

**El Lenguaje de la Depresión:**  
**Categorías Lingüísticas y Temas Usados en Foros de Discusión en el Internet**  
**en Inglés y en Español**

**Nairán Ramírez-Esparza<sup>1</sup>, Cindy K. Chung<sup>2</sup>, Gisela Sierra-Otero<sup>3</sup> y James W.  
Pennebaker<sup>2</sup>**

## Resumen

El estudio de comunicación on-line puede dar a conocer la expresión de la depresión a través de las culturas, sugiriendo una posible información y/o ayuda necesaria y tópicos que necesitan ser discutidos en psicoterapia. Esta investigación explora esta idea mediante dos estudios. En el Estudio 1 se observaron cuáles son los indicadores lingüísticos de depresión extrayendo textos de personas deprimidas (N=160) y personas no deprimidas (N=160) de foros de discusión en inglés y español. Se utilizó un programa de computadora (LIWC2001) para comparar categorías lingüísticas a través de los grupos (i.e., personas deprimidas vs. personas no deprimidas) en cada una de las lenguas. En el Estudio 2 se analizaron los temas que las personas usan cuando hablan de su depresión en foros de discusión en inglés y en español. Se colectaron textos de personas deprimidas de foros en español (N=400) y foros en inglés (N=400). Se utilizó un método automático inductivo y se comparó cómo las personas hablan de su depresión a través de las lenguas. Los resultados del Estudio 1 mostraron que en ambas lenguas los marcadores lingüísticos asociados a la depresión, fueron utilizados más en textos de personas deprimidas que en los textos de las personas no deprimidas. Los resultados del Estudio 2 mostraron que las personas deprimidas que escribieron en español sobre su depresión utilizaron temas relacionados a procesos sociales, que las personas que escribieron en inglés, los cuales a su vez escribieron más acerca de temas relacionados con procesos médicos. Se abordan y discuten las implicaciones que tienen los resultados para necesidades de información y apoyo en psicoterapia vs. apoyo on-line.

**Palabras clave:** Depresión, Internet, transcultural, lenguaje

El propósito de este estudio fue explorar los temas que surgen naturalmente en Internet, a través de Foros de Discusión sobre depresión, y comparar las similitudes y diferencias entre los usuarios de lengua inglesa y de lengua española. Se tiene la conciencia de que a estos foros pueden acceder personas que presenten estados fisiológicos propios de la patología depresiva, y también acceden a ellos usuarios con depresión leve o simplemente en estados prolongados de tristeza. Esto no puede discriminarse fácilmente.

---

Si bien, algunos autores plantean la duda respecto a que, la diferencia entre depresión y tristeza está dada sólo por niveles de intensidad (Wolpert, 1999), así como también Klemann y Good (1985), en su libro *Culture and Depression* hacen surgir la pregunta de la relación entre depresión y tristeza; Barr-Zisowitz, en el *Handbook of Emotions* (2004), confirma que hay una controversia acerca de si la depresión es un desorden biológico y universal, y qué tanto es un estado culturalmente determinado o un conjunto de conductas; aunque la mayoría de los autores coinciden en que no son lo mismo, pero sí son primas cercanas.

Sin embargo, en estudios recientes se encontró que es posible detectar los principales síntomas que caracterizan la depresión a través del uso natural del lenguaje, especialmente a través del programa para computadora LIWC2001 (Pennebaker, Francis, & Booth, 2001) y la versión traducida al español de Ramírez-Esparza, Pennebaker, García, & Suriá (2007). Aun cuando no se identifique el grado de depresión o de tristeza, es un hecho que las emociones y problemas que tienen que afrontar estas personas son algo muy parecido, así como la evidencia que se ha encontrado (para un resumen ver, Pennebaker, Mehl, & Niederhoffer, 2003; Ramírez-Esparza & Pennebaker, 2006), de que las personas con algún tipo de conflicto, cuando lo hablan o escriben tienden a mejorar en forma significativa en comparación con aquellos que la mantienen para su intimidad. Es también un hecho irrefutable que a cualquier nivel del problema, se necesita ayuda, y que identificando la problemática a través del LIWC2001 podría proporcionárseles.

El detectar los síntomas de algún grado de depresión en los textos recuperados puede ser en ocasiones una tarea un tanto obvia. Consideren el ensayo que una persona ha escrito en un foro de discusión de depresión en el Internet:

Tengo 22 años y siempre creí que era normal sentirme tan cansada, triste, agotada y harta de todo, imaginaba formas de matarme, cayendo de un coche andando, durmiéndome con algún collar o algo largo para que me ahorcara, tomando varias pastillas, desafortunadamente me

imaginaba esto desde que tenía 3 años. Mi mamá sólo me decía que estaba amargada, que la familia de mi padre me influenciaba y me comportaba como una niña consentida, que era de un carácter fuerte, por un tiempo pensé que así era y traté de cambiar pero me di cuenta que conforme pasaba el tiempo la gente se me acercaba más pero yo me sentía cada vez más triste, decepcionada. Por un tiempo pensé que era homosexual, tenía la esperanza de serlo porque así significaría que yo no estaba triste sin motivo que había algo de verdad detrás de mi depresión, pero me di cuenta que no lo era, no tenía ningún problema, si era heterosexual y creo que eso me deprimió más, no había motivo para sentirme así. Hasta hace poco las ideas suicidas cambiaron, ya no solo lo pensaba en varias ocasiones trataba de provocarlo, por accidente cortándome, poniendo un poco de agua en las escaleras para caerme, no comía durante varios días, no quería que lo notaran mis padres ni nadie por eso lo hacía así, para que no sospecharan y pensarán que era solo descuidada. Cuando me animé a hacer algo fue en un accidente de coche, vi que el coche que me iba a pegar se acercaba y enés de acelerar o de no avanzar para que no me pegara decidí adelantarme y esperar, desear, que me pegara, destruyó el coche pero solo salí con un pequeño dolor de cuello. No pude morir, esperaba que se dieran cuenta que yo lo había deseado pero nunca se dieron cuenta, no se han dado cuenta que lo hice. No quisiera aceptarlo pero me gusta ser así tengo miedo que se enteren mi familia, amigos, porque sé que tratarían de ayudarme pero no quiero, pero quiero que se enteren, estoy perdida. Vivo en una familia de padres separados, pero eso nunca me afectó al contrario me hizo ver que estaba más madura y concentrada de lo que pensaba, nunca hemos tenido problemas, sin embargo siento que no tengo motivo, estoy mal, enferma, ya quiero irme.

Es difícil imaginar que alguien escriba sobre sus sentimientos más profundos en foros de discusión donde el mundo entero lo puede ver. Sin embargo, el uso de foros de discusión en el Internet ha crecido en popularidad porque ofrece apoyo y ayuda de una manera anónima y económica y además pueden ser accesados por personas que viven en áreas geográficamente remotas (Houston, Cooper, & Ford, 2002; Stjernsward & Ostman, 2006). En los foros de discusión las personas típicamente intercambian sus experiencias en línea con respecto a un tema o lo utilizan para buscar ayuda, especialmente sobre enfermedades estigmatizadas como la depresión (Baker, Wagner, Singer & Bundorf, 2003; Berger, Wagner, & Baker, 2005; Fox & Fallows, 2003). De

acuerdo a lo que la persona menciona en su ensayo, se puede inferir que está deprimida, por ejemplo menciona estar cansada, triste, agotada, tener ideación suicida, decepción, se centra en sí misma, se ve a sí misma y al mundo en forma negativa; todos estos síntomas son mencionados por la *American Psychiatric Association* (APA, 1994) como propios de un trastorno depresivo mayor, esta asociación también menciona otros síntomas como son: melancolía, agitación o retardo psicomotor, pérdida de interés, insomnio, hipersomnia, cambios en el apetito, dificultad para concentrarse e indecisión.

Otras investigaciones han encontrado que las personas que visitan estos foros están deprimidas. Por ejemplo, en un estudio se les invitó a participantes de foros de discusión en depresión a completar cuestionarios y se encontró que muchos participantes del foro tienen altos niveles de síntomas depresivos, están en terapia o han recibido terapia y/o toman antidepresivos y además reportan que platicar en línea les ayuda a aliviar sus síntomas. Más aún, en este estudio se encontró que los síntomas depresivos disminuían con el tiempo, especialmente en aquellos que usaban los foros frecuentemente (Houston et al., 2002).

El hecho de que en los últimos años hayan surgido foros de discusión en línea da la oportunidad de observar y analizar de una manera discreta la forma en que las personas hablan acerca de situaciones específicas (Davison, Pennebaker, & Dickerson, 2000). Los avances recientes en programas de computadora que analizan textos, ha permitido a investigadores explorar el comportamiento natural que sucede en línea, mediante el análisis del lenguaje que usan las personas en foros. Por ejemplo Owen, Yarbrough, Vaga, y Tucker (2003) utilizaron el Buscador Lingüístico y Contador de Palabras (LIWC, por sus siglas en inglés, Pennebaker et al., 2001) para analizar la expresión emotiva y cognitiva en grupos de ayuda de cáncer en el Internet. En otro estudio Lieberman y Goldstein (2006) utilizaron el LIWC para analizar la expresión emotiva en mujeres con cáncer de mama y encontraron que las mujeres que utilizaron

más palabras relacionadas a la ansiedad experimentaron más depresión. Por otra parte Lyons, Mehl, y Pennebaker (2006) encontraron que en foros de discusión sobre la anorexia, que los anoréxicos, que reportan haberse recuperado de la anorexia utilizan más palabras emotivas positivas, menos palabras de ansiedad y menos palabras relacionadas con enfoques en uno mismo. Estas investigaciones demuestran que usar un enfoque de conteo de palabras de manera automática ha sido una manera válida de caracterizar el lenguaje de grupos de discusión en línea de una manera eficiente, teórica y de relevancia para la psicología.

### *Foros de Discusión*

Los foros pueden ser públicos o privados (donde una clave privada es requerida), y son frecuentemente supervisados por un moderador que busca contenido ofensivo. Una serie de respuestas al mensaje resulta en lo que se podría llamar un hilo de comunicación. Los hilos de comunicación están accesibles para los lectores que están navegando la red. Esta característica de archivar hace de los foros un gran lugar para que las personas interesadas puedan acceder información que ha sido discutida en el foro y determinar si es un foro en que les gustaría participar. Para investigadores, estas características de archivo de foros lo hace una fuente ecológicamente válida para coleccionar y examinar comportamientos naturales. El propósito de este estudio fue examinar los marcadores lingüísticos de depresión en línea (Estudio 1), y explorar los temas que surgen naturalmente en foros de discusión de depresión en el Internet y hacer una comparación a través de la lengua inglesa y española (Estudio 2).

En el estudio 1 comparamos marcadores lingüísticos en personas deprimidas y no deprimidas de foros de discusión en inglés y en español. Para lograr este objetivo utilizamos el LIWC (Pennebaker et al., 2001) y la versión traducida al español (Ramírez-Esparza et al., 2007). En el Estudio 2 analizamos los temas que las personas usan al hablar de su depresión. Para lograr este objetivo utilizamos un procedimiento que ha sido desarrollado recientemente y que analiza el lenguaje de una manera

inductiva. Este procedimiento es llamado el Método de Extracción de Significados (MES, Chung & Pennebaker, 2008) el cual usa herramientas de computación de análisis de texto para determinar cuáles son los temas más frecuentemente usados. El método inicialmente determina cuáles son las palabras más frecuentemente usadas a través de un corpus de textos y después se lleva a cabo un análisis factorial el cual agrupa las palabras más frecuentemente usadas en factores que muestran palabras que co-ocurren en el mismo texto. Los factores resultantes tienden a ser un grupo de palabras que son coherentes e intuitivamente comprensibles.

### **Estudio 1: Marcadores Lingüísticos de Depresión**

En estudios anteriores utilizando el LIWC se encontró que es posible caracterizar la depresión a través del uso natural del lenguaje. El LIWC (Pennebaker et al., 2001) es una herramienta de computación que cuenta y categoriza palabras. El LIWC contiene 72 dimensiones lingüísticas, las cuales fueron creadas a partir de conceptos psicológicos y con la ayuda de jueces. Estas dimensiones incluyen categorías de lenguaje estándar (e. g., artículos, preposiciones, pronombres, incluyendo primera persona en singular, primera persona en plural etc.), procesos psicológicos (e. g., categorías de emociones positivas y negativas, procesos cognitivos), palabras relacionadas a la relatividad (e. g., tiempo, tiempo del verbo, moción, espacio), y dimensiones tradicionales de contenido (e. g., sexo, muerte, casa, ocupación). Las categorías del LIWC están organizadas jerárquicamente, por ejemplo, todas las palabras relacionadas con la categoría 'enojo' también se organizan en las categorías de emociones negativas o palabras afectivas. El LIWC utiliza una estrategia de conteo de palabras donde busca sobre 2,300 palabras y raíces de palabras dentro de un texto dado. Cada palabra y/o raíces de palabra es parte de una o más categorías o subcategorías de palabras. Por ejemplo, la palabra 'llorar' caería dentro de las siguientes cuatro categorías: 'tristeza', 'emoción negativa', 'afecto en general', y 'tiempo

pasado'. Por lo tanto, si esta palabra es encontrada en un texto, cada una de las escalas de las mencionadas categorías se incrementará.

La versión final del diccionario en español (Ramírez-Esparza et al., 2007) cuenta con un total de 7,515 palabras y raíces de palabras, y con el mismo número de categorías y subcategorías que el LIWC en inglés. El diccionario en español tiene significativamente más palabras que la versión en inglés debido a que en español existen más conjugaciones verbales, se consideran las palabras en femenino y masculino, palabras acentuadas y no acentuadas.

### *Depresión y Marcadores Lingüísticos*

Utilizando el LIWC2001, Stirman y Pennebaker (2001) encontraron que poetas que habían cometido el suicidio tendían a utilizar más pronombres en primera persona (e.g., Yo, mi, mío) y menos pronombres en primera persona del plural (e.g., nosotros, nuestro) a través de sus carreras como escritores, en comparación con poetas que no habían cometido suicidio. Estos resultados apoyan el modelo de depresión llamado integración-social/desconexión-social de Durkheim (1951). El cual establece que una de las claves que llevan al suicidio es la percepción de que uno mismo no está integrado en la sociedad, o en otras palabras, que está desconectado de la vida social.

En otro estudio, Rude, Gortner, y Pennebaker (2004) utilizando el LIWC, encontraron que en los ensayos literarios libres sobre ingreso a la universidad (i.e., ensayos sobre emociones y pensamientos más profundos de estar en la universidad), escritos por estudiantes deprimidos y no deprimidos, los estudiantes que estaban deprimidos utilizaron el pronombre "yo" con mayor frecuencia que los no deprimidos y también tendieron a utilizar más palabras negativas y menos palabras positivas. Estos resultados apoyaron dos teorías de depresión: La primera menciona que las personas deprimidas tienden a pensar en una gran medida en ellos mismos. Por ejemplo, Pyszczynski y Greenberg (1987) han especulado que las personas deprimidas después de que han perdido su fuente de autovaloración, se ven sumergidos en tratar de ganar



lo que han perdido, por lo tanto tienden a enfocarse en sí mismos. También apoyan la teoría cognitiva de depresión de Beck (1967), la cual postula que las personas que están deprimidas tienden a verse a sí mismos y al mundo en términos negativos. Es importante hacer notar que clínicos también reportan que pacientes deprimidos utilizan más pronombres en primera persona (Bucci & Freedman 1981; Weintraub, 1981).

El lenguaje depresivo se ha estudiado ampliamente en texto escrito (Rude et al., 2004, Striman & Pennebaker, 2001); sin embargo, no se sabe mucho del lenguaje que se utiliza en línea. Más aún, no hay mucha investigación que explore el uso del lenguaje en otras lenguas, específicamente el español. Este estudio fue hecho para explorar los marcadores lingüísticos que las personas utilizan cuando escriben sobre su depresión en foros de discusión en cada una de las lenguas: inglés y español. Para lograr este objetivo, comparamos marcadores lingüísticos usados en foros de depresión en inglés y los comparamos con un grupo control. Seleccionamos como grupo control foros de discusión de cáncer de mama. Se utilizó este grupo como control porque en los foros de cáncer de mama, al igual que los foros de depresión, las mujeres dan su testimonio de vivir con un padecimiento, y expresan sus emociones, sus sentimientos, y platican sobre cómo su padecimiento ha afectado su vida en cuestiones sociales y psicológicas. De esta manera ambos grupos son comparables. El Segundo paso de este estudio fue analizar diferencias en marcadores lingüísticos entre personas deprimidas y un grupo control pero en foros de discusión en español.

### *Método*

Archivos de Texto. Un total de 160 textos se obtuvieron de los foros de discusión en español. De esos 80 textos fueron de mujeres de foros de depresión y 80 de foros de cáncer de mama. Por otro lado, un total de 160 textos se obtuvieron de foros de discusión en inglés. De esos 80 fueron de mujeres de foros de depresión y 80 fueron de mujeres de cáncer de mama (ver Ramírez et al., 2007 para más detalles de la muestra). Es importante hacer notar que un texto corresponde a una mujer. Es decir,

no se obtuvieron múltiples textos por cada participante. No se logró conocer otras características personales de los participantes (i.e., su edad y de dónde eran).

Procedimiento. Se seleccionaron muestras de textos de foros donde las personas escribieran un mínimo de 100 palabras. Por ejemplo, se acudía a un foro de discusión de depresión donde una persona se presentaba al grupo y hablaba sobre su depresión, se tomaba el texto, se corregían los errores de ortografía y se guardaba en formato de texto (e. g., depsujeto1.txt), y así sucesivamente. Lo mismo se hizo para los foros de discusión de cáncer de mama. Los análisis lingüísticos de los textos en ambos idiomas se analizaron con el LIWC correspondiente para cada idioma.

Se seleccionaron categorías relacionadas con las teorías de depresión antes mencionadas y que hubieran sido utilizadas en otras investigaciones (e. g., Rude et al., 2004; Stirman & Pennebaker, 2001). Estas categorías se refieren al uso de palabras en primera persona singular (e. g., yo, mío, mi), el uso de palabras con valencia positiva (e. g., contento, bonito, bueno, amor) y valencia negativa (e. g., odio, enemigo, nervioso, miedo), y el uso de palabras en primera persona del plural (e. g., nosotros, nuestro).

### *Resultados*

La Figura 1 muestra los resultados para los foros de depresión en inglés y la Figura 2 muestra los resultados de los foros de discusión en español. Los resultados mostraron que las mujeres de habla hispana y habla inglesa en foros de discusión de depresión utilizan más palabras en primera persona del singular y menos palabras de primera persona en plural que las mujeres de habla hispana y habla inglesa en foros de discusión de cáncer de mama. Estos resultados replican lo encontrado en otras investigaciones, donde poetas que se suicidaron utilizaban más pronombres en primera persona y menos en primera persona del plural (Stirman & Pennebaker, 2001), así como lo encontrado por Rude et al. (2004) donde estudiantes deprimidos tendieron a usar más palabras en primera persona que los estudiantes no deprimidos. Estos

hallazgos apoyan la teoría de Pyszczynski y Greenberg (1987) que asocia el centrarse en uno mismo con la depresión. Así como el modelo de suicidio de Durkheim (1951) de que las personas deprimidas tienden a aislarse y a desconectarse del mundo.

---Insertar Figuras 1 y 2 por aquí---

Así mismo, los resultados mostraron que las mujeres en foros de depresión utilizan más palabras con valencia negativa y menos palabras con valencia positiva que las mujeres en foros de cáncer de mama. Estos resultados están en línea con los hallazgos de otras investigaciones con respecto al uso de palabras negativas y positivas (e. g., Rude et al., 2004). Los resultados apoyan la teoría de Beck (1967) donde establece que las personas deprimidas tienden enfocarse en aspectos negativos de su vida.

### *Conclusión*

Los grupos de depresión en línea utilizaron un lenguaje similar a lo que se ha encontrado previamente en otros estudios. Interesantemente, las personas que hablan español utilizaron el lenguaje de manera similar que las personas que hablan inglés. Una de las limitaciones de este estudio es que no fue posible comparar diferencias a través de los lenguajes. Esto es porque el LIWC en español fue traducido y por ende no está exento de sesgos de lenguaje (para más detalles ver Ramírez et al., 2007). Otra limitación de este estudio es que fue basado solamente en mujeres y no se obtuvo otra información sociodemográfica. Finalmente, la muestra es relativamente pequeña. El propósito del Estudio 2 fue abarcar estas limitaciones, al introducir un método que extrae temas que las personas usan cuando hablan de su depresión. Además, este método permitió observar diferencias cualitativas en los temas a través de los lenguajes. Finalmente, en el Estudio 2 se colectaron ensayos de más individuos deprimidos y se codificó otros datos sociodemográficos.

## **Estudio 2: Temas que Utilizan las Personas en Foros de Discusión**

Estudios anteriores en el uso de lenguaje en personas deprimidas se han focalizado en el estilo lingüístico del lenguaje depresivo, en comparación de analizar de *qué* hablan las personas. Mientras que analizar el estilo lingüístico tiene muchas características prometedoras que pueden ser generalizadas a través de los temas, descubrir los temas de conversación en una muestra en particular resulta tener implicaciones de gran alcance. Esto nos puede dar una clave acerca del tipo de información y apoyo que las personas naturalmente buscan de distintos grupos en línea.

Métodos tradicionales de análisis de contenido han explorado textos codificando determinados tópicos. Esto provee una confiabilidad entre codificadores moderada y toma mucho tiempo y esfuerzo y además no está exento de decisiones subjetivas hechas por los codificadores. Mientras que una codificación intensiva puede ser beneficiosa, lo que realmente le importa al investigador es un resumen de los temas que abarcan una gran proporción de los textos. Mediante avances recientes de programas de computadora que analizan textos, ahora es posible resumir los temas más importantes en un corpus de manera rápida y confiable.

En la Universidad de Texas en Austin, hemos desarrollado recientemente el Método de Extracción de Significados (MES, Chung & Pennebaker, 2008). El MES utiliza herramientas de computación automáticas para identificar los temas más comúnmente usados tomando en cuenta solamente las palabras de contenido en los textos escritos y después determina cómo estas palabras co-ocurren en un texto escrito. Para apreciar cómo el MES trabaja, imaginen que 500 estudiantes se les piden que describan su personalidad. Asumamos que cada descripción de personalidad toma 15min y que los estudiantes escriben un promedio de 300 palabras. A pesar de que esta tarea va a generar un promedio de 15,000 palabras diferentes, hay probablemente solamente cerca de 3,000 palabras de contenido diferente usadas. Una vez que se

quitan las palabras de relleno (e.g., artículos, verbos auxiliares, preposiciones y pronombres) así como las palabras que no fueron usadas en por lo menos 3% de los participantes, entonces solamente quedarán 100 palabras. Imaginen que ahora usamos estas palabras para ver si fueron usadas en cada uno de los 500 ensayos y determinamos si la palabra existe (codificada como 1) o no existe (codificada como 0), utilizando una aproximación binaria. Lo resultante sería una matriz hecha de 500 filas y 100 palabras en las columnas.

La matriz de 500 x 100 de *unos y ceros* puede ser sometida a un tradicional análisis factorial (ver Chung & Pennebaker, 2008; Ramírez-Esparza, 2007 para más detalles) con rotación Varimax. El número óptimo de factores que emergen es determinado por el método del *scree plot*. Mediante este último análisis uno puede determinar el grado en que cualquier grupo de palabras tendieron a co-ocurrir en un texto, y además el análisis factorial no dice como se agrupan las palabras de manera coherente. Cada grupo, es esencialmente un significado o un tema. Por ejemplo si una persona usa la palabra *inteligente*, también puede usar las palabras, *rápido*, *abusado*, *trabajador*, *cuidadoso*, *minucioso*. Otra persona puede usar los mismos temas con otros sinónimos. Lo atractivo de este método es que todos estos términos pueden dar un factor que sería definido como un tema de inteligencia.

Hay dos ventajas de este método que lo hace particularmente apropiado para investigación transcultural. Primero, las dimensiones que emergen son puramente inductivas. Las palabras no dependen de los puntajes dados por jueces o pre-definido por investigadores que tienen sus propios sesgos culturales. El método, entonces, evita imponer constructos teóricos o listas de términos en los escritores. Mediante el análisis de textos de personas deprimidas, podemos capturar las conexiones naturales entre las palabras, en lugar de agrupar palabras en una dimensión preseleccionada. En segundo lugar, el método no se fundamenta en traducción. El MES simplemente ve a las palabras—o grupo de palabras separadas por espacios en blanco y comas--y

analiza cómo se agrupan las palabras estadísticamente. El proceso de traducción es solamente relevante al *final* del proceso (Ramirez-Esparza, 2007).

### *Método*

Colección de los Datos. Este estudio utilizó el mismo procedimiento del Estudio 1 para coleccionar los textos. Sin embargo, esta vez se tomaron en cuenta ensayos donde se pudiera saber categóricamente la edad del escritor y el sexo. Se utilizaron 5 foros de depresión para coleccionar los ensayos en inglés y 2 para coleccionar los ensayos en español.

Archivos de texto. La muestra incluyó 404 ensayos en inglés y 400 ensayos en español de foros de discusión de depresión. La Tabla 1 muestra la información demográfica que se codificó de los textos.

---ISERTAR TABLA 1 POR AQUI--- (Cf. al final del trabajo)

Análisis de los Datos. El MES puede ser pensado como un análisis factorial de palabras a través de los textos. Como todo análisis factorial, la técnica permite determinar cómo un grupo de ítems (en este caso palabras) covarían naturalmente. Se hizo un análisis por cada lenguaje.

Todas las palabras para cada lenguaje se enumeraron de acuerdo a frecuencia utilizando WordSmith (Scott, 1996). Las palabras de relleno (i.e. pronombres, preposiciones, artículos, conjunciones, y verbos auxiliares) no fueron consideradas para los análisis. Se creó un diccionario para cada lengua que incluía aproximadamente 200 palabras más frecuentemente usadas y sus raíces de palabras (ver Ramírez et al., 2007, para más detalles de cómo hacer un diccionario para LIWC). Se utilizó LIWC y se corrió el diccionario en inglés para observar la ocurrencia de 197 palabras en los 404 ensayos en inglés. Así mismo, se corrió el diccionario en español utilizando LIWC para observar la ocurrencia de las 195 palabras en los 400 ensayos en español. Lo resultante de los foros de depresión en inglés fue una matriz de 197 palabras por 404 archivos de textos y una matriz de 195 por 400 archivos de textos

para los foros de depresión en Español. Cada matriz estaba compuesta de *unos* (para indicar que ocurrió la palabra en el texto) y *ceros* (para indicar que la palabra no ocurrió en el texto).

Se realizó un análisis factorial de componentes principales usando una rotación varimax para cada matriz. El *scree plot* de Eigenvalues (Cattell, 1966; Stevens, 1992) sugirió que 5 factores fueran extraídos para cada muestra. Cada factor representó un grupo de contenido de palabras que tendieron a co-ocurrir dentro de cada lengua.

### *Resultados*

Las Tablas 2 y 3 muestran el grupo de palabras que se conglomeraron para formar temas coherentes. En ambos lenguajes se encontraron 5 factores, de estos, tres fueron similares a través de las culturas: Tratamiento, Familia, Escuela. Sin embargo, hubo pequeñas deferencias en los temas que se hablaron. Por ejemplo, Tratamiento para la muestra en inglés indica términos más específicamente relacionados a efectos secundarios y de medicamentos (e.g., *efecto, secundario, experiencia, sufrir*) que en la muestra en español. El factor Escuela para la muestra en inglés se asocia más a relaciones (i.e., utilizaron palabras tales como *escuela, universidad, colegio*, con palabras tales como *relación, amigo, novia, social*). En contraste, en español los temas de la escuela están más asociados a temas de inseguridad (i.e., utilizaron palabras tales como *colegio, estudios, universidad* con palabras tales como *tímida, incapaz, confianza*). El factor Familia para la muestra en inglés se mencionan en general roles específicos (e.g., *mamá, hija, hermano, hermana*), en cambio en la muestra en Español los términos de familia van unidos a temas específicos como lo indica incluir términos tales como *dinero, estudios, comer y suicidio*.

Los temas que fueron únicos de la muestra en inglés fueron de Divulgación y Síntomas. Divulgación mostró palabras asociadas a querer compartir sentimientos con otros (e.g., *siento, feliz, gente, decir*). El factor de Síntomas habla detalladamente del proceso de dormir (e.g., *noche, dormir, despertar*) y hábitos alimenticios (e.g., *comer*,

*comida, peso*), lo cual ambos son conocidos dentro de episodios depresivos (APA, DSM-IV, 1994).

Los temas que son únicos para la muestra en español son Relaciones y Desesperanza. El factor de Relaciones se asoció con la crónica de una relación específica como lo indica el uso de palabras de relaciones humanas (e.g., *amigos, chico, novio*), con palabras temporales (e.g., *conocí, empecé, dejar, meses, final*) El factor Desesperanza incluye palabras relacionadas a la pérdida de la esperanza (e.g., *morir, terrible, ojos, noche, segundo*) con muestras de esperanza (e.g., *paz, espero, fuerte, corazón, sueños*).

Un análisis cualitativo de los resultados muestra que en los foros de depresión en inglés las personas usan una voz activa. Por ejemplo, mencionan quiénes son ellos y qué rol tienen en su contexto (i.e., en el factor Familia), hablan de cómo experimentan su depresión (i.e., factor Síntomas), qué tratamiento están llevando a cabo, y qué medicinas están tomando (i.e., factor Tratamiento). Así mismo mencionan de una manera activa que están visitando el foro para desahogarse, buscar ayuda y apoyo de las otras personas (i.e., en el factor Divulgación). En cambio en los foros de depresión en español las personas hablan de su depresión de una manera más pasiva. En general hablan más de sus problemas, acontecimientos de vida, y sentimientos que los síntomas de su depresión y tratamiento o actividades que hacen para mejorarse. Por ejemplo, hablan de los problemas familiares (i.e., en el factor Familia), de acontecimientos con relaciones (i.e., en el factor de Relaciones), de sus sentimientos de inseguridad y de desesperanza (i.e., en los factores Escuela y Desesperanza respectivamente).

---INSERTAR TABLA 2 Y 3 POR AQUI--- (Cf. al final del trabajo).



### *Investigaciones Futuras*

La meta de futuras investigaciones deberá ser analizar el uso del lenguaje de aquellas personas que experimentaron mejoría en corto tiempo vs. quienes necesitaron un periodo largo de tiempo participando en foros en línea para personas deprimidas; y ver si hay una relación entre el tiempo (corto o largo) y el lenguaje utilizado: activo o pasivo; ya que esta diferencia en el lenguaje implica actitudes diferentes hacia su estado emocional. Como ya se mencionó, los que utilizan un lenguaje activo parecen tener más claros sus síntomas, los tratamientos y medicamentos adecuados para tratar una depresión. Mientras que los que utilizan un lenguaje considerado como pasivo se refieren más a sus sentimientos, problemas, relaciones con las personas cercanas, sentimientos de inseguridad y desesperanza; en otras palabras, utilizan el foro para desahogarse de las emociones negativas propias de la depresión.

Podrían también implementarse nuevas investigaciones donde participen terapeutas en los foros sobre depresión, que a nivel coloquial puedan dar una orientación sobre lo que es la *regulación emocional* y muy particularmente manejada a través de la terapia cognitivo-conductual, ya que existe suficiente información de que está dentro de las mejores opciones (Beck, 1976; Ellis, 1977). La meta de esta terapia es cambiar los síntomas de los problemas psicológicos a través de modificar directamente el pensamiento disfuncional que acompaña a esos síntomas.

Examinar la conducta natural y los temas discutidos en los foros en línea, deberá ser el primer paso para entender la información y la ayuda necesitada a través de los pacientes de diferentes poblaciones y culturas.

### *Conclusiones*

El Internet puede ser usado como una fuente para estudiar comportamientos diarios y naturales. Así como nos da entrada a poblaciones que solamente tenemos acceso de manera limitada en la práctica clínica o en el laboratorio. Interesantemente,

este estudio encuentra que las características de la depresión en línea son similares a lo que se encuentra en otros análisis que se han hecho en textos escritos en laboratorio o en estudios clínicos. Las estrategias de análisis de texto usadas en este artículo nos permitieron caracterizar el lenguaje que se usa por personas que visitan foros de depresión en línea y además analizar cómo los temas son mencionados de forma natural a través de las culturas. Se recomienda que los marcadores lingüísticos y tópicos de interés sean analizados en otras poblaciones clínicas y en otras culturas.

### Referencias

- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders*. (4th ed.). Washington, DC: Author.
- Baker, L., Wagner, T. H., Singer, S., & Bundorf, M. K. (2003). Use of The Internet and E-Mail for Health Care Information. *Journal of the American Medical Association*, 289, 2400-2406.
- Barr-Zisowits, C. (2004). *Sadness-Is there such a thing?*. En M. Lewis, & J. Haviland-Jones, (Eds.): *Handbook of Emotions* (pp.607-622). New York: The Guilford Press.
- Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects*. New York: Harper & Row.
- Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. New York: International University Press.
- Berger, M., Wagner, T. H., & Baker, L. C. (2005). Internet Use and Stigmatized Illness. *Social Science and Medicine*, 61, 1821-1827.
- Bucci, W., & Freedman, N. (1981). The Language of Depression. *Bulletin of the Menninger Clinic*, 45, 34-358.
- Catell, R. B. (1966). The Scree Test for the Number of Factors. *Multivariate Behavioral Research*, 1, 245-276.

- Chung, K. C., & Pennebaker, J. W. (2008). Revealing Personality in Natural Language: Using an Automated Meaning Extraction Method on Open-Ended Self-Descriptions. *Journal of Research in Personality, 42*, 96-132.
- Davison, K. P., Pennebaker, J. W., & Dickerson, S. S (2000). Who talks? The social psychology of illness support groups. *The American Psychologist, 55*, 205-217.
- Durkheim, E. (1951). *Suicide*. New York, NY: Free Press.
- Ellis, A., & Grieger, R. (Eds.). (1977). *Handbook of rational emotive therapy*. New York: Springer.
- Fox, S., & Fallows, D. (2003, July 16). Internet health resources. Retrieved June 03, 2007, from [http://www.pewinternet.org/pdfs/PIP\\_Health\\_Report\\_July\\_2003.pdf](http://www.pewinternet.org/pdfs/PIP_Health_Report_July_2003.pdf)
- Houston, T. K., Cooper, L. A., & Ford, D. E. (2002). Internet Support Groups for Depression: A 1-year prospective cohort study. *American Journal of Psychiatry, 159*, 2062-2068.
- Kleinman, A. & Good, B. (Eds.). (1985). *Culture and depression*. Berkeley: University of California Press.
- Lieberman, M. A., & Goldstein, B. A. (2006). Not All Negative Emotions are Equal: The Role for Emotional Expression in Online Support Groups for Women with Breast Cancer. *Psycho-Oncology, 15*, 160-168.
- Lyons, E. J., Mehl, M. R., & Pennebaker, J. W. (2006). Pro-Anorexics and Recovering Anorexics Differ in their Linguistic Internet Self-Presentation. *Journal of Psychosomatic Research, 60*, 253-256.
- Owen, J. E., Yarbrough, E. J., Vaga, A., & Tucker, D. A. (2003). Investigation of the Effects of Gender and Preparation on Quality of Communication in Internet Support Groups. *Computers in Human Behavior, 19*, 259-275.

- Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). *Linguistic Inquiry and Word Count: LIWC 2001*. Mahwah, NJ: Erlbaum Publishers ([www.erlbaum.com](http://www.erlbaum.com)).
- Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology*, *54*, 547-577.
- Pyszczynski, T., & Greenberg, J. (1987). Self-Regulatory Perseveration and the Depressive Self-Focusing Style: A Self-Awareness Theory of Depression. *Psychological Bulletin*, *102*, 122-138.
- Ramírez-Esparza, N. (2007). *Cross-Cultural Constructions of Self: Americans and Mexicans*. Ph.D. diss., Dept of Psychology, University of Texas, Austin, TX.
- Ramírez-Esparza, N., & Pennebaker, J. W. (2006). Do good stories produce good health?  
Exploring words, language, and culture. **Narrative Inquiry**, *16*, 221-229.
- Ramírez-Esparza, N., Pennebaker, J. W., García, A. F., & Suriá, R. (2007). La Psicología del Uso de las Palabras: Un Programa de Computadora que Analiza Textos en Español. *Revista Mexicana de Psicología*, *24*, 85-99.
- Rude, S. S., Gortner, E. M., & Pennebaker, J. W. (2004). Language Use of Depressed and Depression-Vulnerable College Students. *Cognition & Emotion*, *18*, 1121-1133.
- Scott, M. (1996). *WordSmith*. New York, NY: Oxford University Press.
- Stevens, J. P. (1992). *Applied Multivariate Statistics for the Social Sciences (2<sup>nd</sup> ed.)*. Hillsdale, NJ: Erlbaum.
- Stirman, S. W., & Pennebaker, J. W. (2001). Word Use in the Poetry of Suicidal and Non-Suicidal Poets. *Psychosomatic Medicine*, *63*, 517-522.

- Stjernsward, S., & Ostman, M. (2006). Potential of E-Health in Relation to Depression: Short Survey of Previous Research. *Journal of Psychiatric and Mental Health Nursing*, 13, 698-703.
- Weintraub, W. (1981). *Verbal Behavior: Adaptation and Psychopathology*. New York, NY: Springer.
- Wolpert, L. (1999). *Malignant Sadness: the anatomy of depression*. New York: Simone & Schuster, Inc. Free Press.

Tabla 1. Información Demográfica para los Foros de Depresión en Inglés y Español.

Foro de Depresión	Sexo		Edad		
	Mujeres	Hombres	<25	25 a 45	>45
Inglés	290	114	168	214	22
Español	252	148	166	207	26

Nota: Los números en las celdas indican frecuencias. Edad, <25 = más joven de 25, 25 a 45 = 25 a 45 años de edad, >45 = más de 45 años.

Tabla 2. Factores Resultantes de los Foros de Discusión de Depresión en Inglés

<b>Factor 1:</b>		<b>Factor 2:</b>		<b>Factor 3:</b>		<b>Factor 4:</b>		<b>Factor</b>	<b>5:</b>
<b>Tratamiento</b>		<b>Divulgación</b>		<b>Familia</b>		<b>Síntomas</b>		<b>Escuela</b>	
MEDICATION/	.62	TELL/	.43	MOM/	.49	SLEEP/	.51	CONSTANT/	.46
Medicamento		Decir		Mamá		Dormir		Constantemente	
EFFECT/	.46	PEOPLE/	.41	DAUGHTER/	.48	HOUR/	.48	RELATIONSHIP/	.45
Efecto		Gente		Hija		Hora		Relación	
DEPRESSION/	.43	KNOW/	.39	CHILD/	.48	FOOD/	.44	SCHOOL/	.40
Depresión		Conocer		Niño/niña		Comer		Escuela	
SIDE/	.35	HAPPY/	.35	FAMILY/	.48	WAKE/	.44	HIGH/	.40
Secundario		Feliz		Familia		Despertar		Prepa	
WEEK/	.34	TALK/	.35	BROTHER/	.43	MORNING/	.44	LACK/	.39
Semana		Hablar		Hermano		Mañana		Falta	
THERAPY/	.34	FEEL/	.34	SISTER/	.43	NIGHT/	.41	UNIVERSITY/	.38
Terapia		Sentir		Hermana		Noche		Universidad	
SUFFER/	.34	WANT/	.34	DAD/	.41	BED	.39	SOCIAL/	.38
Sufrir		Quiero		Papá		Cama		Social	
DISORDER/	.33	SUPPOSE/	.33	SON/	.33	STAY/	.38	COLLEGE/	.36
Trastorno		Suponer		Hijo		Quedar		Colegio	

DOCTOR/	.33	READ/	.32	LOVE/	.33	WEIGHT/	.37	MOVE/	.32
Doctor		Leer		Amor		Peso		Mover	
ANTIDEPRESSANT/	.32	HURT/	.32	GIRL/	.33	EAT/	.36	FRIEND/	.32
Antidepresivo		Herido		Niña		Comer		Amigo	
EXPERIENCE/	.32	WRONG/	.32	YOUNG/	.32	PLACE	.32	GIRLFRIEND/	.32
Experiencia		Equivocar		Jóven		Lugar		Novia	
MAJOR/	.32	EMOTIONAL/	.31	PARENT/	.32			CLASS/	.31
Mayor		Emocional		Padre				Clase	
MENTAL/	.31	MIND/	.31	HOUSE/	.31				
Mental		Mente/importa		Casa					
PSYCHIATRIST/	.31	SAD/	.31	HUSBAND/	.30				
Psiquiatra		Triste		Esposo					
		MAKE/	.31	CRAZY/	.30				
		Hacer		Loco					
		HATE/	.31						
		Odiar							

---

Nota: Solamente se reportan los valores mayores a .30. Traducción en minúsculas



Tabla 3: Factores Resultantes de los Foros de Discusión de Depresión en Español

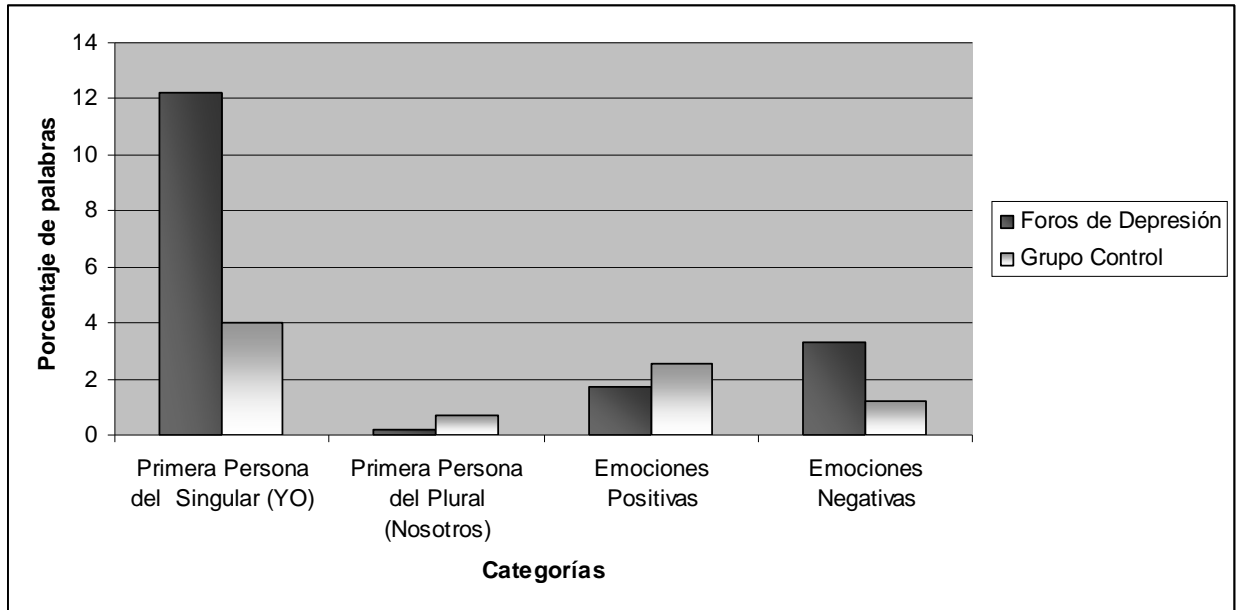
<b>Factor 1:</b>		<b>Factor 2:</b>		<b>Factor 3:</b>		<b>Factor 4:</b>		<b>Factor 5:</b>	
<b>Familia</b>		<b>Relación</b>		<b>Desesperanz a</b>		<b>Escuela</b>		<b>Tratamiento</b>	
MADRE	.48	RELACIÓN	.51	NOCHE	.45	TIMIDA	.41	PSICÓLOGO	.40
HERMAN O	.46	ENAMORA DO	.48	SEGUN DO	.44	CARRERA	.41	ANSIEDAD	.40
ABUELA	.44	CONOCÍ	.45	MORIR	.42	COLEGIO	.40	EMPRESA	.39
PADRES	.44	HABLAR	.45	PAZ	.39	CONFIANZA	.37	ANTIDEPRESIV OS	.39
PAPÁ	.43	CHICO	.44	OJOS	.39	ESTUDIOS	.37	SÍNTOMAS	.34
HORRIBL E	.39	AMIGOS	.44	ESPER O	.38	INCAPAZ	.33	MEDICAMENT O	.30
CASA	.39	NOVIO	.43	TERRIB LE	.36	UNIVERSIDA D	.33		
SUICIDIO	.36	JUNTOS	.37	FUERTE	.34	TONTERÍA	.32		
DINERO	.34	ESPECIAL	.36	CORAZ ÓN	.31				
ESTUDIO S	.33	EMPECÉ	.33	SUEÑO S	.31				

CLASE	.32	TIEMPO	.33
FEA	.32	DEJAR	.33
ASCO	.32	FINAL	.33
COMER	.31	HISTORIA	.31
PEQUEÑ O	.31	MESES	.31
FAMILIA	.31	LLEGAR	.31
HIJOS	.30		

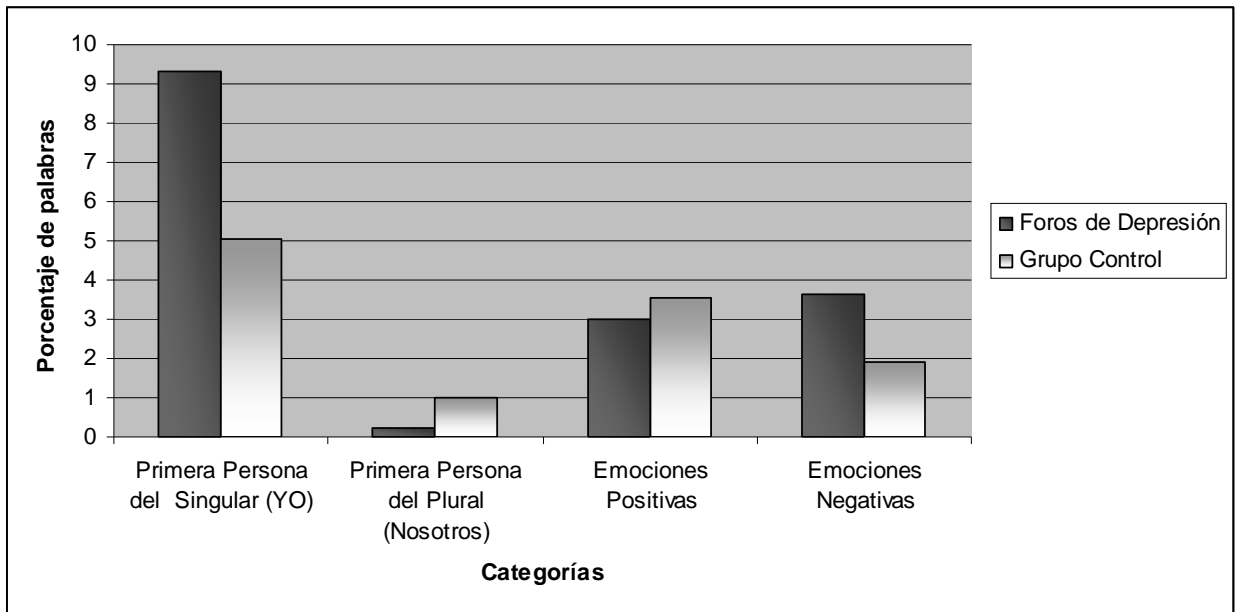
---

Nota: Solamente se reportan los valores mayores a .30

Figura 1: Marcadores Lingüísticos de Foros de Depresión en Inglés



Nota: Todas las pruebas *t*-independiente son de dos colas y se basan en 158 grados de libertad. Las diferencias para cada categoría son significativas al .001



Nota: Todas las pruebas *t*-independiente son de dos colas y se basan en 158 grados de libertad. Las diferencias para cada categoría son significativas al  $p < .001$ , excepto por Emociones Positivas cuya significancia es  $p < .05$

<sup>1</sup>University of Washington

I-LABS, Box 357988

Seattle, WA 98195

nairan@u.washington.edu

<sup>2</sup>University of Texas at Austin

<sup>3</sup>Universidad Nacional Autónoma de México