

Diversos estudios han demostrado que la frecuencia léxica afecta el procesamiento lingüístico¹, y cuando no se controla, puede introducir confusiones en los resultados de un experimento psicolingüístico. La calidad de las medidas de frecuencia pueden ser evaluadas usando datos de procesamiento de palabras⁷. Dada su relevancia para la investigación psicolingüística, este estudio tiene como objetivo explorar los datos de frecuencia subjetiva disponibles para el español rioplatense. Tras recolectar calificaciones subjetivas de frecuencia para 360 palabras y obtener tiempos de respuesta en una tarea de decisión léxica, concluimos que las frecuencias obtenidas a partir de un estudio de asociación libre (SWOWrp²) son las mejores predictoras de los tiempos de respuesta. Asimismo, la herramienta desarrollada para obtener puntuaciones subjetivas de frecuencia demostró ser confiable y pueden ser utilizadas para estimar la frecuencia léxica en ausencia de corpus disponibles.

Participantes

→ Un total de 146 adultos participaron en el estudio de frecuencia subjetiva. Se excluyó a 2 voluntarios porque no utilizaron correctamente la escala de calificación de frecuencia. Se analizaron las puntuaciones proporcionadas por los 144 participantes restantes (114 mujeres; edad promedio = 40 años, DE = 13 años; rango de edad de 22 a 91 años). Cada palabra fue puntuada en promedio por 30 participantes.

→ Un total de 249 adultos participaron en el estudio de frecuencia léxica. Se excluyó a 8 voluntarios porque no tuvieron una suficiente tasa de aciertos. Se analizaron las evaluaciones realizadas por los 241 participantes restantes (183 mujeres; edad promedio = 34 años, DE = 15 años; rango de edad de 18 a 78 años). Cada palabra fue evaluada en promedio por 52 participantes.

Materiales

→ Se seleccionaron 360 palabras del español rioplatense de manera que también se conociera la frecuencia de su equivalente en la Lengua de Señas Uruguaya (LSU) de 277 de ellas. Algunas palabras también fueron seleccionadas porque se disponía de frecuencia a partir de un estudio de asociación libre².

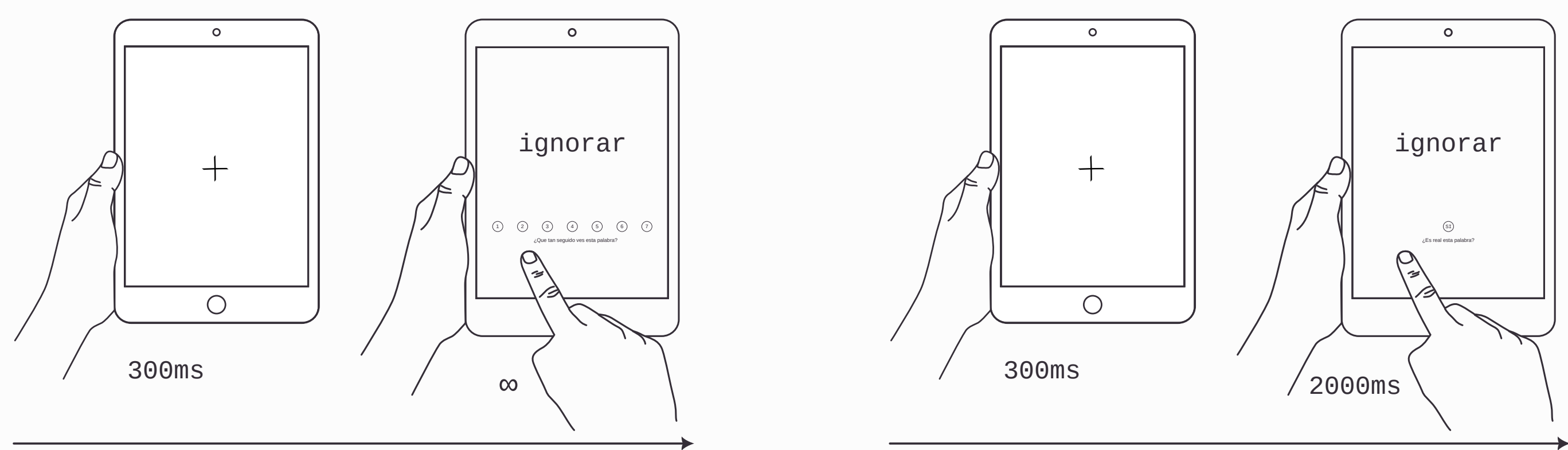
→ Para el estudio de decisión léxica se generaron, además, 360 pseudopalabras usando UniPseudo³.

Procedimiento

Los datos se recopilaron en línea usando dos sitios web accesibles desde un dispositivo móvil o una computadora. El código fuente de los sitios es de acceso público y puede ser utilizado por otros investigadores.

En el estudio de frecuencia subjetiva se informó a los participantes que verían una serie de palabras del español rioplatense y se les instruyó que tocaran (si usaban una pantalla táctil) o hicieran clic en un botón de una escala Likert de 7 puntos según la frecuencia con la que encontraban cada palabra: 1 significaba "nunca vi esta palabra" y 7 "veo esta palabra casi todos los días". Se les asignaron aleatoriamente 90 palabras estímulo. La duración aproximada de la tarea fue de 10 minutos.

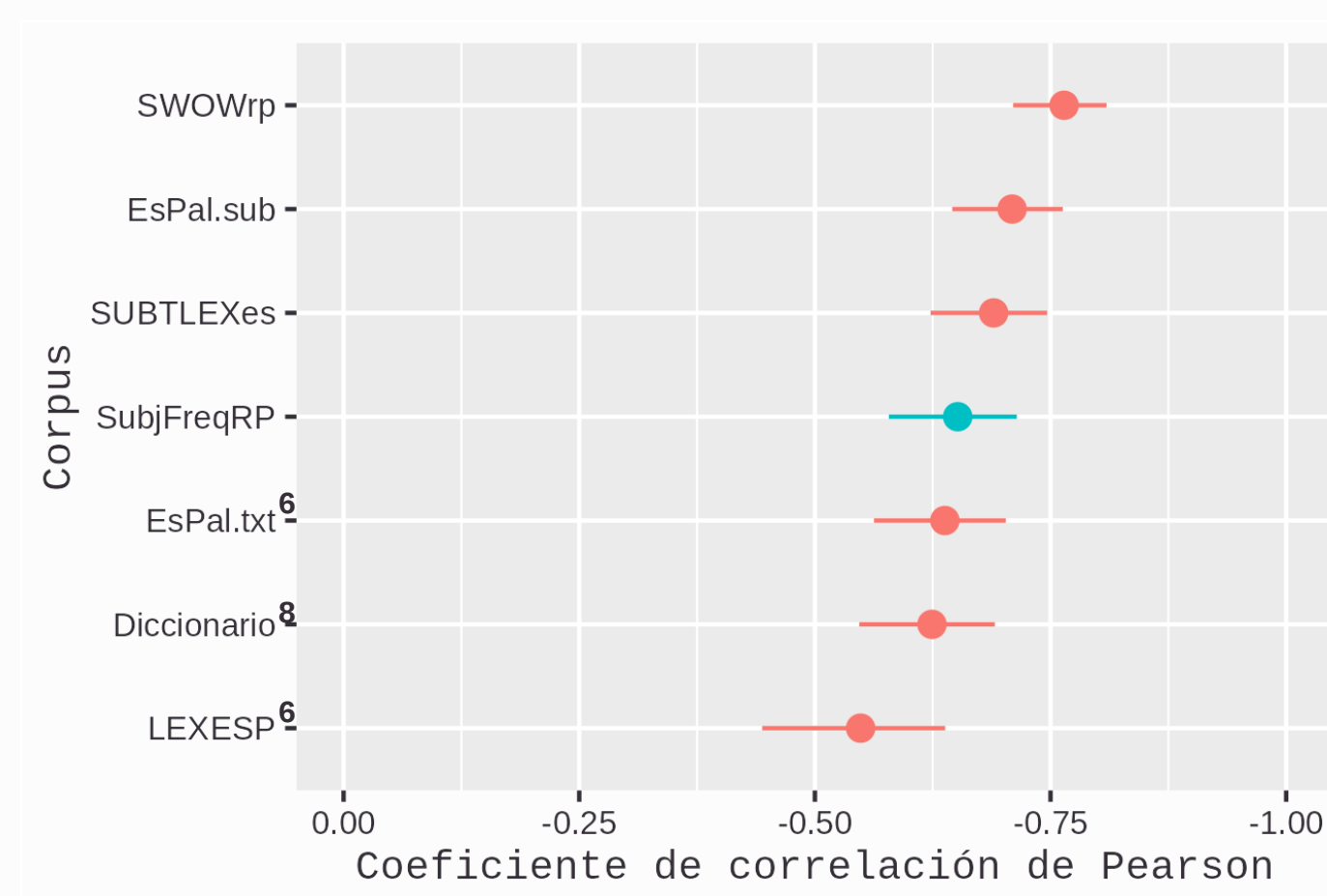
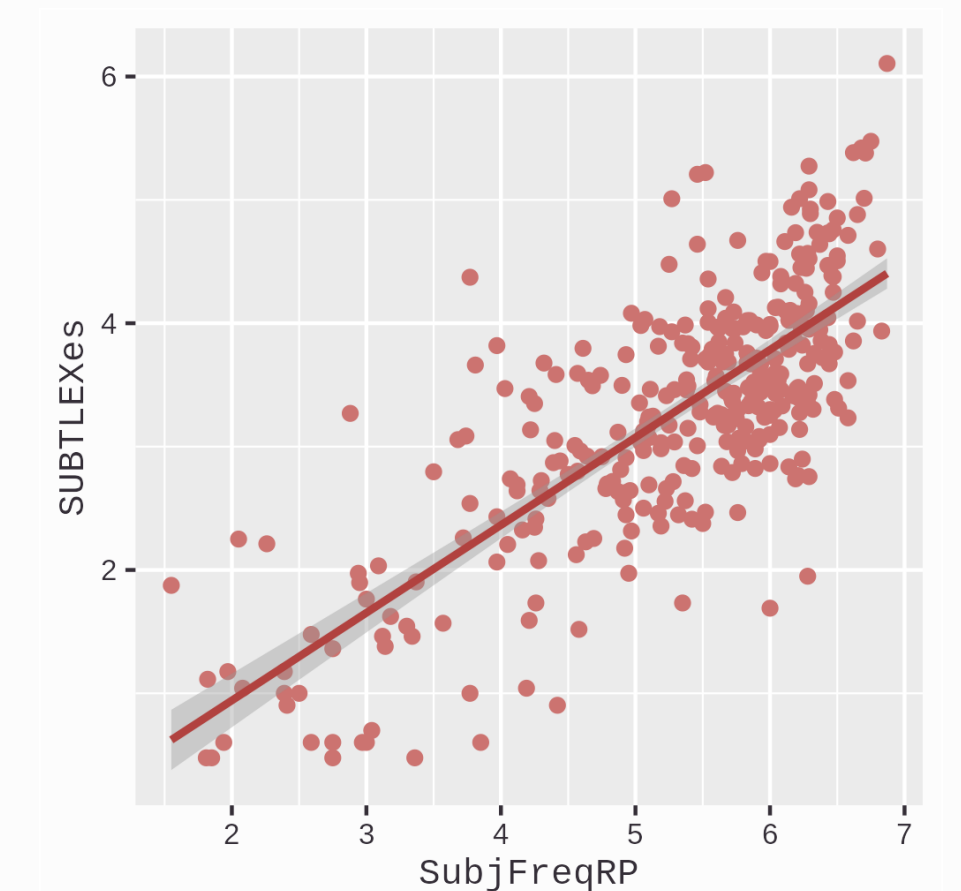
En el estudio de decisión léxica se informó a los participantes que verían una serie de palabras reales o inventadas del español rioplatense y se les instruyó que tocaran (si usaban una pantalla táctil) o hicieran clic en un botón según si la palabra presentada era real o inventada. Se les asignaron aleatoriamente 180 palabras y pseudopalabras estímulo. La duración aproximada de la tarea fue de 10 minutos



Resultados

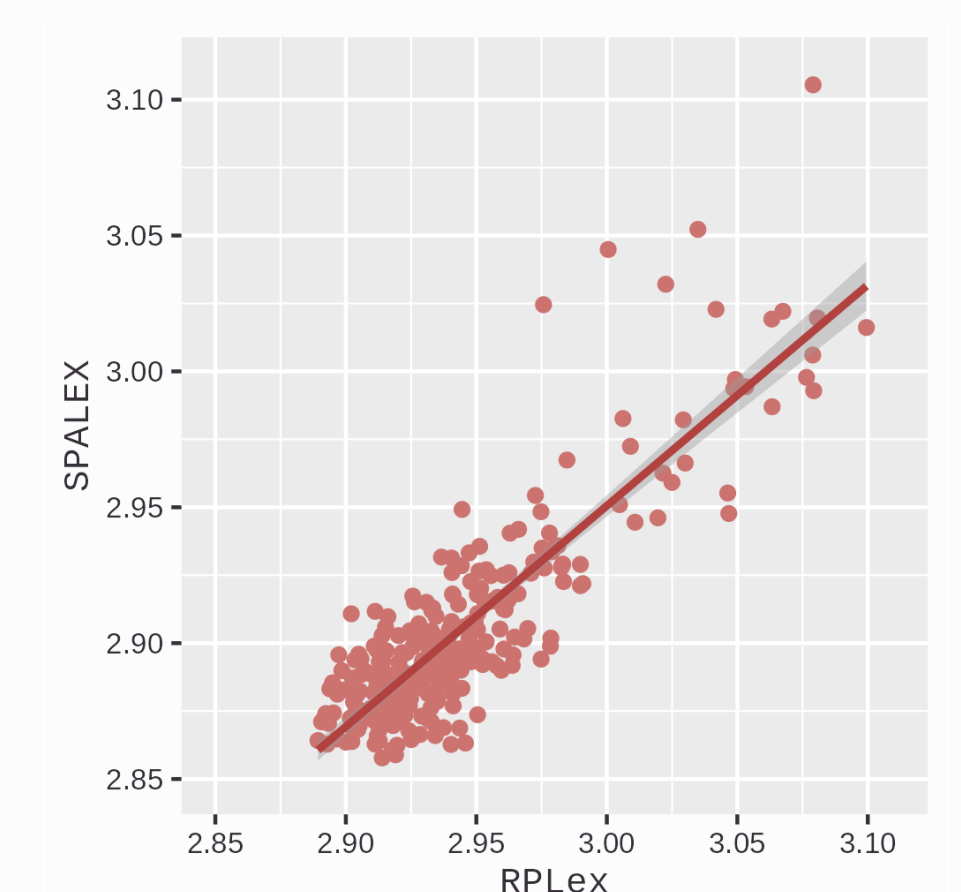
A Correlación de frecuencias subjetivas con SUBTLEXes

Las puntuaciones de frecuencia subjetiva obtenidas muestran una correlación fuerte con SUBTLEXes⁴ ($r = .77$; $p < .001$).



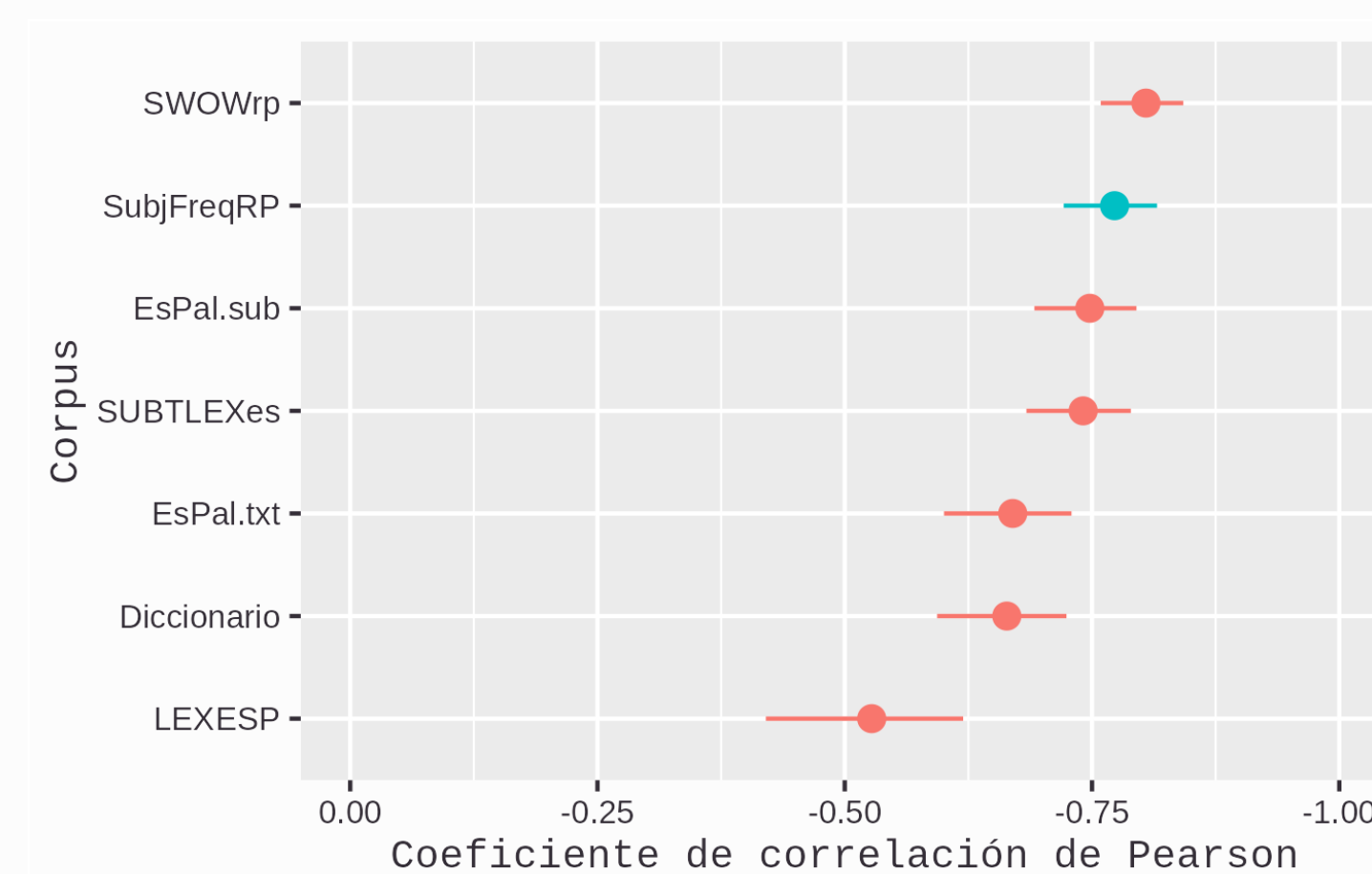
B Correlación de frecuencias con SPALEX

Las frecuencias subjetivas relevadas muestran una correlación fuerte con el corpus de SPALEX⁵ ($r = -0.65$; $p < .001$), pero se ubican debajo de SWOWrp ($r = -0.76$; $p < .001$), EsPal.sub⁶ ($r = -0.71$; $p < .001$) y SUBTLEXes ($r = -0.69$; $p < .001$).



C Correlación de tiempos de respuesta con SPALEX

Los tiempos de respuesta obtenidos muestran una correlación fuerte con SPALEX ($r = -.77$; $p < .001$).



D Correlación de frecuencias con los tiempos de respuesta obtenidos

Las frecuencias subjetivas relevadas muestran una correlación fuerte con los tiempos de respuesta obtenidos en la tarea de decisión léxica ($r = -0.77$; $p < .001$), sólo superadas por SWOWrp ($r = -0.81$; $p < .001$).

SWOWrp	59, 67%
SWOWrp + SUBTLEXes	69, 36%
SWOWrp + SUBTLEXes + SubjFreqRP	71, 89%
SWOWrp + SUBTLEXes + SubjFreqRP + Longitud	74, 83%
SWOWrp + SUBTLEXes + Longitud	71, 41%

E Análisis de regresión múltiple

Cuando se toman en cuenta la frecuencia calculada a partir de una tarea de asociación libre (SWOWrp), la frecuencia calculada a partir de un corpus (SUBTLEXes) y la longitud de las palabras, las frecuencias subjetivas obtenidas son capaces de explicar un 3,42% adicional de la variabilidad de los tiempos de respuesta en una tarea de decisión léxica. Todos juntos son capaces de explicar un 74,82%.

Referencias

- Preston, K. A. (1935). The Speed of Word Perception and Its Relation to Reading Ability. *The Journal of General Psychology*, 13(1), 199–203. doi:10.1080/00221309.1935.9917878
- Cabana, Á., Zugarramurdi, C., Valle-Lisboa, J. C., & De Deyne, S. (2023). The "Small World of Words" free association norms for Rioplatense Spanish. *Behavior Research Methods*. doi:10.3758/s13428-023-02070-z
- New, B., Bourgin, J., Barra, J., & Pallier, C. (2024). UniPseudo: A universal pseudoword generator. *Quarterly Journal of Experimental Psychology*, 77(2), 278–286. doi:10.1177/17470218231164373
- Cuetos, F., Glez-Nosti, M., Barbón, A., & Brysbaert, M. (2012). SUBTLEX-ESP: Spanish word frequencies based on film subtitles. *Psicológica*, 33(2), 133-143.

- Aguasvivas, J. A., Carreiras, M., Brysbaert, M., Mandera, P., Keuleers, E., & Duñabeitia, J. A. (2018). SPALEX: A Spanish lexical decision database from a massive online data collection. *Frontiers in psychology*, 9, 2156.
- Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., & Carreiras, M. (2013). EsPal: One-stop shopping for Spanish word properties. *Behavior research methods*, 45, 1246-1258.
- Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect. *Experimental psychology*.
- Pérez, M. Á., Alameda, J. R., & Cuetos, F. (2003). Frecuencia, longitud y vecindad ortográfica de las palabras de 3 a 16 letras del Diccionario de la Lengua Española (RAE, 1992). *REMA Revista electrónica de metodología aplicada*, 8(2), 1-10.