Presentación centralex

En este documento se presenta una librería de herramientas en lo que pretende ser una guía rápida de uso para el cálculo de la disponibilidad léxica, la centralidad léxica y las diferentes funcionalidades que de esta última se derivan. Aunque la librería permite, a través de programación funcional, implementar múltiples modelos, se proporciona un conjunto de herramientas elementales prediseñadas para que se pueda empezar a usar con las funciones señaladas (disponibilidad, centralidad) de forma casi inmediata. Sin embargo, consideramos necesario hacer notar que estas posibilidades de aplicación directa representan tan solo un subconjunto de las opciones que permite el sistema aquí presentado.

Actualmente está en preparación otro trabajo en el que se discuten en profundidad tanto los modelos posibles y las diferentes opciones de aplicación, como un estudio de diferentes aspectos de las bases teóricas del cálculo de la disponibilidad y su interpretación. Sin embargo, el desarrollo de estas etapas posteriores requiere cierta destreza en el manejo de la sintaxis de R y de conceptos de programación funcional, con lo que su presentación requiere otro ámbito de exposición.

# Exposición y contextualización del sistema

El sistema se ha construido sobre R[[1]](#footnote-21), una implementación abierta del lenguaje de programación S, que lleva en desarrollo desde 1992 y con una primera versión estable en el año 2000. Es fácilmente extensible y permite la reutilización de código y datos. Se trata de un sistema soportado por una amplia comunidad a lo largo del mundo y se usa de forma habitual en entornos de análisis de datos de ingeniería y medicina.

RStudio[[2]](#footnote-23) es un entorno comercial de desarrollo para R que permite el uso personal gratuito y que aumenta la productividad en el uso de este lenguaje. Está disponible para la gran mayoría de plataformas y sistemas operativos. Para facilitar el trabajo, se proporcionan utilidades que se pueden ejecutar dentro de este entorno y que facilitan el proceso de análisis.

Es posible utilizar las funcionalidades proporcionadas por nuestro software sin el uso de RStudio, pero consideramos que su empleo se dificulta y se pierden muchas de las posibilidades que proporciona este marco de software.

Dentro del sistema construido en RStudio nos resulta especialmente interesante la posibilidad de usar el paradigma de la *programación literaria*, según la cual, en un mismo documento, se integran tanto los contenidos textuales como la exposición de hechos o la interpretación de los análisis a través de bloques de código que llevan a cabo las manipulaciones y representaciones de datos. Lo que sigue pretende ser un ejemplo práctico del uso de esta metodología. Como muestra de sus ventajas –al incluir en el mismo documento el análisis que se ha llevado a cabo con el código para realizarlo– se potencia la replicabilidad, característica cada vez más exigida en el mundo académico, en la que se debe permitir que otros investigadores puedan replicar, tanto para verificar como para desarrollar de manera individual, el trabajo expuesto. Este objetivo se cumple al exponer en el mismo documento la disertación y el código con el que se ha realizado los análisis.

A diferencia de soluciones basadas en programas autónomos para el cálculo de la disponibilidad léxica, la implementación de esta solución dentro de un sistema como el formado por R junto con RStudio permite, además, la integración de todo el procedimiento de gestión de datos. De esta forma, los resultados obtenidos a partir de las herramientas proporcionadas se incorporan al sistema, con lo que se permite la integración en nuevos procesos de datos y la aplicación personal de los resultados obtenidos.

La distribución de las soluciones es otro problema que ya ha encontrado soluciones eficientes dentro de la comunidad de usuarios de R. Repositorios de código como GitHub permiten la gestión y el control automatizado de versiones. Es posible llevar a cabo un control exhaustivo de las modificaciones que se van realizando, sin necesidad de descargar e instalar continuamente programas. Con dos líneas de código podemos estar seguros de tener instalada la última versión disponible. Además, este modelo de intercambio de código permite que los proyectos no desaparezcan, ya que se pueden crear ramas de cualquier repositorio de código para continuar su desarrollo, independientemente de los creadores originales.

Entre las múltiples herramientas construidas sobre R hemos considerado muy interesantes y productivas las herramientas que conforman el “Universo Tidyverse”[[3]](#footnote-25). Se trata de un conjunto de librerías que facilitan y permiten una manipulación expresiva de los datos, basada en el concepto de tubería. Este tratamiento se construye mediante la secuenciación de manipulaciones que se van encadenando, de manera que el resultado de una etapa se convierte en la fuente de datos para la siguiente. Esta característica, junto con operadores adaptados a este tipo de trabajo, proporciona una expresividad potente y cómoda de llevar a cabo transformaciones que, de otra forma, serían muy complejas. Las manipulaciones de datos presentadas en este trabajo se realizan utilizando estas herramientas.

# Instalación

## Prerequisitos

En primer lugar es necesario instalar el sistema R. Para ello se accede, en la sección *Download* de la página del proyecto, <https://www.r-project.org/>, al enlace [CRAN](http://cran.r-project.org/mirrors.html), seleccionar un sitio –en nuestro caso vamos a escoger el de la Red Iris, que es una réplica local en España <https://cran.rediris.es/>– y, en la página principal encontramos un enlace para descargar la versión para nuestro sistema operativo. En el caso de elegir la versión para Windows, se elegiría la versión base y accedemos, finalmente, al binario que instala el sistema, que en este momento, corresponde a la versión 4.0.2, en el enlace (<https://cran.rediris.es/bin/windows/base/R-4.0.2-win.exe>)[<https://cran.rediris.es/bin/windows/base/R-4.0.2-win.exe>].

A continuación se recomienda, aunque no es estrictamente necesario, instalar el entorno de productividad RStudio. En (<https://rstudio.com/>)[<https://rstudio.com/>], siguiendo, en este momento, el camino Product > Open Source > RStudio, se llega a la página de descripción del producto, del que hay dos versiones: Desktop y Server. En nuestro caso usaremos la versión Desktop, en su versión *Open Source Edition*, que se puede descargar en (<https://rstudio.com/products/rstudio/download/>)[<https://rstudio.com/products/rstudio/download/>]. La versión actual para Windows se encuentra en (<https://download1.rstudio.org/desktop/windows/RStudio-1.3.1073.exe>)[<https://download1.rstudio.org/desktop/windows/RStudio-1.3.1073.exe>].

## Centralex

Una vez instalados los prerequisitos, para poder utilizar las herramientas que proponemos, el usuario debe llevar a cabo un segundo paso de instalación que, en nuestra opinión, no es más complejo que cualquier proceso de instalación en el sistema *R*. Este paso ha de hacerse únicamente cuando se pretenda instalar o actualizar el paquete en el sistema. Puesto que nuestra intención es seguir trabajando en el paquete, sería recomendable realizar esta acción de forma periódica.

Las siguientes órdenes instalan el paquete *Centralex* desde el repositorio *GitHub*. Todo el código está implementado en *R*, con lo que es fácil de revisar.

install.packages("devtools")  
library(devtools)  
install\_github("jmss70/centralex")

Para facilitar la introducción de la herramienta se han impuesto como requerimientos, para que se instalen de forma automática, las herramientas del universo *Tidyverse*, que facilitan el procesamiento de datos y se usa de forma extensiva en este y otros trabajos, y la librería *flextable* y *kableExtra*, con el objetivo de mejorar la presentación de las tablas.

Una vez que se han llevado a cabo los pasos anteriores, y si no se ha producido ningún contratiempo, el sistema está preparado para ejecutar el cálculo de los índices de disponibilidad y centralidad.

# Carga de las librerías y los datos

Para poder usar las librerías hay que cargarlas en nuestra sesión de trabajo. De esta manera se ponen a disposición del usuario las funciones que proporcionan.

library(centralex)

El formato requerido de los datos coincide con el tradicionalmente usado por los programas más empleados hasta ahora para el cálculo de la disponibilidad léxica (**Lexidisp**, **Dispolex**, esencialmente). Sin embargo, consideramos que parte de la información que contiene su estructura es redundante y sería conveniente plantear la posibilidad de establecer un estándar de codificación que sea más coherente con los modelos de datos normalizados.

De momento, se espera que los datos estén en un archivo de texto, con campos separados por espacios, según el siguiente formato:

* Un campo con información sociológica básica del hablante formada por cinco caracteres sucesivos. Cada uno de ellos representa la codificación de una variante dentro de una variable sociológica
* Un campo de identificación del usuario de tres caracteres
* Un campo de identificación del centro de interés de dos caracteres
* Una lista de palabras separadas por comas y siguiendo el orden de su aparición en las listas originales. Es necesario que no se utilice el carácter “coma” como valor de un elemento, ya que se interpretaría como separador.

Un ejemplo de dos líneas sería:

21131 001 01 mano, pie, brazo, cerebro, pulmón, nariz, extremidad, ojo, boca, diente, pelo, oreja, culo, vagina   
12131 002 01 riñón, corazón, garganta, cabeza, pierna, pie, hígado, estómago, mano, brazo, antebrazo, abdomen, pecho, ojo, boca, oído, dedo, rodilla, costilla

Suponiendo que tenemos todos los datos cargados en un archivo, denominado datos.txt, que estará alojado en el mismo directorio que el script de procesamiento, se podrían cargar los datos como:

data <- read.centralex("datos.txt")  
data %>%   
 head() %>%   
 flextable() %>% autofit() %>%  
 theme\_booktabs()

| infos | users | centers | words |
| --- | --- | --- | --- |
| 21131 | 001 | 01 | mano, pi.... |
| 12131 | 002 | 01 | riñón, c.... |
| 12213 | 003 | 01 | brazo, m.... |
| 22214 | 004 | 01 | brazo, o.... |
| 12214 | 005 | 01 | cabeza, .... |
| 22213 | 006 | 01 | pie, man.... |

# Cálculo de la disponibilidad

La función general del cálculo de la disponibilidad es build.availability. Sin embargo, el uso de esta función requiere del uso de varios parámetros. Se han construidos dos funciones de utilidad que encapsulan su uso y ofrecen los dos modelos que consideramos, por el momento, más interesantes (Modelo López-Strassburger y Modelo Ávila-Sánchez). Consideramos que es preferible exponer su uso mediante ejemplos.

## Modelo de López-Strassburger

Este modelo, que es el más conocido y el último de una serie de iteraciones para la cuantificación de la disponibilidad, se puede usar como:

disponibilidad <- build.lopezstrass.availability(data)

El resultado es un nuevo marco de datos cuyos campos son el centro de interés, la palabra y la disponibilidad calculada, junto con otros datos necesarios para su obtención como la posición y las frecuencias relativas y acumuladas dentro del centro de interés:

head(disponibilidad) %>%  
 flextable() %>%   
 set\_header\_labels(centers = "Centro de interés",   
 words="Palabra",   
 order="Orden",   
 availability="Disponibilidad",   
 freq.abs="Frecuencia absoluta",   
 freq.abs.cum="Frecuencia absoluta acumulada",   
 freq.rel="Frecuencia relativa",   
 freq.rel.cum="Frecuencia relativa acumulada") %>%  
 colformat\_num(j=c(4,6,8), digits=6) %>%  
 width(j=c(1,3,5,7),width=.65) %>%  
 width(j=c(2,4,6,8),width=.9) %>%  
 theme\_booktabs()

| Centro de interés | Palabra | Orden | Disponibilidad | Frecuencia absoluta | Frecuencia relativa | Frecuencia absoluta acumulada | Frecuencia relativa acumulada |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 01 | ojo | 1 | 0.632987 | 58 | 0.037395 | 58 | 0.037395 |
| 01 | brazo | 2 | 0.551808 | 50 | 0.032237 | 108 | 0.069632 |
| 01 | pierna | 3 | 0.507215 | 49 | 0.031593 | 157 | 0.101225 |
| 01 | mano | 4 | 0.489482 | 42 | 0.027079 | 373 | 0.240490 |
| 01 | cabeza | 5 | 0.478054 | 37 | 0.023856 | 410 | 0.264346 |
| 01 | pie | 6 | 0.450128 | 43 | 0.027724 | 331 | 0.213411 |

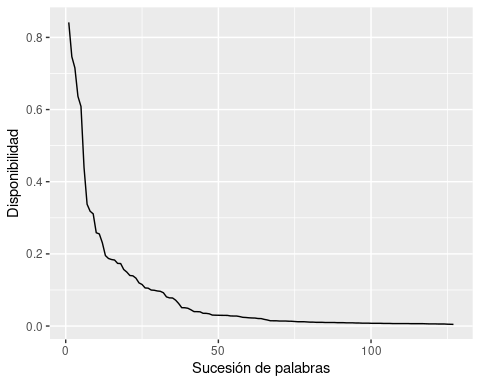
Al ser un marco de datos de *R* estándar, se pueden realizar sobre él todas las operaciones que permite el sistema. En el siguiente ejemplo hemos seleccionado los resultados obtenidos en el centro de interés “03”, los hemos ordenado en orden decreciente de disponibilidad y mostramos los 10 vocablos más disponibles

disponibilidad %>%   
 filter(centers=="03") %>%   
 arrange(-availability) %>%   
 head(10) %>%  
 flextable() %>%   
 set\_header\_labels(centers = "Centro de interés",   
 words="Palabra",   
 order="Orden",   
 availability="Disponibilidad",   
 freq.abs="Frecuencia absoluta",   
 freq.abs.cum="Frecuencia absoluta acumulada",   
 freq.rel="Frecuencia relativa",   
 freq.rel.cum="Frecuencia relativa acumulada") %>%  
 colformat\_num(j=c(4,6,8), digits=6) %>%  
 width(j=c(1,3,5,7),width=.65) %>%  
 width(j=c(2,4,6,8),width=.9) %>%  
 theme\_booktabs()

| Centro de interés | Palabra | Orden | Disponibilidad | Frecuencia absoluta | Frecuencia relativa | Frecuencia absoluta acumulada | Frecuencia relativa acumulada |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 03 | cocina | 1 | 0.841569 | 69 | 0.066925 | 69 | 0.066925 |
| 03 | (cuarto de) baño | 2 | 0.746670 | 66 | 0.064016 | 135 | 0.130941 |
| 03 | salón | 3 | 0.715562 | 58 | 0.056256 | 255 | 0.247333 |
| 03 | dormitorio | 4 | 0.636728 | 55 | 0.053346 | 310 | 0.300679 |
| 03 | terraza | 5 | 0.608507 | 62 | 0.060136 | 197 | 0.191077 |
| 03 | comedor | 6 | 0.439152 | 39 | 0.037827 | 349 | 0.338506 |
| 03 | salita | 7 | 0.337618 | 31 | 0.030068 | 447 | 0.433560 |
| 03 | lavadero | 8 | 0.318056 | 33 | 0.032008 | 416 | 0.403492 |
| 03 | pasillo | 9 | 0.310978 | 34 | 0.032978 | 383 | 0.371484 |
| 03 | (cuarto de) aseo | 10 | 0.258557 | 24 | 0.023278 | 500 | 0.484966 |

También es posible utilizar los mecanismos de *R* para representar la curva de disponibilidad, esto es, la sucesión de valores de disponibilidad, una vez ordenados en valor decreciente de disponibilidad:

disponibilidad %>%   
 filter(centers=="03") %>%   
 arrange(-availability) %>%   
 ggplot(aes(x=order, y=availability)) + geom\_line() +  
 xlab("Sucesión de palabras") + ylab("Disponibilidad")



## Modelo de Ávila-Sánchez

Ávila y Sánchez (2014 Fuzzy sets and Prototype Theory: Representational model of cognitive community structures based on lexical availability trials) propusieron un macro-modelo para el estudio de la disponibilidad léxica a partir de la Teoría de los Conjuntos Difusos y mediante la modelización de los conceptos que se pretenden representar. Esta modelización se produce en dos etapas. En la primera, se cuantifica la relevancia de cada término en las pruebas obtenidas para cada hablante y centro de interés según una ley descendente según se avanza en cada listado, y en la segunda etapa se integra esa información con una ley aditiva que considera los distintos valores alcanzados para cada palabra en cada centro de interés.

Hay múltiples posibles elecciones, pero en distintas pruebas resultaron más prometedoras las que utilizaban en la primera etapa una ley de Zipf-Mandelbrot y en la segunda una adición probabilística. La interpretación de los valores obtenidos corresponde al concepto de ‘centralidad’ de cada término en cada centro de interés. Un valor de 1, o muy cercano, respondería a la pertenencia al núcleo del vocabulario específico del centro de interés, mientras que un valor próximo a 0 indicaría que se trataría de un término poco accesible y, por tanto, descentralizado.

La función que lleva a cabo este análisis, build.avilasanchez.availability, se utiliza de la misma forma que la mostrada anteriormente en el modelo López-Strassburger:

disponibilidad <- build.avilasanchez.availability(data)

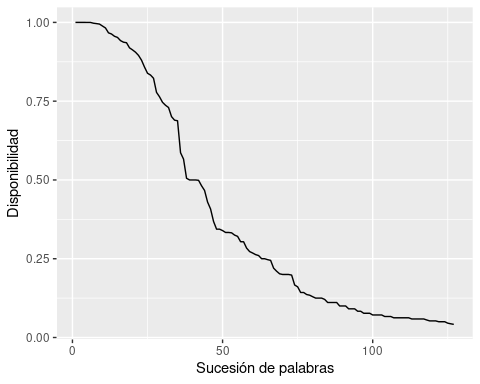
El resultado es, de igual modo, un nuevo marco de datos con la disponibilidad de cada término en cada centro de interés

disponibilidad %>%   
 filter(centers=="03") %>%   
 arrange(-availability) %>%   
 head(10) %>%   
 flextable() %>%  
 set\_header\_labels(centers = "Centro de interés",   
 words="Palabra",   
 order="Orden",   
 availability="Disponibilidad",   
 freq.abs="Frecuencia absoluta",   
 freq.abs.cum="Frecuencia absoluta acumulada",   
 freq.rel="Frecuencia relativa",   
 freq.rel.cum="Frecuencia relativa acumulada") %>%  
 colformat\_num(j=c(4,6,8), digits=6) %>%  
 width(j=c(1,3,5,7),width=.65) %>%  
 width(j=c(2,4,6,8),width=.9) %>%  
 theme\_booktabs()

| Centro de interés | Palabra | Orden | Disponibilidad | Frecuencia absoluta | Frecuencia relativa | Frecuencia absoluta acumulada | Frecuencia relativa acumulada |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 03 | cocina | 1 | 1.000000 | 69 | 0.066925 | 69 | 0.066925 |
| 03 | salón | 2 | 1.000000 | 58 | 0.056256 | 255 | 0.247333 |
| 03 | (cuarto de) baño | 3 | 1.000000 | 66 | 0.064016 | 135 | 0.130941 |
| 03 | dormitorio | 4 | 1.000000 | 55 | 0.053346 | 310 | 0.300679 |
| 03 | comedor | 5 | 0.999928 | 39 | 0.037827 | 349 | 0.338506 |
| 03 | terraza | 6 | 0.999900 | 62 | 0.060136 | 197 | 0.191077 |
| 03 | salita | 7 | 0.997438 | 31 | 0.030068 | 447 | 0.433560 |
| 03 | entrada | 8 | 0.996197 | 23 | 0.022308 | 523 | 0.507274 |
| 03 | (cuarto de) aseo | 9 | 0.994715 | 24 | 0.023278 | 500 | 0.484966 |
| 03 | pasillo | 10 | 0.988474 | 34 | 0.032978 | 383 | 0.371484 |

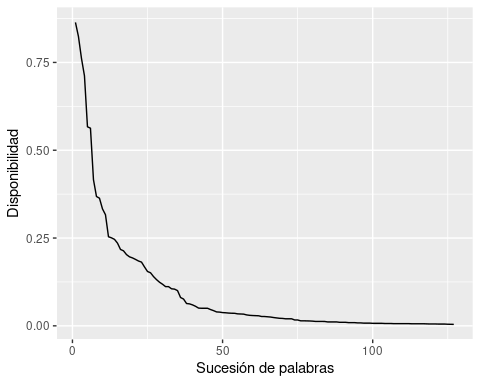
Que se puede procesar como cualquier marco de datos del sistema *R*:

disponibilidad %>%   
 filter(centers=="03") %>%   
 arrange(-availability) %>%   
 ggplot(aes(x=order, y=availability)) + geom\_line() +  
 xlab("Sucesión de palabras") + ylab("Disponibilidad")



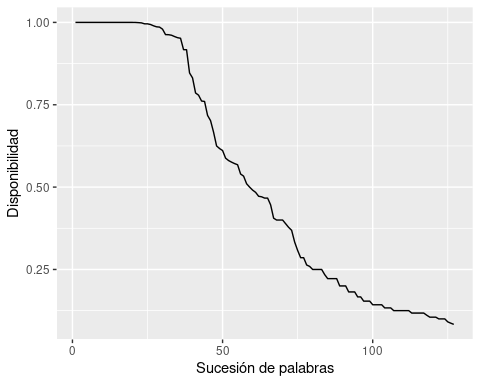
Debido a las características de los operadores aditivos de la teoría de los conjuntos difusos, es posible que la forma de la curva obtenida para el cálculo de la disponibilidad clásica: quizás no aparezcan términos con valores cercanos a 1 (por ejemplo, si se tienen pocos datos y estos son relativamente dispersos) o, por el contrario, encontremos demasiados términos con valoraciones cercanas a 1 (por ejemplo, si se trabaja con muchas muestras). Aunque las opciones anteriores son improbables, pero posibles, hemos optado por regular la curva mediante un parámetro adicional, k, que va a modificar la curva (ascendente-descendente), pero manteniendo su forma y clasificación. El valor por defecto de k es 1. Si se da un valor entre 0 y 1, la curva bajará (menos términos con valores cercanos a 1), mientras que si a k se le da un valor mayor que la unidad la curva subirá (más términos cercanos a la unidad).

disponibilidad <- build.avilasanchez.availability(data, k = 0.1)  
disponibilidad %>%  
 filter(centers=="03") %>%   
 arrange(-availability) %>%   
 ggplot(aes(x=order, y=availability)) + geom\_line() +  
 xlab("Sucesión de palabras") + ylab("Disponibilidad")



El resultado final será, en cualquier caso, un nuevo marco de datos con la disponibilidad de cada término en cada centro de interés

disponibilidad <- build.avilasanchez.availability(data, k = 2)  
disponibilidad %>%  
 filter(centers=="03") %>%   
 arrange(-availability) %>%   
 ggplot(aes(x=seq\_along(availability), y=availability)) + geom\_line() +  
 xlab("Sucesión de palabras") + ylab("Disponibilidad")



Téngase en cuenta que los ejemplos mostrados representan de forma artificial, manipulada y consciente, valores muy extremos usados a modo de exposición. En nuestras pruebas, realizadas siempre con datos reales procedentes de investigaciones previas, hemos encontrado que el valor de referencia ofrece resultados adecuados en todos los casos que hemos encontrado.

# Niveles de disponibilidad

Una pregunta recurrente en casi todos los estudios previos de disponibilidad léxica es considerar cuál es el tamaño del conjunto de elementos que se habría que considerar para establecer el núcleo de un centro de interés. Para responder a esta cuestión, y a partir del nuevo marco teórico creado en nuestra propuesta, se proporciona una herramienta que etiqueta los términos por niveles de centralidad. El nivel 0 correspondería a aquellos elementos que no pertenecen al núcleo, es decir, aquellos términos que no son generalmente accesibles. Los níveles 1, 2, 3, … y sucesivos representarían un mayor grado de centralidad y aproximación al centro de interés.

disponibilidad <- build.avilasanchez.availability(data)  
levels <- classify.availability.levels(disponibilidad)  
levels %>%   
 head(20) %>%  
 arrange(-availability) %>%  
 select(-order) %>%  
 flextable() %>%  
 set\_header\_labels(centers = "Centro de interés",   
 words="Palabra",   
 availability="Disponibilidad",  
 level="Nivel de disponibilidad",  
 cutlevel="Nivel de corte",  
 freq.abs="Frecuencia absoluta",   
 freq.abs.cum="Frecuencia absoluta acumulada",   
 freq.rel="Frecuencia relativa",   
 freq.rel.cum="Frecuencia relativa acumulada") %>%  
 colformat\_num(j=c(5,7,9), digits=5) %>%  
 width(j=c(1,4,6,8),width=.65) %>%  
 width(j=c(2,3,5,6,7,9),width=.75) %>%  
 theme\_booktabs()

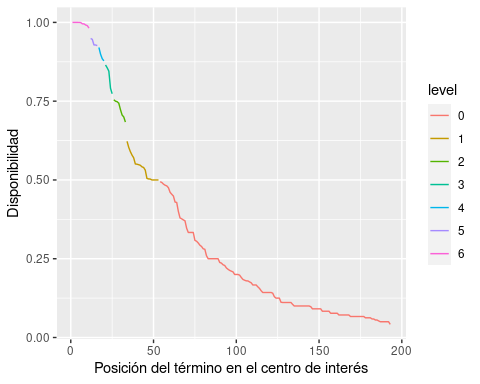
| Centro de interés | Palabra | Disponibilidad | Nivel de disponibilidad | Nivel de corte | Frecuencia absoluta | Frecuencia relativa | Frecuencia absoluta acumulada | Frecuencia relativa acumulada |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 14 | perro | 1 | 6 | 0.94971 | 67 | 0.04206 | 67 | 0.04206 |
| 12 | coche | 1 | 6 | 0.95891 | 65 | 0.05717 | 65 | 0.05717 |
| 19 | rojo | 1 | 6 | 0.96142 | 66 | 0.05255 | 203 | 0.16162 |
| 03 | cocina | 1 | 6 | 0.96381 | 69 | 0.06693 | 69 | 0.06693 |
| 19 | azul | 1 | 6 | 0.96142 | 68 | 0.05414 | 137 | 0.10908 |
| 14 | gato | 1 | 6 | 0.94971 | 64 | 0.04018 | 131 | 0.08223 |
| 04 | mesa | 1 | 6 | 0.95622 | 54 | 0.05182 | 230 | 0.22073 |
| 03 | salón | 1 | 6 | 0.96381 | 58 | 0.05626 | 255 | 0.24733 |
| 17 | euro | 1 | 6 | 0.94392 | 45 | 0.04014 | 45 | 0.04014 |
| 09 | aire acondicionado | 1 | 6 | 0.95832 | 58 | 0.07827 | 118 | 0.15924 |
| 06 | tenedor | 1 | 6 | 0.95511 | 59 | 0.05530 | 59 | 0.05530 |
| 11 | árbol | 1 | 6 | 0.93846 | 51 | 0.04048 | 51 | 0.04048 |
| 08 | pizarra | 1 | 6 | 0.95052 | 62 | 0.04769 | 62 | 0.04769 |
| 10 | calle | 1 | 6 | 0.95278 | 51 | 0.03783 | 51 | 0.03783 |
| 20 | ola | 1 | 6 | 0.94369 | 46 | 0.03430 | 97 | 0.07233 |
| 04 | silla | 1 | 6 | 0.95622 | 63 | 0.06046 | 63 | 0.06046 |
| 09 | ventilador | 1 | 6 | 0.95832 | 60 | 0.08097 | 60 | 0.08097 |
| 02 | pantalón | 1 | 6 | 0.95766 | 62 | 0.04829 | 62 | 0.04829 |
| 08 | mesa | 1 | 6 | 0.95052 | 46 | 0.03538 | 385 | 0.29615 |
| 01 | cabeza | 1 | 6 | 0.95760 | 37 | 0.02386 | 410 | 0.26435 |

levels %>%   
 filter(centers=="01") %>%   
 arrange(-availability) %>%   
 filter(level > 0) %>%  
 select(-order) %>%   
 flextable() %>%   
 set\_header\_labels(centers = "Centro de interés",   
 words="Palabra",   
 availability="Disponibilidad",  
 level="Nivel de disponibilidad",  
 cutlevel="Nivel de corte",  
 freq.abs="Frecuencia absoluta",   
 freq.abs.cum="Frecuencia absoluta acumulada",   
 freq.rel="Frecuencia relativa",   
 freq.rel.cum="Frecuencia relativa acumulada") %>%  
 colformat\_num(j=c(3,5,7,9), digits=5) %>%  
 width(j=c(1,4,6,8),width=.65) %>%  
 width(j=c(2,3,5,6,7,9),width=.75) %>%  
 theme\_booktabs()

| Centro de interés | Palabra | Disponibilidad | Nivel de disponibilidad | Nivel de corte | Frecuencia absoluta | Frecuencia relativa | Frecuencia absoluta acumulada | Frecuencia relativa acumulada |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 01 | cabeza | 1.00000 | 6 | 0.95760 | 37 | 0.02386 | 410 | 0.26435 |
| 01 | brazo | 1.00000 | 6 | 0.95760 | 50 | 0.03224 | 108 | 0.06963 |
| 01 | ojo | 1.00000 | 6 | 0.95760 | 58 | 0.03740 | 58 | 0.03740 |
| 01 | mano | 1.00000 | 6 | 0.95760 | 42 | 0.02708 | 373 | 0.24049 |
| 01 | pierna | 0.99997 | 6 | 0.95760 | 49 | 0.03159 | 157 | 0.10123 |
| 01 | pie | 0.99992 | 6 | 0.95760 | 43 | 0.02772 | 331 | 0.21341 |
| 01 | corazón | 0.99937 | 6 | 0.95760 | 44 | 0.02837 | 201 | 0.12959 |
| 01 | dedo | 0.99810 | 6 | 0.95760 | 44 | 0.02837 | 245 | 0.15796 |
| 01 | nariz | 0.99786 | 6 | 0.95760 | 43 | 0.02772 | 288 | 0.18569 |
| 01 | boca | 0.99490 | 6 | 0.95760 | 33 | 0.02128 | 479 | 0.30883 |
| 01 | pelo | 0.99239 | 6 | 0.95760 | 27 | 0.01741 | 595 | 0.38362 |
| 01 | oreja | 0.98634 | 6 | 0.95760 | 29 | 0.01870 | 568 | 0.36622 |
| 01 | hueso | 0.97596 | 6 | 0.95760 | 21 | 0.01354 | 686 | 0.44230 |
| 01 | hígado | 0.97017 | 6 | 0.95760 | 29 | 0.01870 | 539 | 0.34752 |
| 01 | músculo | 0.96741 | 6 | 0.95760 | 19 | 0.01225 | 705 | 0.45455 |
| 01 | tronco | 0.96647 | 6 | 0.95760 | 10 | 0.00645 | 1040 | 0.67054 |
| 01 | uña | 0.96456 | 6 | 0.95760 | 36 | 0.02321 | 446 | 0.28756 |
| 01 | pulmón | 0.92715 | 5 | 0.91971 | 31 | 0.01999 | 510 | 0.32882 |
| 01 | riñón | 0.91971 | 5 | 0.91971 | 23 | 0.01483 | 643 | 0.41457 |
| 01 | estómago | 0.91195 | 4 | 0.85694 | 25 | 0.01612 | 620 | 0.39974 |
| 01 | extremidad | 0.88274 | 4 | 0.85694 | 8 | 0.00516 | 1084 | 0.69890 |
| 01 | ceja | 0.86702 | 4 | 0.85694 | 18 | 0.01161 | 723 | 0.46615 |
| 01 | oído | 0.85913 | 4 | 0.85694 | 13 | 0.00838 | 851 | 0.54868 |
| 01 | órgano | 0.85694 | 4 | 0.85694 | 7 | 0.00451 | 1134 | 0.73114 |
| 01 | cara | 0.84609 | 3 | 0.76528 | 10 | 0.00645 | 1000 | 0.64475 |
| 01 | cerebro | 0.84392 | 3 | 0.76528 | 12 | 0.00774 | 876 | 0.56480 |
| 01 | codo | 0.81224 | 3 | 0.76528 | 18 | 0.01161 | 741 | 0.47776 |
| 01 | páncreas | 0.78089 | 3 | 0.76528 | 18 | 0.01161 | 777 | 0.50097 |
| 01 | rodilla | 0.77772 | 3 | 0.76528 | 22 | 0.01418 | 665 | 0.42876 |
| 01 | diente | 0.77321 | 3 | 0.76528 | 18 | 0.01161 | 759 | 0.48936 |
| 01 | sangre | 0.76715 | 3 | 0.76528 | 16 | 0.01032 | 809 | 0.52160 |
| 01 | hombro | 0.76653 | 3 | 0.76528 | 14 | 0.00903 | 838 | 0.54030 |
| 01 | cuello | 0.73645 | 2 | 0.61961 | 16 | 0.01032 | 793 | 0.51128 |
| 01 | esqueleto | 0.70370 | 2 | 0.61961 | 4 | 0.00258 | 1216 | 0.78401 |
| 01 | pestaña | 0.69882 | 2 | 0.61961 | 13 | 0.00838 | 864 | 0.55706 |
| 01 | cintura | 0.69805 | 2 | 0.61961 | 9 | 0.00580 | 1058 | 0.68214 |
| 01 | tibia | 0.69390 | 2 | 0.61961 | 10 | 0.00645 | 1030 | 0.66409 |
| 01 | antebrazo | 0.68921 | 2 | 0.61961 | 11 | 0.00709 | 935 | 0.60284 |
| 01 | lengua | 0.67137 | 2 | 0.61961 | 12 | 0.00774 | 912 | 0.58801 |
| 01 | tobillo | 0.66121 | 2 | 0.61961 | 15 | 0.00967 | 824 | 0.53127 |
| 01 | vello | 0.61961 | 2 | 0.61961 | 6 | 0.00387 | 1176 | 0.75822 |
| 01 | vena | 0.61347 | 1 | 0.46263 | 11 | 0.00709 | 990 | 0.63830 |
| 01 | enfermedad | 0.60225 | 1 | 0.46263 | 5 | 0.00322 | 1186 | 0.76467 |
| 01 | espalda | 0.59569 | 1 | 0.46263 | 12 | 0.00774 | 888 | 0.57253 |
| 01 | pecho | 0.59098 | 1 | 0.46263 | 12 | 0.00774 | 924 | 0.59574 |
| 01 | labio | 0.59024 | 1 | 0.46263 | 10 | 0.00645 | 1020 | 0.65764 |
| 01 | garganta | 0.56908 | 1 | 0.46263 | 7 | 0.00451 | 1120 | 0.72211 |
| 01 | peroné | 0.56348 | 1 | 0.46263 | 9 | 0.00580 | 1076 | 0.69375 |
| 01 | muñeca | 0.56318 | 1 | 0.46263 | 11 | 0.00709 | 968 | 0.62411 |
| 01 | bazo | 0.54471 | 1 | 0.46263 | 11 | 0.00709 | 946 | 0.60993 |
| 01 | esófago | 0.54466 | 1 | 0.46263 | 11 | 0.00709 | 957 | 0.61702 |
| 01 | culo | 0.53444 | 1 | 0.46263 | 7 | 0.00451 | 1113 | 0.71760 |
| 01 | fémur | 0.52785 | 1 | 0.46263 | 12 | 0.00774 | 900 | 0.58027 |
| 01 | torso | 0.51124 | 1 | 0.46263 | 4 | 0.00258 | 1260 | 0.81238 |
| 01 | salud | 0.50403 | 1 | 0.46263 | 4 | 0.00258 | 1248 | 0.80464 |
| 01 | barriga | 0.50263 | 1 | 0.46263 | 4 | 0.00258 | 1208 | 0.77885 |
| 01 | cadera | 0.48391 | 1 | 0.46263 | 9 | 0.00580 | 1049 | 0.67634 |
| 01 | falange | 0.46904 | 1 | 0.46263 | 10 | 0.00645 | 1010 | 0.65119 |
| 01 | piel | 0.46263 | 1 | 0.46263 | 11 | 0.00709 | 979 | 0.63121 |

Con esta información se puede construir una representación en la que se observa la distribución de las disponibilidades en el centro de interés y los diferentes conjuntos de cortes.

levels %>%  
 filter(centers=="04") %>%   
 mutate(level=factor(level)) %>%   
 arrange(-availability) %>%   
 ggplot(aes(x=order,y=availability,color=level)) + geom\_line() +  
 xlab("Posición del término en el centro de interés") +  
 ylab("Disponibilidad")

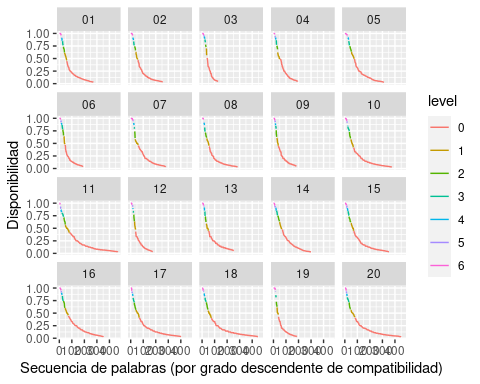


Se han construido utilidades para ayudar a la representación de estos conjuntos de cortes. El siguiente ejemplo muestra los principales conjuntos de corte obtenidos en los 16 centros de interés usados en el trabajo de Villena Ponsoda y Ávila Muñoz (2010)

clasificacion <- build.availability.levels(levels)  
clasificacion %>%   
 filter(level> 0) %>%  
 flextable() %>%   
 set\_header\_labels(centers = "Centro de interés",   
 words="Palabras",   
 level="Nivel de disponibilidad",  
 count="Recuento") %>%  
 width(j=c(1,2,3),width=.75) %>%  
 width(j=4,width=4) %>%  
 align(i=4,align="right") %>%  
 theme\_booktabs()

| Centro de interés | Nivel de disponibilidad | Recuento | Palabras |
| --- | --- | --- | --- |
| 01 | 6 | 17 | cabeza, brazo, ojo, mano, pierna, pie, corazón, dedo, nariz, boca, pelo, oreja, hueso, hígado, músculo, tronco, uña |
| 01 | 5 | 2 | pulmón, riñón |
| 01 | 4 | 5 | estómago, extremidad, ceja, oído, órgano |
| 01 | 3 | 8 | cara, cerebro, codo, páncreas, rodilla, diente, sangre, hombro |
| 01 | 2 | 9 | cuello, esqueleto, pestaña, cintura, tibia, antebrazo, lengua, tobillo, vello |
| 01 | 1 | 18 | vena, enfermedad, espalda, pecho, labio, garganta, peroné, muñeca, bazo, esófago, culo, fémur, torso, salud, barriga, cadera, falange, piel |
| 02 | 6 | 16 | pantalón, camisa, falda, camiseta, calcetín, jersey, vestido, zapato, abrigo, chaqueta, blusa, braga, calzoncillo, sujetador, bufanda, corbata |
| 02 | 5 | 1 | (pantalón) vaquero |
| 02 | 4 | 5 | moda, tejido, media, sombrero, bañador |
| 02 | 3 | 7 | zapatilla, chándal, traje, camisón, anorak, bermudas, tanga |
| 02 | 2 | 10 | cinturón, biquini, rebeca, gorro, color, pijama, bolso, guante, necesario, chaleco |
| 02 | 1 | 18 | estilo, vestir, chaquetón, pañuelo, felpa, rebajas, diseño, comodidad, gabardina, botón, lana, (pantalón) pirata, falda vaquera, vestimenta, bota, sudadera, slip, algodón |
| 03 | 6 | 12 | cocina, salón, (cuarto de) baño, dormitorio, comedor, terraza, salita, entrada, (cuarto de) aseo, pasillo, lavadero, ventana |
| 03 | 5 | 4 | hall, recibidor, (cuarto) trastero, habitación |
| 03 | 4 | 6 | tejado, jardín, techo, escalera, sala de estar, vestidor |
| 03 | 3 | 5 | patio, garaje, pared, porche, balcón |
| 03 | 2 | 5 | despensa, buhardilla, sótano, suelo, vestíbulo |
| 03 | 1 | 5 | chimenea, puerta, despacho, azotea, estudio |
| 04 | 6 | 11 | mesa, silla, cama, sofá, sillón, armario, mesita de noche, aparador, cómoda, estantería, lámpara |
| 04 | 5 | 5 | ropero, mesa de comedor, espejo, vitrina, encimera |
| 04 | 4 | 4 | cuadro, lavadora, librería, frigorífico |
| 04 | 3 | 5 | tresillo, mueble de cocina, mueble-bar, lavavajillas, zapatero |
| 04 | 2 | 8 | butaca, tele(visión), mesilla de noche, escritorio, repisa, taquillón, tocador, cajonera |
| 04 | 1 | 20 | sinfonier, hornilla, lámpara de techo, mueble de salón, alacena, mesa camilla, cama de matrimonio, mesa de tele(visión), taburete, comodín, mecedora, televisor, mesa de salón, ordenador, cortina, armario de niños, barbacoa, cenefa, mueble de entrada, rinconero |
| 05 | 6 | 21 | pan, carne, leche, tomate, huevo, lenteja, arroz, patata, garbanzo, verdura, pescado, fruta, manzana, cebolla, lechuga, agua, plátano, pera, pimiento, mantequilla, naranja |
| 05 | 5 | 2 | aceite, pasta |
| 05 | 4 | 7 | queso, jamón, melón, sandía, azúcar, pollo, judía |
| 05 | 3 | 7 | yogur, melocotón, chocolate, legumbre, fresa, carne de cerdo, alcachofa |
| 05 | 2 | 15 | zanahoria, carne de ternera, ensalada, maíz, café, uva, coliflor, calabacín, harina, chorizo, espinaca, hamburguesa, espagueti, sopa, bocadillo |
| 05 | 1 | 24 | kiwi, refresco, salchichón, cerdo, macarrón, berenjena, galleta, carne de pollo, vino, zumo, dulce, gazpacho, hortaliza, sal, cereal, fideo, filete, limón, mandarina, pepino, col, calabaza, fibra, trigo |
| 06 | 6 | 10 | tenedor, mantel, cuchara, cuchillo, plato, servilleta, vaso, plato llano, plato hondo, ensaladera |
| 06 | 5 | 5 | copa, cucharilla, botella, cuchillo de carne, cuchillo de pescado |
| 06 | 4 | 6 | plato de postre, fuente, tenedor de pescado, jarra, cubierto, sopera |
| 06 | 3 | 5 | cucharón (de servir), salero, tenedor de carne, cuchara sopera, pan |
| 06 | 2 | 8 | salvamanteles, panera, bandeja, cuenco, taza, plato sopero, servilletero, jarra de agua |
| 06 | 1 | 11 | copa de vino, vela, pimentero, cuchara de postre, cucharilla de postre, bol, candelabro, centro floral, tenedor de ensalada, frutero, vaso de vino |
| 07 | 6 | 20 | sartén, horno, hornilla, micro(ondas), olla, cacerola, frigorífico, lavadora, cuchillo, nevera, cucharón, espumadera, fregadero, olla exprés, (placa) vitro(cerámica), plato, lavavajillas, mesa, encimera, cuchara |
| 07 | 5 | 3 | cazo, mueble, batidora |
| 07 | 4 | 4 | silla, congelador, tenedor, cafetera |
| 07 | 3 | 3 | vaso, paleta, armario |
| 07 | 2 | 5 | colador, freidora, tostadora, cacillo, cazuela (de barro) |
| 07 | 1 | 24 | placa, despensa, plato hondo, grifo, escurridor, bandeja, tortilla, cocina, cubierto, poyo, especiero, exprimidor, salero, barra americana, cocinar, cortina, cuchillo de servir, fluorescente, lavadero, mueble bajo, tabla para picar, paño, plancha, cubo de (la) basura |
| 08 | 6 | 11 | pizarra, mesa, pupitre, silla, lápiz, libro, tiza, bolí(grafo), goma (de borrar), libreta, cuaderno |
| 08 | 5 | 2 | ordenador, sacapuntas |
| 08 | 4 | 7 | folio, mesa del profesor, borrador (de tiza), lápiz de color, aula, regla, pluma |
| 08 | 3 | 4 | rotulador, lapicero, mapa, carpeta |
| 08 | 2 | 13 | tarima, mochila, plastilina, banca, hoja (de papel), puerta, cuadro, cartabón, escuadra, profesor, banco, lápiz de cera, libro de matemáticas |
| 08 | 1 | 18 | compás, papel, clase, estuche, armario, estantería, ventana, timbre, diccionario, proyector, tintero, calculadora, bloc, curso, libro de conocimiento del medio, librería, maleta, papelera |
| 09 | 6 | 12 | aire acondicionado, ventilador, lámpara, bombilla, estufa, radiador, (aparato) calefactor, ventana, chimenea, calefacción, caldera, vela |
| 09 | 5 | 2 | abrir ventanas, linterna |
| 09 | 4 | 3 | foco, brasero, abanico |
| 09 | 3 | 4 | flexo, fluorescente, (foco) halógeno, extractor (de humos) |
| 09 | 2 | 5 | electricidad, luz, abrir puertas, climatizador, carbón |
| 09 | 1 | 12 | puerta, tubo fluorescente, refrigeración, estufa eléctrica, lámpara de pie, leña, interruptor, calentador, aparato de calefacción, foco de jardín, bombilla de bajo consumo, cerilla |
| 10 | 6 | 17 | calle, coche, edificio, carretera, plaza, semáforo, parque, jardín, avenida, acera, farola, tienda, casa, (auto)bús, tráfico, ruido, moto(cicleta) |
| 10 | 5 | 4 | árbol, banco, museo, gente |
| 10 | 4 | 3 | callejón, bici(cleta), iglesia |
| 10 | 3 | 12 | polución, colegio, papelera, bar, fuente, bloque de pisos, ayuntamiento, paso de peatones, cine, contaminación, catedral, teatro |
| 10 | 2 | 9 | piso, asfalto, centro comercial, restaurante, policía, paso de cebra, puerto, humo, trabajo |
| 10 | 1 | 24 | flor, parada de (auto)bús, persona, peatón, atasco, comercio, barrio, chalé, población, camión, calzada, distrito, especulación, lugar para vivir, playa, rotonda, polideportivo, perro, paseo, hospital, pasaje, señal de tráfico, metro, campo de fútbol |
| 11 | 6 | 8 | árbol, flor, hierba, río, tierra, planta, animal, montaña |
| 11 | 5 | 4 | olivo, verde, vaca, gallina |
| 11 | 4 | 9 | pájaro, insecto, caballo, granja, monte, piedra, casa, cerdo, huerta |
| 11 | 3 | 13 | tractor, arado, paisaje, cabra, agua, tranquilidad, naranjo, bosque, agricultura, alegría, peral, aire puro, abeja |
| 11 | 2 | 13 | conejo, césped, camino, cultivo, sembrado, olivar, fruta, naturaleza, lago, mariposa, árbol frutal, toro, sol |
| 11 | 1 | 35 | pradera, aire limpio, oveja, vegetación, campo, terreno, manzano, avispa, hormiga, campesino, alergia, brisa, descanso, hierba seca, hierba verde, suelo, plantación, carril, limonero, paella, cortijo, roca, rama, arroyo, aire, arbusto, matorral, rosa, almendro, burro, margarita, amapola, vereda, pantano, invernadero |
| 12 | 6 | 12 | coche, avión, (auto)bús, bici(cleta), tren, moto(cicleta), barco, patín, metro, helicóptero, caballo, camión |
| 12 | 5 | 4 | avioneta, tranvía, carro, patinete |
| 12 | 4 | 2 | taxi, a pie |
| 12 | 3 | 5 | furgoneta, monopatín, globo (aerostático), coche de caballos, carreta |
| 12 | 2 | 6 | automóvil, burro, barca, triciclo, autocar, lancha |
| 12 | 1 | 10 | camioneta, andar, tractor, (tren de) cercanías, (tren) AVE, ciclomotor, gasoil, quad, moto acuática, sidecar |
| 13 | 6 | 12 | arar, sembrar, regar, podar, plantar, segar, recolectar, jardinero, abonar, labrar, cortar césped, cavar |
| 13 | 5 | 5 | recoger, agricultor, fumigar, siembra, talar |
| 13 | 4 | 3 | cortar, injertar, recolección |
| 13 | 3 | 4 | cultivar, recoger aceitunas, trasplantar, trillar |
| 13 | 2 | 7 | poda, cosechar, vendimiar, pastor, riego, arado, podador |
| 13 | 1 | 20 | varear, ganadero, hortelano, sembrar flores, limpiar, siega, recolector, roturar, arreglar flores, botánica, huerta, huerto, jardinería, recogida de algodón, segadora, recoger almendras, aventar, quitar malas hierbas, echar abono, regador |
| 14 | 6 | 11 | perro, gato, león, tigre, caballo, elefante, vaca, toro, pájaro, cerdo, conejo |
| 14 | 5 | 6 | ratón, oveja, jirafa, oso, canario, leopardo |
| 14 | 4 | 6 | burro, serpiente, gallina, rata, paloma, águila |
| 14 | 3 | 10 | mono, cebra, cabra, rinoceronte, abeja, pato, loro, ardilla, cucaracha, ballena |
| 14 | 2 | 14 | ave, avestruz, pez, gorrión, hormiga, buey, delfín, jabalí, cordero, periquito, búfalo, pollo, tortuga, camello |
| 14 | 1 | 21 | araña, lobo, ciervo, hipopótamo, pantera, guepardo, zoo(lógico), tiburón, hámster, cocodrilo, foca, pingüino, lagartija, lombriz, dogo, felino, liebre, pitón, murciélago, mosquito, gallo |
| 15 | 6 | 14 | parchís, fútbol, oca, ajedrez, cartas, dominó, baloncesto, cine, tenis, póquer, pilla-pilla, leer, tele(visión), escondite |
| 15 | 5 | 4 | pasear, mus, balonmano, comba |
| 15 | 4 | 6 | videojuego, Monopoly, damas, teatro, cinquillo, solitario |
| 15 | 3 | 9 | leer libros, juego de mesa, crucigrama, lotería, música, bailar, tute, internet, deporte |
| 15 | 2 | 14 | tres en raya, natación, puzle, pelota, correr, dados, golf, volei(bol), Play(-Station), naipe, Trivial, nadar, ping-pong, lectura |
| 15 | 1 | 23 | ordenador, playa, sudoku, muñeco, Tetris, canasta, escribir, rueda, baile, saltar, baraja, brisca, senderismo, gallinita ciega, guiso, balón, coleccionar sellos, hacer el tonto, hobby, parque acuático, PS2, radio, billar |
| 16 | 6 | 14 | médico/a, profesor/ora, albañil, carpintero/a, maestro/a, abogado/a, fontanero, electricista, enfermero/a, jardinero, ingeniero, camarero/a, mecánico, bombero |
| 16 | 5 | 3 | dependiente/a, policía, pintor |
| 16 | 4 | 6 | arquitecto/a, cocinero/a, conductor/ora, panadero, juez/eza, peluquero/a |
| 16 | 3 | 10 | carnicero, zapatero, chapista, celador, ama de casa, escayolista, secretario/a, vendedor/ora, administrativo, informático |
| 16 | 2 | 12 | obrero, economista, administrador, limpiador/ora, escritor, farmacéutico/a, comercial, policía local, ingeniero de caminos, psicólogo, solador, agricultor |
| 16 | 1 | 27 | torero, modisto[a], pescadero, ebanista, dentista, taxista, tapicero, camionero, notario/a, representante, frigorista, delineante, óptico/a, cajero/a, fiscal, auxiliar de enfermería, cetrero, doctor, ingeniero de minas, maestro de música, peón, piloto, técnico, perito, sastre, maquinista, ingeniero agrónomo |
| 17 | 6 | 10 | euro, banco, moneda, bolsa, billete, peseta, dólar, hipoteca, libra, acción |
| 17 | 5 | 2 | cuenta corriente, céntimo |
| 17 | 4 | 6 | economista, comprar, cheque, ahorro, inversión, franco |
| 17 | 3 | 10 | préstamo, caja de ahorros, interés, marco, letra (de cambio), lira (italiana), IVA, cambio, cajero, tarjeta |
| 17 | 2 | 13 | banquero, necesario, trabajo, monedero, dinero, cartilla, factura, íbex 35, comercio, cartera, IPC, divisa, oferta |
| 17 | 1 | 27 | sueldo, hacienda, gastar, gasto, trueque, transacción, ruina, vender, impuesto, tarjeta de crédito, deuda, contable, importante, llegar a fin de mes, piso, pobreza, warrant, libreta, talón, demanda, banca, coche, yen, mercado, rublo, finanza, bancarrota |
| 18 | 6 | 11 | ratón, pantalla, teclado, ordenador, chatear, Messenger, torre, página web, correo (electrónico), impresora, disco duro |
| 18 | 5 | 4 | programa, monitor, chat, Google |
| 18 | 4 | 8 | red, PC, Web, información, navegar, software, Chat, Windows |
| 18 | 3 | 9 | altavoz, internet, web, mensaje, escáner, ADSL, virus, CD-ROM, placa base |
| 18 | 2 | 15 | computadora, bit, micro(fono), juego, E-mail, conexión, router, Internet, webcam, hardware, archivo, Bill Gates, Software, buscar trabajo, E-Mule |
| 18 | 1 | 26 | grabadora, Microsoft, Pentium, amigo, cable, informática, automatismo, desesperación, Hispachat, interesante, Java, nueva era, punto com, procesador, tecla, CPU, invento, novio, Terra, comunicación, cámara web, ventilador, alfombrilla, foro, portátil, Basic |
| 19 | 6 | 15 | rojo, azul, blanco, amarillo, verde, negro, rosa, violeta, naranja, celeste, marrón, gris, morado, arco iris, añil |
| 19 | 5 | 2 | beis, fucsia |
| 19 | 4 | 1 | turquesa |
| 19 | 3 | 4 | azul marino, primario, cálido, magenta |
| 19 | 2 | 6 | burdeos, pastel, secundario, luz, lila, malva |
| 19 | 1 | 11 | ocre, alegría, cian, plata, azul celeste, cromatismo, tono, violeta claro, granate, dorado, verde botella |
| 20 | 6 | 11 | ola, agua, barco, pez, arena, sal, playa, alga, mar, roca, azul |
| 20 | 5 | 4 | orilla, tiburón, pesca, marea |
| 20 | 4 | 4 | espuma, oleaje, océano, pescador |
| 20 | 3 | 14 | ballena, barca, nadar, horizonte, velero, arrecife, navegar, paz, boquerón, coral, sol, calamar, tranquilidad, delfín |
| 20 | 2 | 13 | maremoto, concha, color, inmenso, mejillón, medusa, verano, almeja, caracola, gaviota, resaca, barco de vela, vida |
| 20 | 1 | 34 | acantilado, viento, sombrilla, rompeolas, red, moto de agua, fondo, lancha (motora), brisa, puerto, sardina, bucear, piedra, pescado, bonito, mar tranquilo, Resolution, veranear, cabo, yodo, pulpo, pescar, marinero, paseo, sunami, crucero, submarino, flotador, espigón, golfo, relajación, arrecife de coral, chiringuito, marejada |

levels %>%  
 mutate(level=factor(level)) %>%   
 arrange(-availability) %>%   
 ggplot(aes(x=order,y=availability,color=level)) + geom\_line() + facet\_wrap(~centers) +  
 xlab("Secuencia de palabras (por grado descendente de compatibilidad)") +   
 ylab("Disponibilidad")

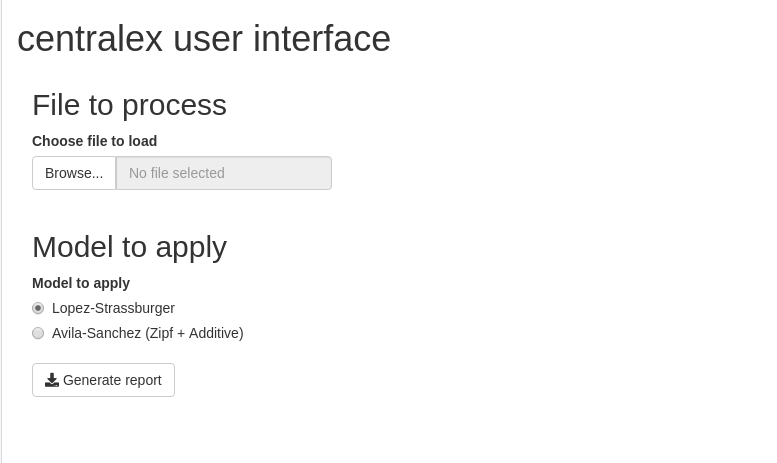


# Herramienta propuesta

Para facilitar el acceso a esta herramienta se ha dispuesto una pequeña utilidad gráfica que construye un documento básico con construcciones que opinamos que cubren las necesidades generales para el estudio de la disponibilidad. Este documento puede ser fácilmente modificable para adaptarlo a las necesidades de cada usuario.

Incluyendo la carga de la librería se puede invocar como:

library(centralex)  
centralex::runUtility()

abriéndose un interfaz gráfico con el siguiente aspecto:  Proporcionado un archivo de datos, seleccionado el modelo de procesamiento (inicialmente se proporcionan dos opciones, el modelo de disponibilidad propuesto por López-Strassburger y el de centralidad propuesto por Ávila-Sánchez) se confirman los datos y se ofrece la descarga de un documento que ha de guardarse junto al archivo de datos.

Este documento incluye los bloques de código que llevan a cabo los procedimientos que hemos considerado útiles y que ofrecen sus resultados en el formato expuesto en este trabajo. Muchos bloques están etiquetados con la opción eval=FALSE, que hace que no se ejecuten por defecto. Si se está interesado en los resultados correspondientes, tan sólo hay que eliminar esa opción.

Se proporcionan las siguientes secciones:

* Carga de datos y modelo de disponibilidad: estas secciones son imprescindibles, o no se podría lleva a cabo el análisis. Se proporcionan una serie de bloques que permiten exponer los datos obtenidos por centro de interés, en forma de tabla y de gráfico, y un resumen gráfico de todos los centros de interés.
* Niveles de centralidad/disponibilidad: se clasifican las palabras por niveles y se construyen los listados de términos por nivel. Se posibilita obtener los listados con la clasificación y otros datos asociados y las tablas resumen. Se han habilitado por defecto la presentación de los niveles de los términos por encima del nivel básico, de menor centralidad.

Nótese que este documento es un documento RMarkdown que es fácilmente modificable y personalizable. Se procesa entonces, dando lugar a un documento PDF, Word o HTML, entre otras opciones, que puede utilizarse como base para completar el trabajo.

# Ejemplos de uso

Hemos decidido, de forma predeterminada, no considerar el uso de las primeras columnas para datos sociológicos en los datos porque su limitado formato y redundancia no constituye una buena práctica, siguiendo los criterios de la normalización de bases de datos y de análisis de datos.

Así, presentamos algunas propuestas que consideramos más productivas para llevar a cabo estos datos.

## Los datos sociológicos

En primer lugar, es importante evitar la redundancia de los datos, tanto por la dificultad que implica mantener la coherencia de los mismos cuando se encuentran duplicados en varios lugares como por facilitar mantenerlos actualizados.

bs <- read.csv("BaseSociologica.csv", colClasses = "character")  
bs <- bs %>%  
 mutate(EDAD=strtoi(EDAD),  
 AÑOSDEESTUDIO = strtoi(AÑOSDEESTUDIO))

bs %>%  
 head() %>%  
 select(SUJETO,SEXO,EDAD,AÑOSDEESTUDIO) %>%  
 flextable() %>%  
 autofit()

| SUJETO | SEXO | EDAD | AÑOSDEESTUDIO |
| --- | --- | --- | --- |
| 001 | 1 | 69 | 4 |
| 002 | 0 | 71 | 4 |
| 003 | 0 | 20 | 12 |
| 004 | 1 | 24 | 17 |
| 005 | 0 | 24 | 17 |
| 006 | 1 | 26 | 15 |

## Análisis separados por sexo

Si se desea llevar a cabo un análisis de centralidad léxica, separados por sexos, se pretende llevar a cabo la integración de la información proporcionada por hombres y mujeres. Así que se filtrarían las realizaciones a procesar seleccionando a los sujetos que cumplen esas características:

idHombres <-  
 bs %>%   
 filter(SEXO=="1") %>%  
 select(SUJETO) %>%  
 unlist()  
idHombres

## SUJETO1 SUJETO2 SUJETO3 SUJETO4 SUJETO5 SUJETO6 SUJETO7 SUJETO8   
## "001" "004" "006" "010" "012" "014" "015" "022"   
## SUJETO9 SUJETO10 SUJETO11 SUJETO12 SUJETO13 SUJETO14 SUJETO15 SUJETO16   
## "026" "029" "032" "035" "036" "039" "041" "043"   
## SUJETO17 SUJETO18 SUJETO19 SUJETO20 SUJETO21 SUJETO22 SUJETO23 SUJETO24   
## "046" "049" "050" "051" "052" "053" "055" "058"   
## SUJETO25 SUJETO26 SUJETO27 SUJETO28 SUJETO29 SUJETO30 SUJETO31 SUJETO32   
## "061" "064" "065" "070" "073" "075" "076" "097"   
## SUJETO33 SUJETO34   
## "098" "099"

idMujeres <-  
 bs %>%   
 filter(SEXO=="0") %>%  
 select(SUJETO) %>%  
 unlist()  
idMujeres

## SUJETO1 SUJETO2 SUJETO3 SUJETO4 SUJETO5 SUJETO6 SUJETO7 SUJETO8   
## "002" "003" "005" "008" "009" "013" "019" "021"   
## SUJETO9 SUJETO10 SUJETO11 SUJETO12 SUJETO13 SUJETO14 SUJETO15 SUJETO16   
## "023" "024" "025" "027" "028" "030" "031" "033"   
## SUJETO17 SUJETO18 SUJETO19 SUJETO20 SUJETO21 SUJETO22 SUJETO23 SUJETO24   
## "034" "037" "038" "040" "042" "045" "047" "048"   
## SUJETO25 SUJETO26 SUJETO27 SUJETO28 SUJETO29 SUJETO30 SUJETO31 SUJETO32   
## "056" "059" "060" "062" "063" "066" "067" "068"   
## SUJETO33 SUJETO34 SUJETO35 SUJETO36 SUJETO37 SUJETO38   
## "069" "071" "072" "074" "077" "096"

Entonces, se pueden separar las realizaciones, si los sujetos pertenecen al grupo de hombres o mujeres

data %>%   
 filter(users %in% idMujeres) %>%  
 head()

## infos users centers words  
## 1 12131 002 01 riñón, c....  
## 2 12213 003 01 brazo, m....  
## 3 12214 005 01 cabeza, ....  
## 4 11232 008 01 cabeza, ....  
## 5 11221 009 01 músculo,....  
## 6 12224 013 01 cabeza, ....

data %>%   
 filter(users %in% idHombres) %>%  
 head()

## infos users centers words  
## 1 21131 001 01 mano, pi....  
## 2 22214 004 01 brazo, o....  
## 3 22213 006 01 pie, man....  
## 4 22224 010 01 esquelet....  
## 5 22224 012 01 cabeza, ....  
## 6 22212 014 01 mano, pi....

Y se pueden procesar de la forma habitual:

disponibilidadHombres <-   
 build.avilasanchez.availability(data %>%  
 filter(users %in% idHombres))  
head(disponibilidadHombres) %>%  
 select(centers, words, availability) %>%  
 flextable() %>%  
 autofit()

| centers | words | availability |
| --- | --- | --- |
| 01 | cabeza | 0.9999756 |
| 01 | mano | 0.9987419 |
| 01 | brazo | 0.9984684 |
| 01 | ojo | 0.9978411 |
| 01 | pierna | 0.9907053 |
| 01 | pie | 0.9894805 |

disponibilidadMujeres <-   
 build.avilasanchez.availability(data %>%  
 filter(users %in% idMujeres))  
head(disponibilidadMujeres) %>%  
 select(centers, words, availability) %>%  
 flextable() %>%  
 autofit()

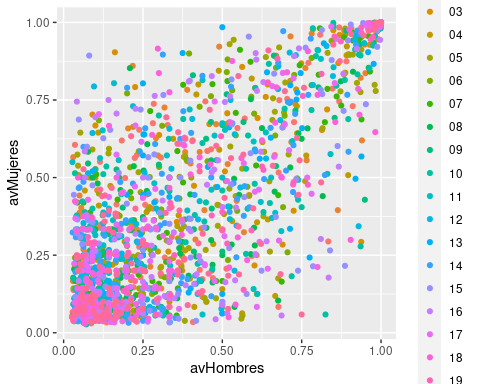
| centers | words | availability |
| --- | --- | --- |
| 01 | brazo | 0.9999198 |
| 01 | ojo | 0.9998306 |
| 01 | cabeza | 0.9986111 |
| 01 | mano | 0.9970720 |
| 01 | pierna | 0.9963897 |
| 01 | pie | 0.9926613 |

Esta información se puede combinar para realizar análisis comparativos, enlazando los datos por centros de interés y términos

dispComp <-  
 inner\_join(disponibilidadHombres %>%   
 select(centers,words,availability) %>%  
 rename(avHombres=availability),   
 disponibilidadMujeres %>%  
 select(centers,words,availability) %>%  
 rename(avMujeres=availability),  
 by = c("centers", "words"))  
dispComp %>%  
 head() %>%  
 flextable() %>%  
 autofit()

| centers | words | avHombres | avMujeres |
| --- | --- | --- | --- |
| 01 | cabeza | 0.9999756 | 0.9986111 |
| 01 | mano | 0.9987419 | 0.9970720 |
| 01 | brazo | 0.9984684 | 0.9999198 |
| 01 | ojo | 0.9978411 | 0.9998306 |
| 01 | pierna | 0.9907053 | 0.9963897 |
| 01 | pie | 0.9894805 | 0.9926613 |

dispComp %>%  
 ggplot(aes(x=avHombres,y=avMujeres,color=centers)) + geom\_point()



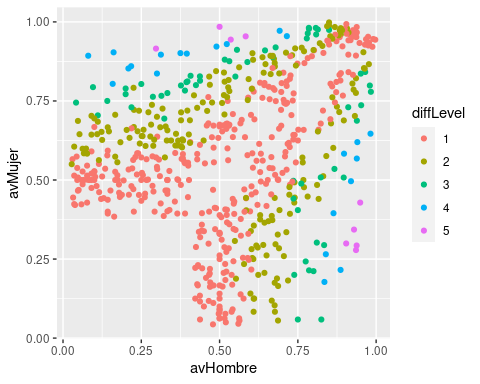
Según este gráfico es posible observar que los términos muy disponibles y poco disponibles tienden a ser similares. Sin embargo, los términos que aparecen en los diversos niveles intermedios pueden variar de forma muy significativa.

levelsHombres <- classify.availability.levels(disponibilidadHombres)  
clasificacionHombres <- build.availability.levels(levelsHombres)  
levelsMujeres <- classify.availability.levels(disponibilidadMujeres)  
clasificacionMujeres <- build.availability.levels(levelsMujeres)  
clasificacionSexo <-  
 inner\_join(clasificacionHombres %>%  
 select(centers,level,words) %>%  
 rename(wordsHombres=words),  
 clasificacionMujeres %>%   
 select(centers,level,words) %>%  
 rename(wordsMujeres=words),   
 by=c("centers","level"))   
   
clasificacionSexo %>%  
 filter(level> 0) %>%  
 flextable() %>%   
 set\_header\_labels(centers = "Centro de interés",   
 wordsHombres="Palabras hombres",   
 wordsMujeres="Palabras mujeres",   
 level="Nivel de disponibilidad",  
 count="Recuento") %>%  
 width(j=c(1,2),width=.75) %>%  
 width(j=c(3,4),width=2.5) %>%  
 align(j=c(3,4),align="right") %>%  
 theme\_booktabs()

| Centro de interés | Nivel de disponibilidad | Palabras hombres | Palabras mujeres |
| --- | --- | --- | --- |
| 01 | 6 | cabeza, mano, brazo, ojo, pierna, pie, corazón | brazo, ojo, cabeza, mano, pierna, pie, nariz, dedo, corazón, pelo, boca |
| 01 | 5 | hueso |  |
| 01 | 4 | dedo, oreja, boca, tronco, nariz | uña, oreja, hígado |
| 01 | 3 | músculo, extremidad, pelo, hígado | músculo, pulmón, riñón, oído, tronco |
| 01 | 2 | estómago, pulmón, riñón, cerebro, esqueleto, uña, órgano | codo, hueso, ceja, estómago, hombro, rodilla, cara, cuello, antebrazo, vello |
| 01 | 1 | tibia, sangre, ceja, cara, vena, pestaña, peroné, páncreas, diente, oído, barriga | páncreas, diente, lengua, órgano, cintura, tobillo, enfermedad, cerebro, espalda, labio, sangre |
| 02 | 6 | pantalón, falda, camisa, vestido, camiseta, jersey, calcetín, blusa | camisa, pantalón, camiseta, calcetín, falda, zapato, jersey, chaqueta, calzoncillo |
| 02 | 5 | abrigo, sujetador, braga | vestido, abrigo, corbata |
| 02 | 4 | zapato, moda | sombrero |
| 02 | 3 | chaqueta, tejido, bufanda | bufanda, blusa, braga, bañador |
| 02 | 2 | (pantalón) vaquero, media, calzoncillo, estilo, bolso, bermudas | (pantalón) vaquero, traje, zapatilla, chándal, gorro, rebeca, vestir |
| 02 | 1 | pijama, biquini, camisón, zapatilla, comodidad, falda vaquera, vestimenta, corbata, chándal, diseño, (pantalón) pirata, chaleco, chaquetón | anorak, sujetador, media, tanga, cinturón, guante, felpa, tejido, color, camisón, rebajas, necesario |
| 03 | 6 | cocina, salón, (cuarto de) baño, dormitorio, terraza, comedor, salita | cocina, salón, dormitorio, (cuarto de) baño, comedor, terraza, entrada |
| 03 | 5 | recibidor | (cuarto de) aseo, salita |
| 03 | 4 | pasillo, (cuarto de) aseo | techo, pasillo, hall |
| 03 | 3 | lavadero, entrada, (cuarto) trastero, jardín | lavadero, ventana, tejado |
| 03 | 2 | ventana, sala de estar, habitación, porche, vestidor, despensa | pared, habitación, escalera, (cuarto) trastero, garaje, vestíbulo, patio |
| 03 | 1 | hall, escalera, patio, despacho, tejado | balcón, chimenea, suelo, vestidor, jardín, sala de estar, sótano |
| 04 | 6 | mesa, cama, silla, sofá, sillón, armario | silla, mesa, sofá, sillón, cama, aparador, armario, mesita de noche |
| 04 | 5 |  | cómoda |
| 04 | 4 | estantería, ropero, cómoda, lámpara, mesita de noche | estantería |
| 04 | 3 | aparador, lavadora, frigorífico | cuadro, lámpara, encimera, mesa de comedor, librería, espejo |
| 04 | 2 | vitrina, mesa de comedor, taquillón, zapatero, espejo, mueble de cocina, mesilla de noche | tresillo, vitrina, mueble-bar, cajonera, butaca |
| 04 | 1 | encimera, lámpara de techo, escritorio, tocador, mueble-bar, tele(visión), alacena, armario de niños, barbacoa, mueble de entrada, lavavajillas, cuadro, repisa | mueble de cocina, lavavajillas, lavadora, televisor, ropero, mecedora, cama de matrimonio, cenefa, comodín, rinconero, tele(visión) |
| 05 | 6 | pan, leche, carne, lenteja, arroz | pan, carne, leche, tomate, huevo, plátano |
| 05 | 5 | patata, fruta, huevo, manzana, garbanzo | verdura, pimiento, agua, patata |
| 05 | 4 | pescado, tomate, lechuga, pera, verdura | cebolla, pescado, garbanzo, mantequilla, naranja |
| 05 | 3 | cebolla, pasta | lechuga, aceite, manzana, arroz, jamón, lenteja, pera, fruta, queso |
| 05 | 2 | judía, aceite, melón, azúcar, agua, maíz, mantequilla, sandía, queso, pollo, carne de ternera, chocolate, naranja, yogur, carne de cerdo | sandía, melón, pollo, azúcar, pasta, hamburguesa, melocotón, legumbre |
| 05 | 1 | jamón, alcachofa, berenjena, fresa, carne de pollo, melocotón, plátano, legumbre, pimiento, zanahoria, harina, fibra, trigo, espagueti, calabacín, filete | sopa, refresco, espinaca, yogur, bocadillo, judía, cerdo, ensalada, café, uva, gazpacho, fresa, kiwi, chorizo, coliflor, chocolate, zanahoria, galleta |
| 06 | 6 | mantel, tenedor, servilleta, cuchara, cuchillo, plato, vaso | tenedor, cuchara, cuchillo, plato, mantel, vaso, servilleta |
| 06 | 5 | plato llano | plato llano, plato hondo |
| 06 | 4 | plato hondo | cuchillo de carne, ensaladera, botella |
| 06 | 3 | ensaladera, cucharón (de servir) | cucharilla, cuchillo de pescado, copa, tenedor de pescado, tenedor de carne |
| 06 | 2 | copa, pan, plato de postre, cuchara sopera, cucharilla, cubierto, jarra de agua | sopera, fuente, salero, jarra, plato de postre, plato sopero, panera |
| 06 | 1 | jarra, fuente, salvamanteles, taza, tenedor de pescado, cuchara de postre, botella, cuchillo de pescado, sopera | cubierto, cuenco, bandeja, servilletero, pimentero, vela, candelabro, centro floral, tenedor de ensalada |
| 07 | 6 | sartén, horno, hornilla, micro(ondas), olla, espumadera, cacerola | sartén, micro(ondas), horno, cuchillo, olla, hornilla, frigorífico, cacerola, lavadora |
| 07 | 5 | olla exprés, lavadora, nevera | fregadero, nevera, cucharón |
| 07 | 4 | frigorífico, cucharón, fregadero, lavavajillas, plato | cuchara, mueble, mesa |
| 07 | 3 | encimera, batidora, cuchillo, (placa) vitro(cerámica), cazo | espumadera, (placa) vitro(cerámica), silla, congelador, tenedor, cazo, encimera |
| 07 | 2 | mesa, colador, vaso, paleta | plato, lavavajillas, armario, cafetera, olla exprés, batidora |
| 07 | 1 | mueble, cuchara, cafetera, tostadora, silla, barra americana, cocinar, cortina, mueble bajo, placa, tabla para picar, plato hondo, congelador | freidora, paleta, tortilla, cocina, poyo, grifo, cuchillo de servir, fluorescente, lavadero, cacillo, cubo de (la) basura, despensa |
| 08 | 6 | mesa, pizarra, silla, pupitre, lápiz, libro, tiza, goma (de borrar), bolí(grafo) | pizarra, pupitre, mesa, lápiz, libro, silla, tiza, bolí(grafo) |
| 08 | 5 | libreta |  |
| 08 | 4 | cuaderno | goma (de borrar), aula, libreta |
| 08 | 3 | lápiz de color, sacapuntas, mesa del profesor, folio | cuaderno, borrador (de tiza), ordenador |
| 08 | 2 | pluma, ordenador, plastilina, rotulador, regla, tarima, carpeta | banca, regla, sacapuntas, folio, mesa del profesor, mochila, libro de matemáticas |
| 08 | 1 | borrador (de tiza), hoja (de papel), mapa, clase, lapicero, cuadro, banco, librería, ventana, papel, lápiz de cera, profesor | escuadra, pluma, cartabón, curso, libro de conocimiento del medio, timbre, lapicero, proyector, estantería, compás, rotulador, puerta, mapa, lápiz de color |
| 09 | 6 | aire acondicionado, ventilador, lámpara, estufa, bombilla, radiador | ventilador, aire acondicionado, bombilla, lámpara, estufa, (aparato) calefactor |
| 09 | 5 | calefacción, ventana | radiador, ventana |
| 09 | 4 | (aparato) calefactor | chimenea, caldera |
| 09 | 3 | abrir ventanas, vela, chimenea | foco, vela, brasero |
| 09 | 2 | linterna, flexo, caldera | linterna, abrir ventanas, abanico, (foco) halógeno, fluorescente, climatizador, extractor (de humos) |
| 09 | 1 | abanico, brasero, carbón, foco de jardín, leña, fluorescente, luz, electricidad, (foco) halógeno | electricidad, estufa eléctrica, abrir puertas, flexo, tubo fluorescente, calentador, aparato de calefacción, refrigeración |
| 10 | 6 | calle, coche, carretera, parque | calle, coche, edificio, plaza, semáforo |
| 10 | 5 | ruido, jardín, acera, farola, avenida | carretera, avenida, acera |
| 10 | 4 | tráfico, tienda, (auto)bús, plaza, edificio | jardín, casa, moto(cicleta), parque, farola |
| 10 | 3 | banco, polución, casa, semáforo, árbol, callejón, bici(cleta) | tienda, (auto)bús, museo, gente, árbol |
| 10 | 2 | contaminación, gente, bar, moto(cicleta), paso de peatones, humo, museo, fuente | bloque de pisos, colegio, tráfico, iglesia, catedral, ayuntamiento, cine |
| 10 | 1 | iglesia, teatro, papelera, asfalto, piso, cine, flor, centro comercial, restaurante, ayuntamiento, comercio, barrio | trabajo, papelera, parada de (auto)bús, población, chalé, banco, peatón, distrito, especulación, lugar para vivir, polideportivo, fuente, policía, piso, paso de cebra, bici(cleta), campo de fútbol, paseo, centro comercial, bar |
| 11 | 6 | árbol, flor, hierba, tierra | árbol, flor, río, hierba |
| 11 | 5 | verde, río | tierra, planta |
| 11 | 4 | animal, olivo, montaña | montaña, animal, vaca, monte |
| 11 | 3 | planta, peral, aire puro, insecto, gallina | arado, bosque, agricultura, cabra, casa |
| 11 | 2 | agua, paisaje, tranquilidad, huerta, pájaro, naranjo, olivar, caballo, fruta, piedra | granja, cerdo, cultivo, pájaro, piedra, abeja, tractor, gallina, caballo, olivo, césped, toro, aire limpio, lago, pradera |
| 11 | 1 | terreno, camino, cerdo, alegría, brisa, granja, hierba seca, hierba verde, suelo, tractor, conejo, vaca, roca, avispa, manzano, sembrado, limonero, amapola, corral, casa | alegría, alergia, campo, descanso, huerta, plantación, insecto, árbol frutal, conejo, oveja, naturaleza, sembrado, naranjo, invernadero, secano, lluvia, paella, sol, mariposa |
| 12 | 6 | coche, (auto)bús, avión, bici(cleta), tren, moto(cicleta), barco | coche, avión, bici(cleta), (auto)bús, tren, moto(cicleta), barco |
| 12 | 5 | carro | patín, metro |
| 12 | 4 | patín | caballo, camión, tranvía |
| 12 | 3 | helicóptero, metro, avioneta | helicóptero, avioneta, carreta, patinete |
| 12 | 2 | camión, caballo, taxi, tranvía, patinete, furgoneta | a pie, coche de caballos, taxi, monopatín, automóvil, autocar |
| 12 | 1 | globo (aerostático), burro, andar, a pie, monopatín, barca, (tren) AVE | globo (aerostático), furgoneta, ciclomotor, triciclo, lancha, (tren de) cercanías, gasoil, burro |
| 13 | 6 | sembrar, arar, regar, podar, segar, recolectar | arar, regar, podar, sembrar, jardinero, plantar |
| 13 | 5 | labrar, plantar, cortar césped | abonar |
| 13 | 4 |  | recolectar, siembra, segar, recoger |
| 13 | 3 | abonar, recoger aceitunas, cortar, cavar | cavar, recolección, agricultor, fumigar, talar |
| 13 | 2 | fumigar, talar, agricultor, trasplantar, injertar, recoger, cultivar | poda, riego, cortar, cosechar, injertar |
| 13 | 1 | hortelano, sembrar flores, arreglar flores, jardinería, jardinero, roturar, siega, vendimiar, trillar, echar abono | pastor, labrar, cultivar, trillar, limpiar, arado, botánica, huerta, huerto, recogida de algodón, segadora, cortar césped, vendimiar, regador |
| 14 | 6 | perro, gato, león | perro, gato, león, tigre, vaca, caballo, toro |
| 14 | 5 | caballo, tigre, elefante, conejo | elefante |
| 14 | 4 | pájaro, leopardo | oveja, cerdo, burro |
| 14 | 3 | toro, ratón, águila, canario, gallina, mono | serpiente, rinoceronte, pájaro, ratón, jirafa, oso |
| 14 | 2 | jirafa, oso, paloma, vaca, cerdo, rata, pato, cabra, ave | abeja, canario, cebra, búfalo, rata, avestruz, loro, cucaracha, buey, gallina, araña |
| 14 | 1 | ardilla, hormiga, zoo(lógico), pez, jabalí, tortuga, ballena, cebra, serpiente, delfín, pingüino, reptil, gorrión, burro, mamífero | cordero, cabra, conejo, paloma, pollo, leopardo, gorrión, dogo, felino, pato, lombriz, guepardo, ballena, pantera, ardilla, pitón, murciélago, mono |
| 15 | 6 | parchís, oca, cartas, cine, fútbol | parchís, fútbol, ajedrez, dominó, oca, póquer |
| 15 | 5 | ajedrez, baloncesto | baloncesto, cartas, pilla-pilla |
| 15 | 4 | comba, leer, pasear, tenis, teatro | videojuego, tenis, mus |
| 15 | 3 | lotería, escondite, dominó, tele(visión), juego de mesa, pilla-pilla, bailar | balonmano, damas, tele(visión), escondite, leer libros, cine |
| 15 | 2 | cinquillo, Monopoly, música, mus, pelota, muñeco, crucigrama | tute, leer, puzle, solitario, Monopoly, natación, volei(bol), pasear, ping-pong, Tetris |
| 15 | 1 | póquer, nadar, internet, solitario, balonmano, deporte, Play(-Station), guiso, baile, balón, baraja, Trivial, rueda | crucigrama, correr, golf, tres en raya, ordenador, coleccionar sellos, dados, hacer el tonto, hobby, naipe, parque acuático, PS2, cinquillo, deporte, senderismo, radio, internet |
| 16 | 6 | médico/a, maestro/a, profesor/ora, albañil, carpintero/a, abogado/a, fontanero | médico/a, albañil, carpintero/a, profesor/ora, abogado/a, fontanero, enfermero/a |
| 16 | 5 | electricista | maestro/a, ingeniero, electricista |
| 16 | 4 | jardinero, dependiente/a, enfermero/a, peluquero/a | mecánico, camarero/a, bombero, jardinero |
| 16 | 3 | juez/eza, ingeniero, camarero/a | arquitecto/a, conductor/ora, policía, pintor, zapatero, cocinero/a, chapista, escayolista, panadero |
| 16 | 2 | farmacéutico/a, informático, policía, carnicero, pintor, bombero, mecánico, escritor, modisto[a] | celador, secretario/a, administrador, policía local, vendedor/ora, ebanista, solador, carnicero |
| 16 | 1 | economista, psicólogo, cocinero/a, notario/a, óptico/a, pescadero, panadero, comercial, doctor, ingeniero de minas, maestro de música, representante, ama de casa, fiscal, administrativo, arquitecto/a, obrero, limpiador/ora | tapicero, ama de casa, dependiente/a, administrativo, obrero, cetrero, frigorista, ingeniero de caminos, peón, limpiador/ora, agricultor, aparejador, comercial, economista |
| 17 | 6 | euro, banco, bolsa, peseta, moneda, billete | euro, moneda, bolsa, billete, banco, peseta, dólar |
| 17 | 5 |  |  |
| 17 | 4 | hipoteca, dólar | acción, libra |
| 17 | 3 | cheque, economista, libra, comprar, céntimo | hipoteca, cuenta corriente, ahorro, cambio |
| 17 | 2 | acción, franco, marco, inversión, cuenta corriente, tarjeta, cajero, caja de ahorros, cartilla, interés, préstamo | céntimo, IVA, trabajo, íbex 35, banquero, letra (de cambio), cartera, comprar, dinero, inversión, préstamo, trueque |
| 17 | 1 | gastar, lira (italiana), monedero, contable, importante, llegar a fin de mes, pobreza, ahorro, sueldo, oferta | franco, economista, necesario, piso, warrant, lira (italiana), IPC, factura, caja de ahorros, interés, hacienda, cheque, tarjeta de crédito, divisa, coche, economía, comercio |
| 18 | 6 | ratón, pantalla, chatear, teclado | teclado, ratón, pantalla, ordenador |
| 18 | 5 | torre, Messenger, chat, ordenador | monitor, página web |
| 18 | 4 | correo (electrónico) | programa, disco duro, PC |
| 18 | 3 | impresora, Windows, página web, información, disco duro | Google, Messenger, navegar, impresora, web, placa base, bit, torre |
| 18 | 2 | Chat, Google, red, Web, Bill Gates, E-mail, altavoz, juego | correo (electrónico), software, red, computadora, chatear, virus, internet, Web, hardware, escáner, ADSL, altavoz |
| 18 | 1 | Software, amigo, programa, software, webcam, Microsoft, conexión, nueva era, punto com, router, mensaje, ADSL, CD-ROM, micro(fono) | mensaje, CD-ROM, Pentium, micro(fono), informática, automatismo, buscar trabajo, desesperación, Hispachat, interesante, Java, Chat, información, procesador, CPU, Internet, invento, grabadora, ventilador, Basic, ozu |
| 19 | 6 | rojo, azul, amarillo, verde, blanco, rosa, negro, naranja, celeste | azul, rojo, blanco, amarillo, verde, negro, violeta, rosa, marrón |
| 19 | 5 | marrón, violeta | gris, naranja, celeste, morado |
| 19 | 4 | morado | añil |
| 19 | 3 | gris, arco iris, turquesa | arco iris, fucsia |
| 19 | 2 | beis, azul marino, añil, luz, fucsia, cálido | beis, magenta |
| 19 | 1 | primario, burdeos, malva, alegría, lila, pastel, azul celeste, cromatismo, violeta claro, magenta | primario, cálido, tono, azul marino, turquesa |
| 20 | 6 | ola, agua, barco, pez, arena, sal | ola, barco, agua, pez, arena, playa, sal |
| 20 | 5 | playa | alga, roca |
| 20 | 4 | mar, alga | mar, tiburón |
| 20 | 3 | pesca, azul, paz, espuma, oleaje, marea | azul, orilla, océano, horizonte |
| 20 | 2 | orilla, sol, pescador, roca, calamar, boquerón, barco de vela, ballena, mejillón | color, velero, maremoto, marea, navegar, arrecife, coral, barca, vida, ballena, nadar, viento, espuma |
| 20 | 1 | concha, tiburón, nadar, caracola, rompeolas, brisa, delfín, pescado, inmenso, mar tranquilo, veranear, yodo, barca, océano, acantilado, medusa, relajación, sardina, tranquilidad, gaviota, marejada | pescador, sombrilla, tranquilidad, bonito, Resolution, pesca, red, lancha (motora), oleaje, puerto, bucear, delfín, almeja, boquerón, moto de agua, pescar, verano, profundidad, resaca, pulpo |

Mediante un diagrama de puntos, en el que se representan en cada eje las disponibilidades para hombres y mujeres, pero sólo de aquellos términos que no se han clasificado en el mismo nivel, y etiquetando por diferencia de nivel, se obtendría:

levelsSexo <-  
 inner\_join(levelsHombres %>%  
 select(centers,level,words,availability) %>%  
 rename(avHombre=availability,levelHombre=level),  
 levelsMujeres %>%   
 select(centers,level,words,availability) %>%  
 rename(avMujer=availability,levelMujer=level),   
 by=c("centers","words"))   
levelsSexo %>%  
 filter(levelHombre != levelMujer) %>%  
 mutate(diffLevel=factor(abs(levelHombre-levelMujer))) %>%  
 ggplot(aes(x=avHombre,y=avMujer,color=diffLevel)) + geom\_point()



Se puede observar que en la mayor parte de los casos en los que se clasifican en niveles distintos, lo hacen con poca diferencia.

diffs <- tapply(levelsSexo$levelHombre == levelsSexo$levelMujer,   
 levelsSexo$centers,   
 mean) \* 100   
diffs

## 01 02 03 04 05 06 07 08   
## 71.66667 57.31707 60.00000 66.27907 66.20690 61.64384 71.42857 64.83516   
## 09 10 11 12 13 14 15 16   
## 66.66667 65.57377 71.42857 74.68354 65.15152 69.11765 64.95726 73.68421   
## 17 18 19 20   
## 71.30435 66.00000 73.77049 67.17557

summary(diffs)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 57.32 65.10 66.47 67.44 71.43 74.68

De hecho, se puede observar que el porcentaje de términos clasificados en el mismo nivel en cada centro de interés va entre el 57 y el 74 por ciento.

1. <https://www.r-project.org/> [↑](#footnote-ref-21)
2. <https://rstudio.com/> [↑](#footnote-ref-23)
3. <https://www.tidyverse.org/> [↑](#footnote-ref-25)