En el presente trabajo se busca contrastar el lenguaje utilizado por YouTubers en video ensayos en diferentes temas, con el lenguaje académico encontrado en seminarios y clases en universidades de habla inglesa. Para esto se creará un corpus con los subtítulos de video ensayos de diferentes autores y el corpus de *The British Academic Spoken English (BASE)*, específicamente los apartados de *Arts and Humanities* y de *Social Science*. Para el primer caso se buscará subtítulos que han sido transcritos manualmente y no autogenerados por la inteligencia artificial de Google, para así reducir el margen de error de lo expresado por estas personas. En el segundo caso, se seleccionaron estas dos áreas en específico ya que los temas discutidos por los YouTubers tienden a recaer dentro de estas dos categorías, por lo que para hacer una comparación justa solo se utilizará esta sección del corpus.

Teniendo ambos corpus listos, se espera hacer una comparación del lenguaje utilizado. Se asume que el lenguaje utilizado por YouTubers, aunque haya sido creado para un público general no necesariamente especializado, se asemeje al encontrado en la universidad. En otras palabras, el lenguaje utilizado en los video ensayos será de un carácter académico. Éste se entenderá como un lenguaje más especializado, con uso de tecnicismos y palabras menos comunes en el lenguaje coloquial (Bunch & Martin, 2020); oraciones más complejas, con mayor nominalización y menor cantidad de frases verbales; se utilizan expresiones como “*appears to*”, “*has the possibility of*” o “*to some extent*” para así proteger las afirmaciones, dándoles un margen de error, entre otros factores.

La razón detrás de este trabajo es estudiar si el uso de lenguaje académico es preferido por YouTubers, cuya audiencia es un público general no necesariamente especializado en los temas abordados. Este tema es de interés por diferentes connotaciones el lenguaje académico conlleva. Éste tiene a ser relacionado con un cierto prestigio, sin embargo, también es utilizado como una manera de discriminación hacia otras variedades menos especializadas. Existe la idea de que individuos sin conocimiento del lenguaje académico no pueden hacer trabajos académicos propiamente tal (Bunch & Martin, 2020). Por esto, es interesante estudiar como este lenguaje sigue siendo utilizado cuando quienes pueden acceder al contenido pueden ser personas de diferentes orígenes y experiencias, no necesariamente relacionados a la academia, pero aun así interesados en los temas de los cuales se habla en estos videos, o como los creadores, pudiendo optar por un lenguaje más coloquial y accesible, siguen prefiriendo la variedad con mayor prestigio.

Python será utilizado para preparar los textos para crear los corpus. También servirá para medir la distancia entre ambos tipos de habla y comparar la frecuencia de palabras utilizadas.

Corpus BASE

<http://www.reading.ac.uk/AcaDepts/ll/base_corpus/ah/index.htm>

<http://www.reading.ac.uk/AcaDepts/ll/base_corpus/ss/index.htm>

**Referencia**

Bunch, G. C., & Martin, D. (2020). From “academic language” to the “language of ideas”: a disciplinary perspective on using language in K-12 settings. *Language and Education*, *0*(0), 1–18. https://doi.org/10.1080/09500782.2020.1842443