# Caveats generales/técnicos sobre las gráficas

Para analizar los resultados, primero tengo que aclarar un par de cosas.

## A. los ejes Y reales [REESCRIBIR ESTO PORQUE ESTÁ HORRIBLE]

Primero quiero escribir [REVISAR] un poco sobre los ejes y [NO ESTOY SEGURO DE CÓMO ESCRIBIR EJE Y]. La gráfica de estilo tiene un rango que va de 2.745657 a 3.300022, y la de tema uno que va de 0.4275955 a 0.5543502. Estos números pueden parecer muy pequeños, pero antes de hacer ese juicio hay que entender qué significan. Como ya describí en el capítulo anterior, lo que representan los ejes Y en ambas gráficas es el resultado de un cálculo de distancias en un espacio multidimensional. Veamos la gráfica de estilo. En el caso del año 2020, por ejemplo, las distancias promedio entre los trabajos de titulación resultaron ser de 2.75. ¿2.75... qué? Lo que estamos viendo cuando vemos ese 2.75 es una condensación de datos. Recordemos que lo que medimos en el capítulo pasado fueron las frecuencias relativas de las 150 palabras más comunes. Es tentador pensar que ese 2.75 es 2.75%, puesto que las frecuencias relativas se miden en porcentajes. Sin embargo, estaríamos cometiendo un error. Ese 2.75 es resultado de, dicho burdamente, sacar la hipotenusa de un tríangulo. El número 2.75 no representa nada que podamos ver en el mundo real: es más bien el resultado de varios procesos de simplificación de datos. Es por esto que no es acertado pensar que 2.75 es mucho ni que es muy poco. La información valiosa que sale de esta gráfica viene de comparar los datos a lo largo de los años y de observar la tendencia. Entonces, en 2020 hubo 2.75 ¿qué?. Nada, no hubo 2.75 nada. Esto mismo aplica para la gráfica de temas. No es que en el 2011 haya habido 0.55 temas, ni que hayan sido 55%[[1]](#footnote-1) diferentes entre sí los temas de los trabajos de ese año. No es posible interpretar los rangos de los ejes y como lo haríamos en una gráfica más común: lo que nos da información son la variación a lo largo del tiempo y las tendencias.

[NECESITARÍAMOS UN GRUPO CONTROL PARA PODER DIMENSIONAR LAS ESCALAS DE LAS GRÁFICAS]

## B. Las fechas de publicación no indican las fechas de escritura: las gráficas apuntan al profesorado más que al estudiantado.

Otra cosa que hay que reconocer es que las fechas de publicación no son un indicador de cuándo fue escrito el trabajo. Bien pudo haber sido escrito en el año en el que fue publicado, o bien pudo haber sido escrito quince años antes. Lo que es claro es que fue aceptado en el año de publicación o el año anterior (pues la Forma de Examen Profesional 1 tiene vigencia de un año). Esto quiere decir que cuando ordenamos los datos por fecha de publicación, las tendencias que resultan apuntan al profesorado (que decide cuáles trabajos se publican), más que al alumnado. Claro está que no se puede ignorar al alumnado, pero me parece mucho más acertado interpretar estas tendencias como reflejos de quienes abren las puertas que de quienes las cruzan.

## C. Tres etapas

Me parece que en ambas gráficas hay tres etapas homólogas, que son 2006-2011, 2011-2017 y 2017-2023. En la gráfica de estilo se ve cómo en la primera etapa el estilo se diversifica un poco, luego cómo en la segunda se homogeneiza rápidamente y en la tercera cómo continúa con la homogeneización pero de manera más lenta. En la de tema se ven los mismos primeros dos movimientos (aunque el primero es un poco más precipitado) y en el tercero vuelve a aumentar la diferencia. Es evidente es que algo pasó entre el 2011 y el 2017 que hizo que tanto el estilo como los temas se homogeneizaran. [NO TENGO IDEA DE QUÉ ES]

# Estilo: Hipótesis respecto a por qué se homogeneiza

## A. Caveat de que no tengo ninguna hipótesis comprobada y que esto es más un trabajo de interpretación subjetiva.

Antes de presentar ḿis hipótesis respecto a por qué se homogeneiza el estilo, tengo que dejar claro que no tengo datos que apoyen a ninguna. Para comprobarlas hacen falta investigaciones futuras.

## B. Autocorrector y otras herramientas que corrigen el estilo.

Mi primera hipótesis (y la que intuyo que es la menos acertada y menos interesante) es que el aumento de herramientas digitales que corrigen errores de ortografía (y a veces de gramática) es un factor en la homogeneización del estilo de los trabajos de letras inglesas. Estoy pensando específicamente en los correctores ortográficos de MS Word y Google Docs, así como en los gramáticos de Google Docs y de servicios como Grammarly o QuillBot.[[2]](#footnote-2) Digo que no creo que esta hipótesis sea tan acertada porque, número uno, no me parece que estos programas sean tan intrusivos como para generar un cambio así de notorio. Es decir, ni el corrector de ortografía de Word ni el de Docs incide en el uso de palabras funcionales. Hablando de correctores gramáticos, sí sugieren a veces la inserción de ciertas palabras funcionales pero la verdad es que son todavía muy poco confiables (Docs corrigió "Hablando correctores gramáticos" a "Hablando correctores cromáticos", cuando yo quería que insertara "de" entre "Hablando" y "correctores"), por lo que no creo que sean herramientas que lxs estudiantes hayan usado, en especial entre el 2011 y el 2017.[[3]](#footnote-3) Si hoy no son confiables, no me imagino que en el 2011 hayan sido de gran ayuda a la hora de escribir un trabajo de titulación de letras inglesas.

## C. Aumento en el rigor del sínodo

Mi segunda hipótesis, en la cual confío un poco más, es que a lo largo del tiempo, lxs profesorxs que han sido parte de los sínodos han ido aumentando el rigor de la escritura de los trabajos de titulación. Más que rigor respecto a los métodos de investigación (que estoy seguro que ha aumentado) me refiero a rigor en la forma de exponer argumentos.

## D. Deseo de que se titule más gente: escritura formularia

cor\_pear: -0.1033474

cor\_pear2: -0.3838292 (quité el 2020)

cor\_spea: -0.179463

cor\_spea2: -0.3906406 (quité el 2020)

1. Recordemos que el software de modelado de temas que usé califica los temas del 0 al 1, que se puede interpretar como de 0% al 100%, por lo que 0.55 sería lo mismo 55%. [↑](#footnote-ref-1)
2. QuillBot fue lanzado en el 2017, por lo que no es el culpable de esta homogeneización, mientras que Grammarly fue lanzado en el 2009, haciéndolo un candidato más probable. [↑](#footnote-ref-2)
3. Será interesante ver cómo, con las innovaciones recientes en modelos extensos de lenguaje (*large language models*), evoluciona la homogeneización en el estilo. Predigo que en los próximos años sí va a haber una homogeneización precipitada. [↑](#footnote-ref-3)