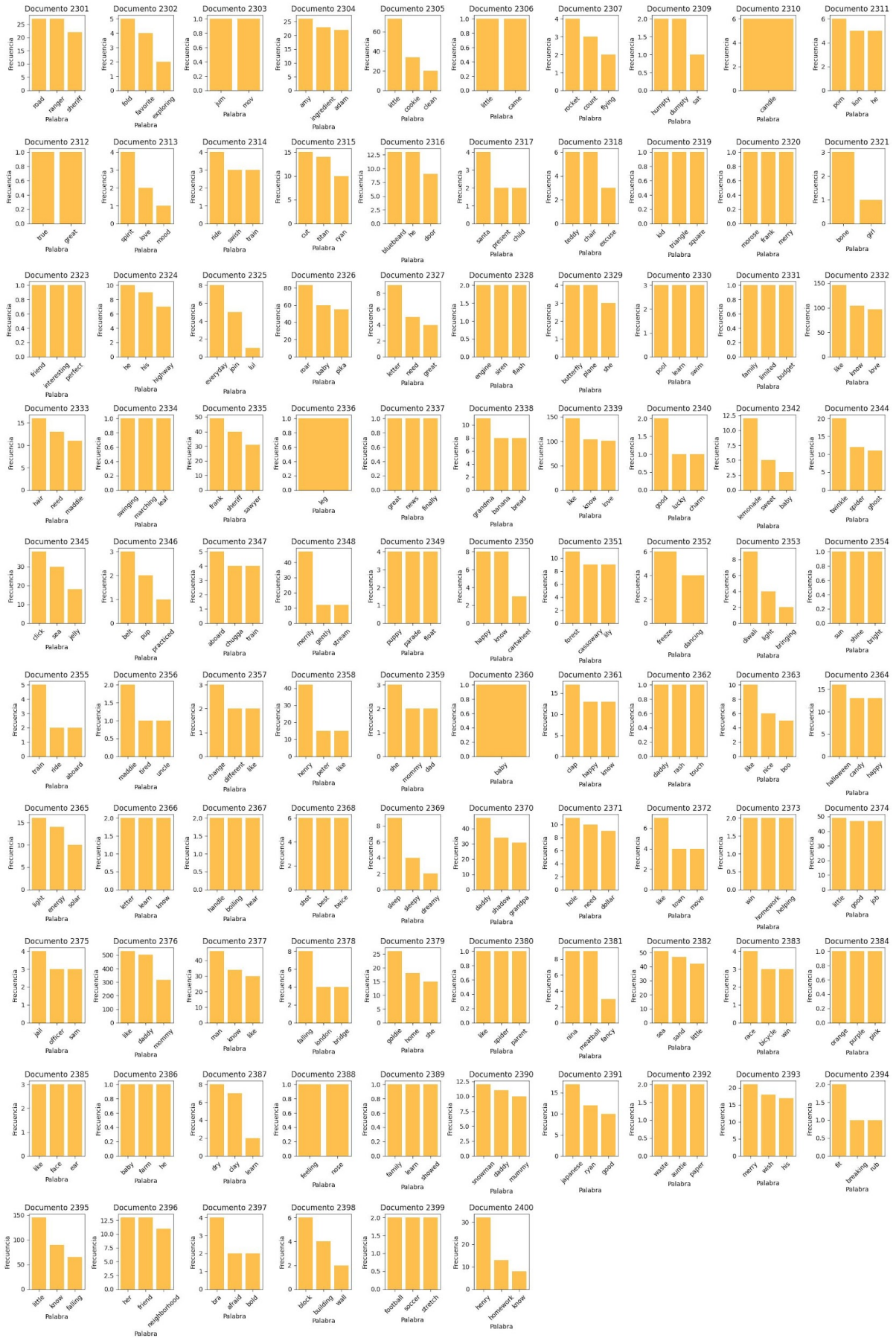


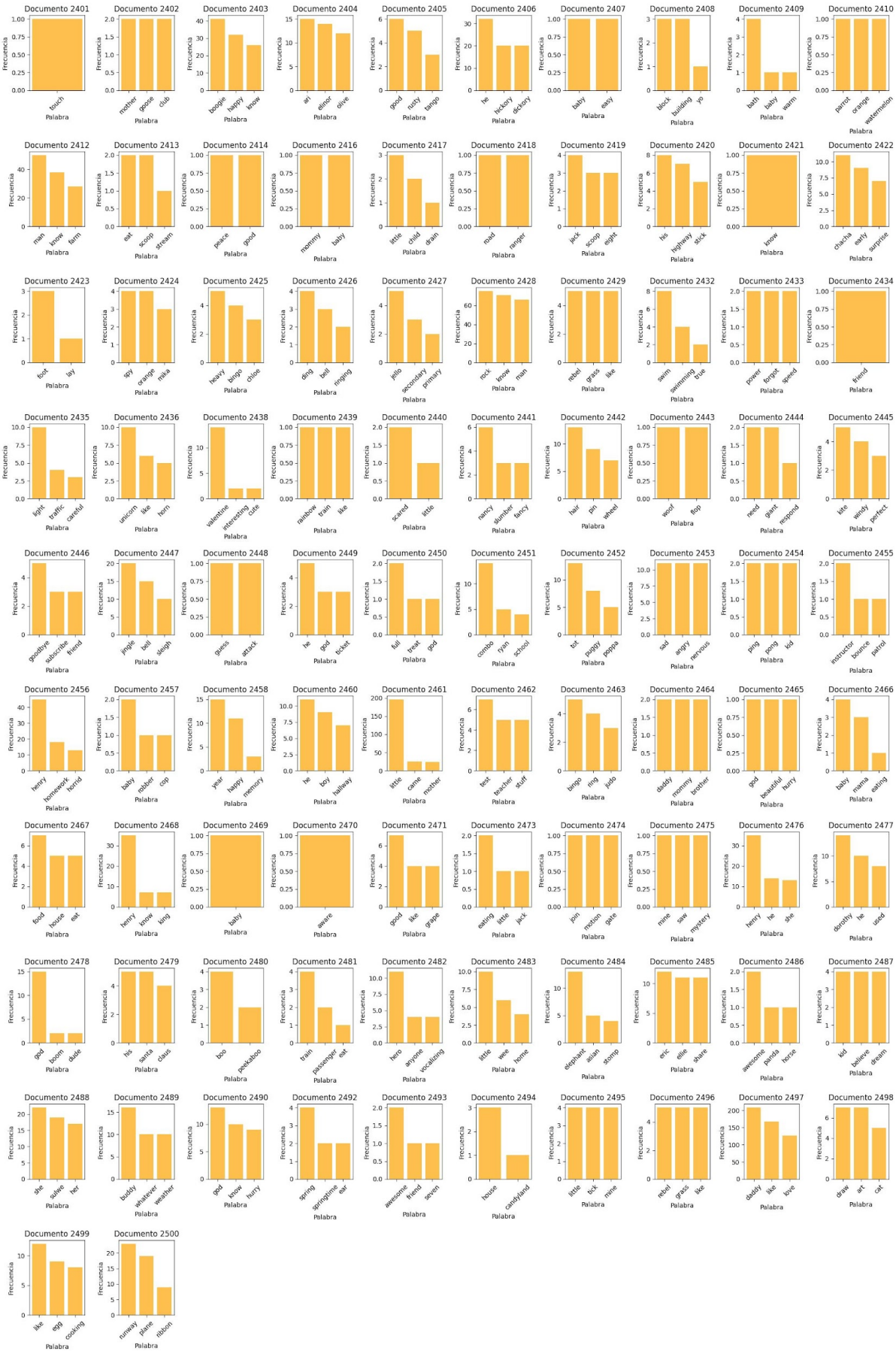












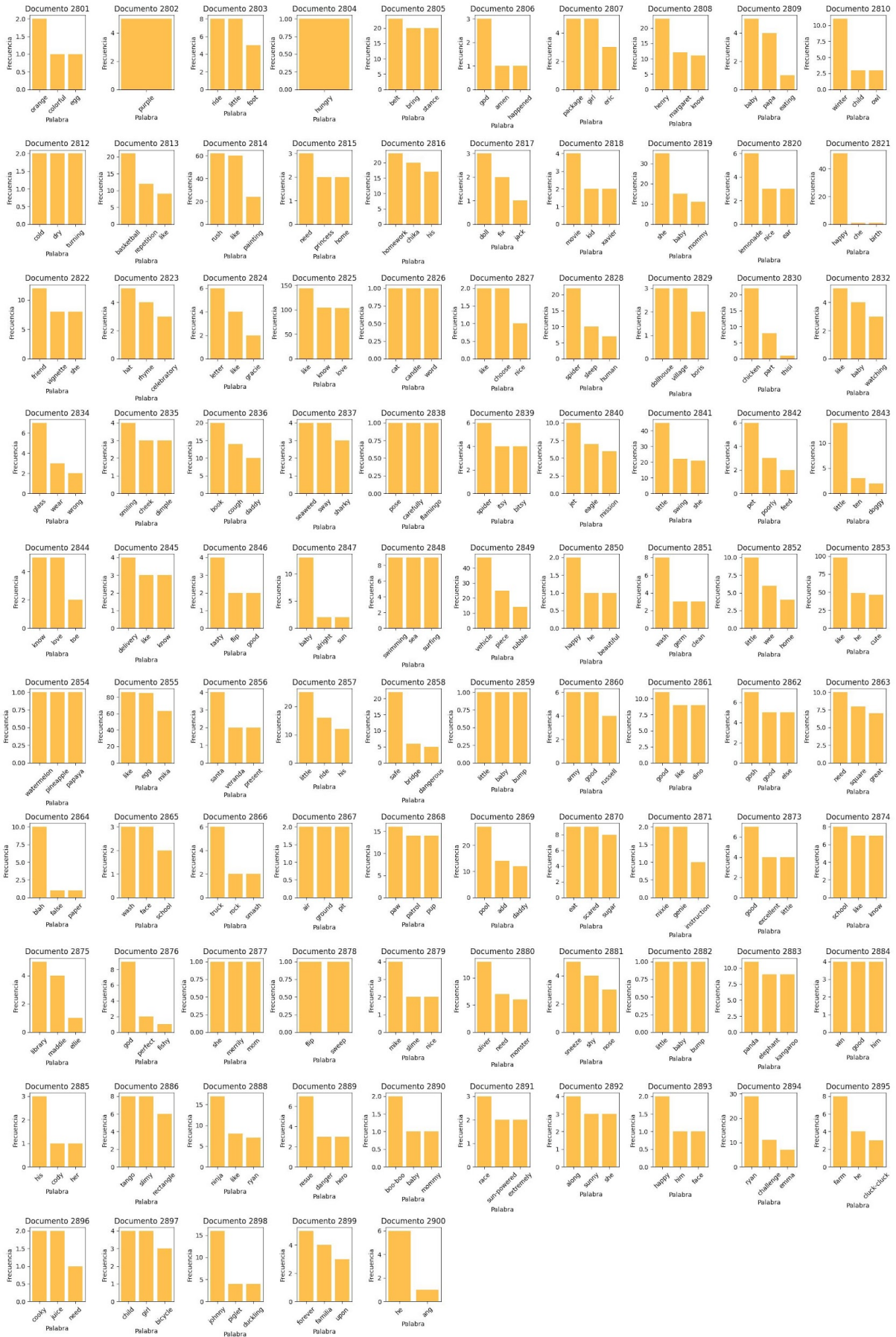






















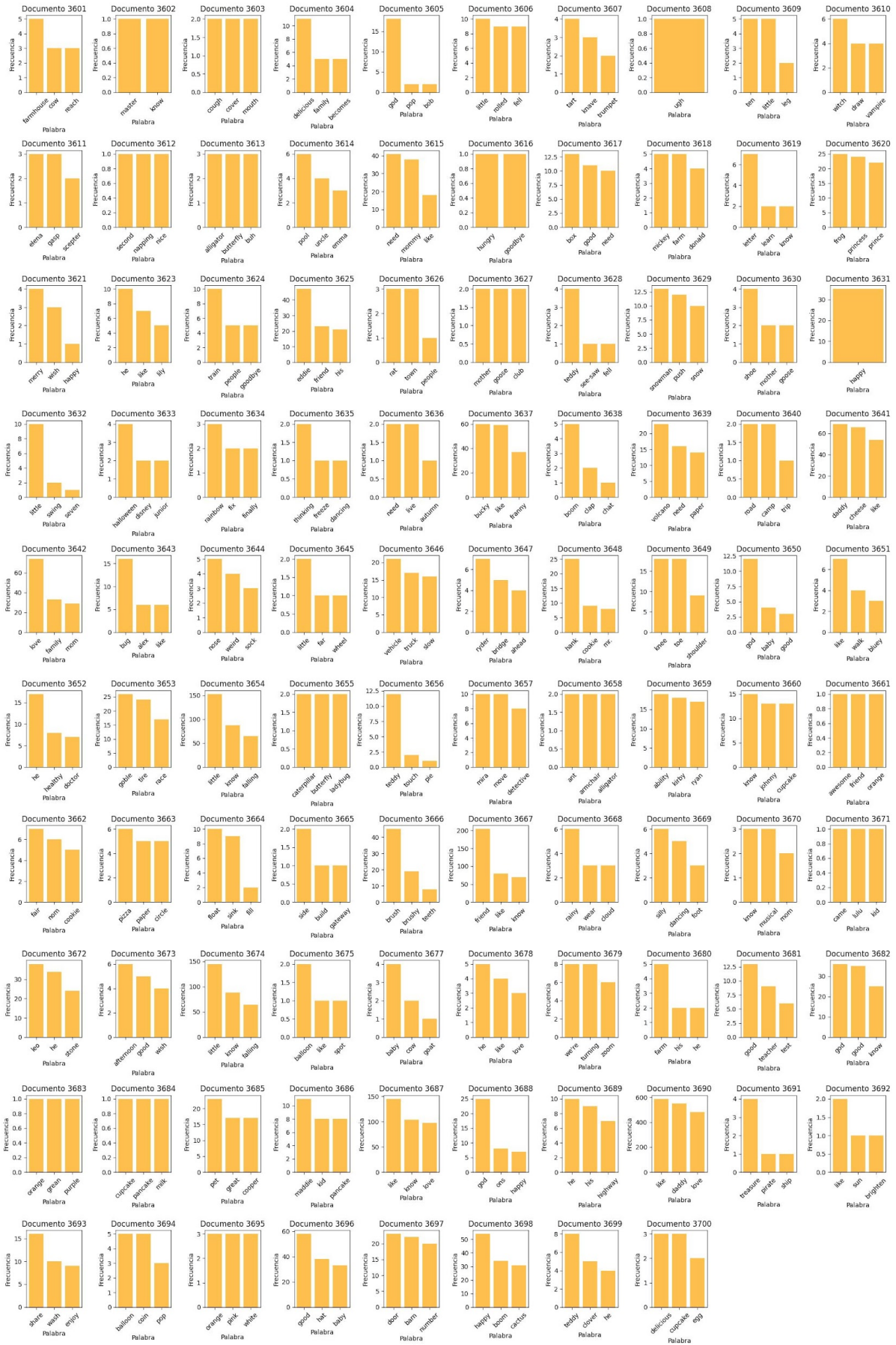






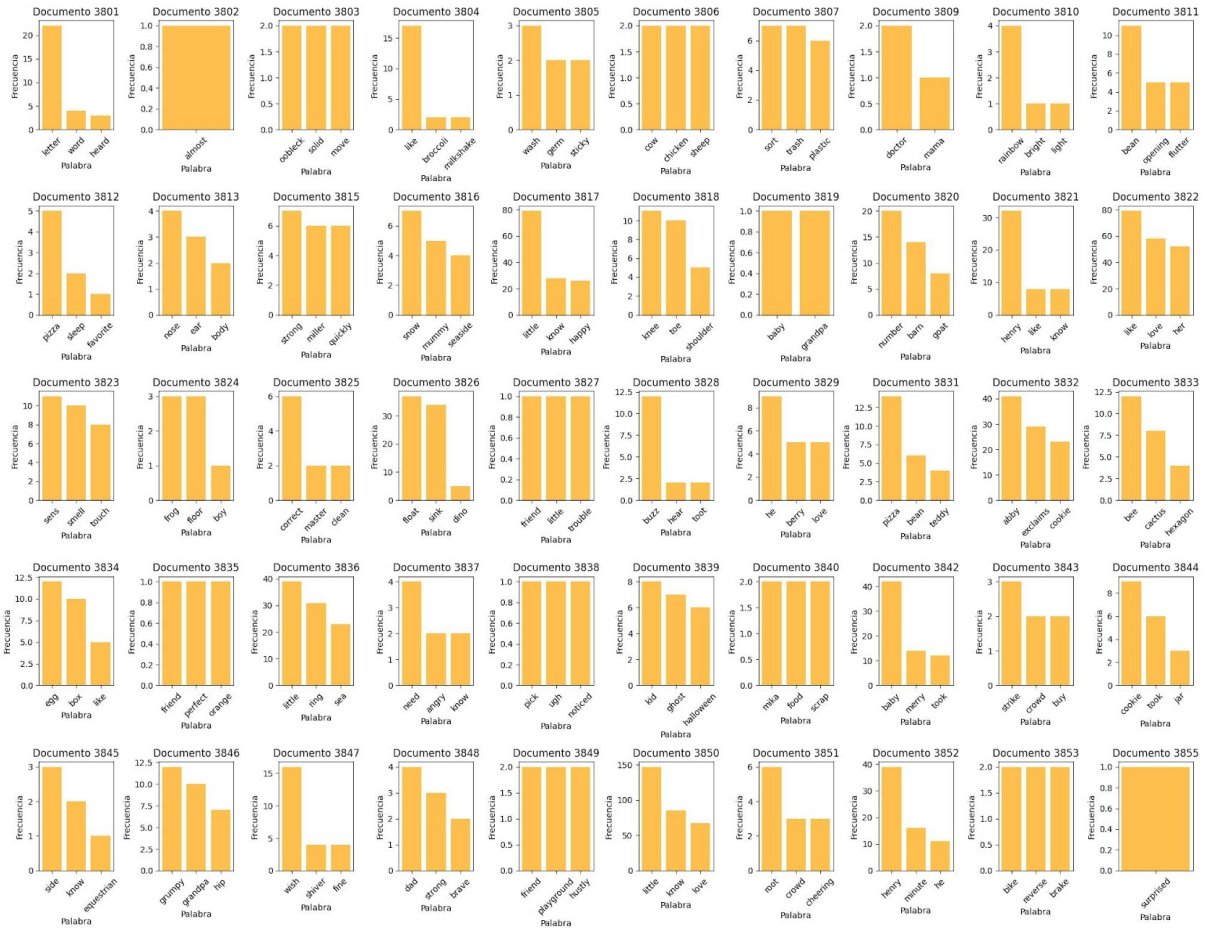












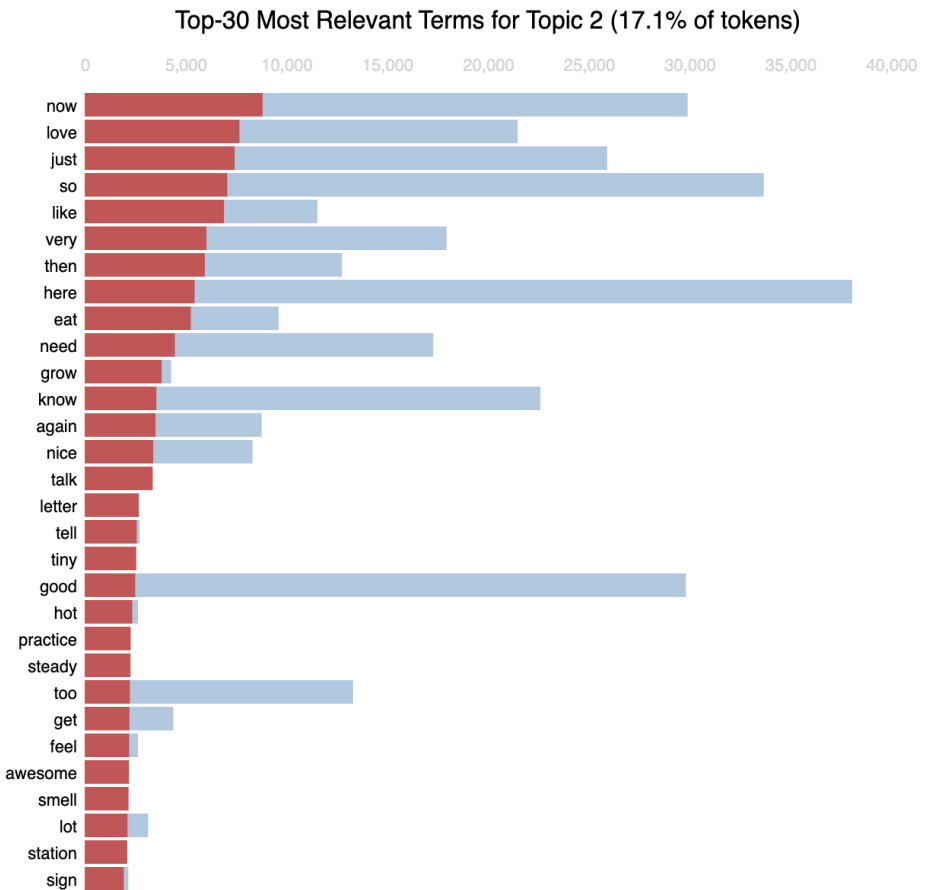
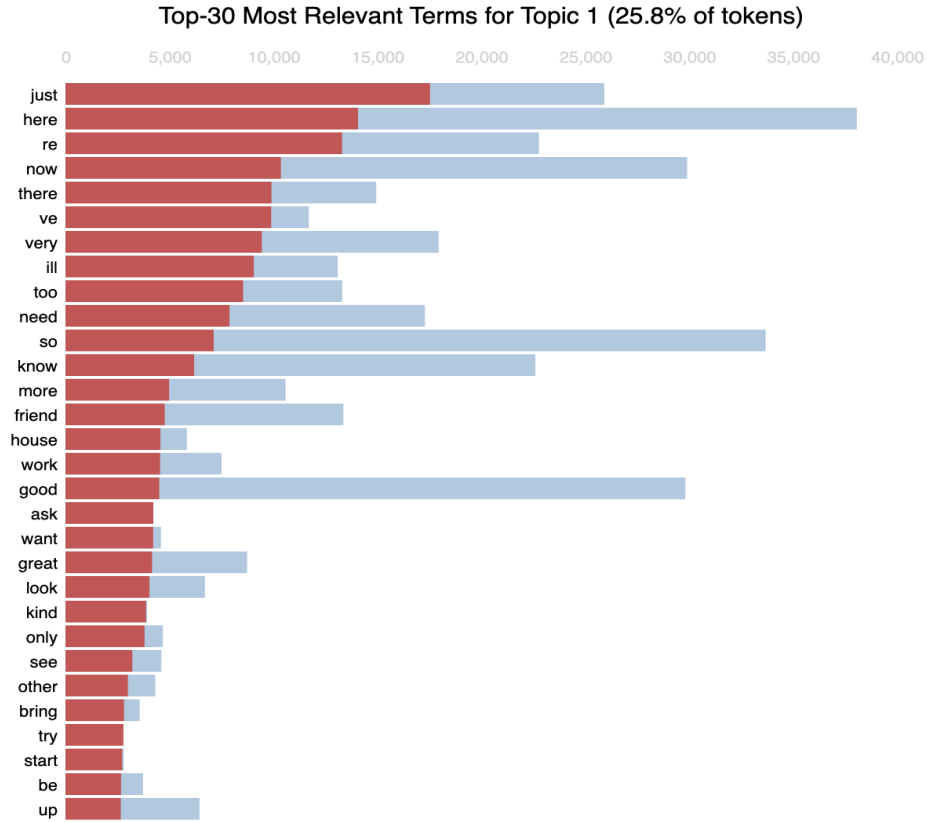
**8.5. Frecuencia de tokens de la lista de palabras**

| <b>palabra</b> | <b>ocurrencias</b> |
|----------------|--------------------|
| little         | 23,490             |
| poor           | 687                |
| healthy        | 1,133              |
| rescue         | 1,053              |
| feeling        | 1,322              |
| kind           | 2,794              |
| great          | 6,079              |
| care           | 2,499              |
| clean          | 2,456              |
| sad            | 1,086              |
| hurt           | 971                |
| cry            | 425                |
| cried          | 100                |
| scare          | 196                |
| hero           | 388                |
| violent        | 4                  |
| irrational     | 1                  |

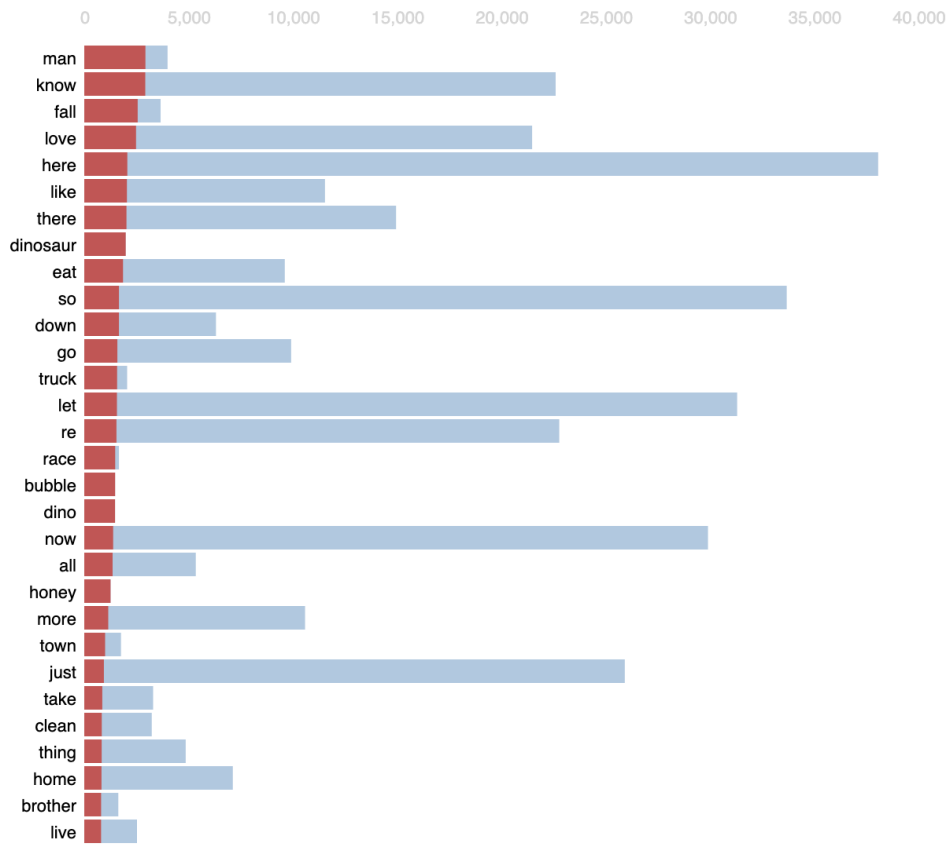


| <b>palabra</b> | <b>ocurrencias</b> |
|----------------|--------------------|
| school         | 2,414              |
| job            | 4,963              |
| doctor         | 1,838              |
| teacher        | 691                |
| house          | 4,641              |
| vacuum         | 178                |
| cooking        | 597                |
| cooked         | 41                 |
| working        | 1,047              |
| teach          | 367                |
| responsibility | 20                 |
| medicine       | 94                 |
| healer         | 43                 |
| worked         | 208                |
| police         | 3,391              |
| nurse          | 40                 |
| professional   | 38                 |
| chief          | 101                |
| babysit        | 17                 |
| football       | 286                |
| architecture   | 1                  |
| engineer       | 3                  |
| surgeon        | 4                  |
| technician     | 2                  |
| technologist   | 1                  |
| science        | 470                |

**8.6. Términos más frecuentes de los 4 tópicos más prevalentes**



**Top-30 Most Relevant Terms for Topic 4 (7.7% of tokens)**



**Top-30 Most Relevant Terms for Topic 3 (11.1% of tokens)**

