

Estudios Filológicos Alemanes

Revista de investigación
en Lingüística, Literatura y Cultura alemanas

Volumen 25

SEPARATA

Sevilla, 2012

ANÁLISIS LINGÜÍSTICO DE TUIITS

MARIO FRANCO
UNIVERSITÄT MADEIRA
PORTUGAL

Recibido: 21/06/2012

Aceptado: 07/09/2012

Zusammenfassung

In diesem Beitrag werde ich meine Teilnahme am internationalen „Twitter-Projekt“ der Universität Hannover beschreiben. Das Projekt wurde von Torsten Siever und Peter Schlobinski geleitet.

Es wurden Tweets in fünfzehn Sprachen analysiert, wobei ich die Analyse der spanischen Tweets übernahm. Die Teilnehmer bekamen ein vom Deutschen ausgehendes Analyseraster mit unterschiedlichen Spach- und Textmerkmalen. Dieses Raster musste ich im Hinblick auf das Spanische modifizieren. Es wurden sowohl spanische Merkmale ergänzt, als auch deutsche Merkmale, die im Spanischen nicht vorkommen, nicht angewandt.

Neben diesem kontrastiven Aspekt werde ich auch die Ergebnisse meiner Analyse bekannt geben. Ich werde vor allem die Merkmale erläutern, bei denen die Zahl der Belege bei männlichen und weiblichen Twitter-Usern sehr unterschiedlich ist.

Schlüsselwörter: Twitter - Tweets - Sprache in den Neuen Medien - Nähesprechen

Resumen

En esta comunicación describiré mi participación en el proyecto internacional de Twitter (“Internacionales Twitter Projekt”), coordinado por Torsten Siever y Peter Schlobinski, de la Universidad de Hannover. En este proyecto fueron analizados tuits en quince lenguas, encargándome yo del análisis de los tuits en lengua española.

A los participantes del proyecto se les entregó un listado con características lingüísticas divididas en diferentes categorías en alemán. Una vez que comenzamos con el análisis, pronto nos dimos cuenta de que era necesario ampliar la lista con características lingüísticas más propias del español. Asimismo, nos percatamos rápidamente de que ciertos fenómenos lingüísticos típicos del alemán no se daban en el español.

También expondré en este artículo los resultados obtenidos de aquellas características lingüísticas en las que el número de ejemplos encontrados sea significativamente diferente entre tuits escritos por usuarios femeninos y aquellos escritos por usuarios masculinos.

Palabras clave: Twitter, Tuits, Lenguaje en los nuevos medios de comunicación, Lenguaje de proximidad

1. Breve descripción del proyecto

El proyecto „Twitter International“ fue dirigido por Torsten Siever y Peter Schlobinski, de la Universidad de Hannover. Su objetivo fue analizar tuits de la red social Twitter <http://twitter.com> en diferentes lenguas. En total fueron analizados

640 tuits en cada lengua, 10 por cada usuario, 32 hombres y 32 mujeres. Las lenguas analizadas son: alemán, alemán de Suiza, chino, español, francés, holandés, inglés, italiano, japonés, polaco, portugués y ruso; habiendo analizado el autor de este artículo los tuits españoles.

Los parámetros de análisis son: ortografía, lengua hablada, palabras, formas reducidas, sintaxis, grafo-estilística, interacción, extensión de las anotaciones, plurilingüismo y aspectos funcionales.

El proyecto será publicado en formato libro y e-book bajo el título *Microblogs global: Eine internationale Studie zu Twitter & Co. Aus der Perspektive von Sprachen und Ländern*.

La recogida de datos se realizó entre Agosto de 2010 y Marzo de 2011. En el buscador *tweepz.com* fueron introducidos los nombres de 32 ciudades españolas (incluyendo por lo menos una ciudad de las diecisiete comunidades autónomas y las dos ciudades autónomas, Ceuta y Melilla), incorporando a nuestro corpus tanto el primer usuario femenino como masculino. En el caso de cuentas de empresas u organizaciones, fueron integradas en el corpus aquellas en las cuales constaba el administrador, indicando sus datos.

2. Análisis

Por cuestión de espacio y por no ser la finalidad de este artículo, no comentaré la totalidad de parámetros, ni todas las características lingüísticas que tuvimos en cuenta en nuestro análisis. Trataré únicamente los siguientes parámetros: ortografía, oralidad concepcional, economía lingüística y funcional, sintaxis, características grafo-estilísticas, interacción y aspectos funcionales.

2. 1. Ortografía

La gran mayoría de tuits, un 87% del total, presenta una escritura estándar. Tuits escritos en su totalidad en minúsculas representan un 12,7%, por el contrario, los escritos en su totalidad en mayúsculas sólo son un 0,5%. Suponemos que la escritura en minúscula se debe a que favorece la rapidez a la hora de escribir. Por el contrario, la escritura en mayúsculas es poco representativa. De aquí interpretamos que la letra mayúscula dificulta la lectura tanto en Twitter como en otras nuevas formas de comunicación (e-mail, chat, sms ...), y que además se interpreta como que el usuario grita.

En cuanto a los signos de puntuación, hay que destacar que un 25,1% de las palabras que deberían llevar la tilde no la llevan. El 62,5% de las frases interrogativas no tienen el signo inicial y en un 6,6% faltan incluso ambos signos de interrogación. No se utiliza la coma en un 27,8% de todos los casos donde debería ponerse, en el caso del punto el porcentaje es más alto: 63,9%.

Una violación de la regla ortográfica que, presumiblemente en la mayoría de los casos se debe a la limitación de los 140 caracteres, es el escribir las palabras juntas sin dejar espacio. Algunos ejemplos de este fenómeno en nuestro corpus son *twitteandoenlos90*, *huellasocial* ... Muchos ejemplos se dan en los nicks o seudónimos de los usuarios o en los hashtags¹ que, en Twitter, son cadenas de caracteres formadas por una o más palabras concatenadas y precedidas por el símbolo #. Ejemplos de nicks consistentes en una cadena de palabras serían p. ej.: *@soscismobiz*, *@MarinaNavas* o *@empleocero*. Los hashtags representan temas que cualquier usuario puede crear o temas que ya están abiertos por otros usuarios a los cuales podemos contribuir aportando nuestra opinión incluyendo en nuestro tuit el símbolo almohadilla y la cadena de caracteres que dan nombre a ese tema, p. ej.: “*#cancionesquemarcaron* Marta tiene un marcapasos. Hombres G (no me lo neguéis, todos la hemos cantado alguna vez)”. El usuario podrá buscar después la cadena *#cancionesquemarcaron* y este mensaje estará presente en los resultados de la búsqueda junto con otros mensajes con el mismo hashtag. Los hashtags más utilizados también se muestran en algunas páginas web de *trending topics* (temas del momento) así como la propia página de inicio de Twitter. Los hashtags no distinguen entre mayúsculas y minúsculas, pero a menudo se emplea el fenómeno llamado *CamelCase* o en alemán *Binnenmajuskel* (o uso de mayúsculas en una posición distinta de la inicial) para mejorar la legibilidad p.ej.: *#RedesSociales*. En ocasiones, un hashtag puede pasar a convertirse en un fenómeno de Internet a gran escala.

Llama también la atención el porcentaje de los nombres propios se escriben en minúscula, el 27,5% del total.

2. 2. Oralidad Concepcional

Con oralidad concepcional nos estamos refiriendo al concepto lingüístico “Konzeptionelle Mündlichkeit” introducido por los romanistas alemanes Koch/Oesterreicher, que distinguen entre medio y concepción. Así, existen formas de comunicación escritas (medio), pero concebidas de forma oral (p.ej. una entrevista en una revista, una conversación en un chat ...), o, por el contrario,

¹ Del inglés *hash* (almohadilla) nombre del símbolo # que precede al *tag* (etiqueta).

formas de comunicación orales (medio), pero concebidas de forma escrita (p. ej. el sermón de un cura en una misa, el discurso de un político ...). Las formas de comunicación que, en el lenguaje utilizado por sus usuarios, muestran más características del lenguaje hablado se aproximan, en el continuo presentado por Koch/Oesterreicher (1985, p. 18), más al polo del *lenguaje de proximidad*. Así, una conversación familiar sería el prototipo de forma de comunicación del *lenguaje de proximidad*. Por el otro lado, un texto jurídico, que tanto en su concepción como su medio es escrito, sería prototipo del polo del *lenguaje de la distancia*.

Debemos indicar que los tuits aquí analizados muestran características del lenguaje hablado, a pesar de ser Twitter un medio escrito,. No obstante, debemos señalar que se trata de un número sensiblemente menor que el de los tuits escritos en alemán.

Uno de los fenómenos más frecuentes es el de la ausencia de la forma verbal conjugada en construcciones verbales complejas como por ejemplo: “*A media hora de trabajar*” o “*Cambiando diseño de Twitter*”.

También parecen en nuestro corpus casos de estructuras de *Operador-Escopo* (“Operator-Skopus-Struktur”). Los lingüistas alemanes Fiehler/Barden/Elstermann/Kraft (2004, 262) definen la función de un *operador* con la expresión *vorausgehende Verstehensanleitung*, expresión que podemos traducir como instrucción previa para la descodificación por parte del oyente del contenido en el escopo del enunciado. Puede parecer complicado, pero simplemente significa que el *operador* ejerce influencia sobre el oyente en el acto de descodificación del mensaje en el escopo del enunciado, al cual se refiere el operador. Entendemos esta función como *instrucción previa* (SIEBERG 2011, p. 56) de los operadores en el sentido de una prioridad estratégica en la orientación del oyente, en el proceso de descodificación del mensaje. Veamos los siguientes tres ejemplos de nuestro corpus: (1) “*Ahora, la acidez de las patatas de Blas del día anterior casi me mata*”; (2) “*pobrete el mal que te estoy dando*”; (3) “*Por fin: ya puedes reservar habitación, Baño Relajante, desayunos o Jacuzzi privado desde Facebook!*”. En (1), *Ahora* anticipa que el emisor va a realizar una pequeña objeción en el escopo a algo dicho en un tuit anterior. En (2), *pobrete* anticipa la disculpa al receptor en el escopo. En (3), *Por fin* indica previamente al receptor que el emisor esperaba y deseaba la información que viene después en el escopo.

También destacamos aquí la presencia en nuestro corpus de lo que, en la investigación del alemán hablado, se denomina *construcción en el margen de la frase* (*Konstruktion am Satzrand*), categoría que abarca varias formas con particularidades sintácticas y clasificaciones complejas (HENNIG 2006, p. 166 y ss.). Teniendo esto como base, en el ejemplo “*el corte ingles de constitución. Pero el de*

enfrente el de libros/informática” interpretamos *el de enfrente* y *el de libros/informática* como una segunda realización del constituyente *el corte ingles de constitucion*. De forma que podemos clasificar esta estructura en análisis como *desplazamiento a la izquierda* (*Linksversetzung*) y en relación a su estatuto de *unidad básica del habla*, como elemento de la *estructura frásica interna*. Otros ejemplos: “*quieres café mi vida? Pues sírvetelo! @grillismo ;) en casa lo decíamos con todo el cariño del mundo*”; “*Tarde de viernes, hoy no es día para agobios de empleo*”; “*viendo documentos históricos con hell girl, alguno sacaremos a la luz*”.

En el siguiente ejemplo “*En la estación del Bus. Se me hace muy extraño que nadie me acompañe*” se trata, al contrario de los ejemplos anteriores, de un sintagma que en nuestra opinión pertenece al grupo de *Freies Thema* (IBID, p. 172 y ss.), término que, en español, corresponde a *tópico discursivo*, dado que la estructura *En la estación del bus* constituye apenas una *proyección pragmática* en relación a la oración más próxima, pero no retoma un constituyente interno de esta oración, como sería el caso de una *proyección sintáctica*. En la gramática española ambas estructuras forman parte del concepto de *tópico*, que está clasificado en *inicial*, *final*, *medial* y *contrastivo* (vinculo formal) y *discursivo* (vinculo funcional) (REAL ACADEMIA ESPAÑOLA 2010, p. 755 y ss.). No es difícil constatar que las estructuras y las funciones de este tipo de arquitectura frásica son analizadas bajo el prisma de la frase escrita canónica y sus normas sin reconocer, por otro lado, sus potencialidades funcionales en la comunicación oral.

Por último, indicar que tenemos menos ejemplos de otras características del lenguaje hablado, como p.ej. elisiones *stoy* (en vez de “estoy”), *pa* (en vez de “para”) ..., palabras emotivas “*Como pipas por la calle. Mola!*”, “*Dioos*”, *Puto teléfono, ha sonado 6 veces seguidas*, “*Cómo runeo, CÓMO RUNEO!*” ... o formas diminutivas de sustantivos, adjetivos y adverbios (*pelín, fresquito, despacito* ...). En el alemán hay muchos más ejemplos de ‘*palabras fónicas*’ (*phonisches Wort*), como p. ej. la apócope del Schwa final -e en la 1ª persona del singular del presente de indicativo de los verbos monosilábicos: *hab, glaub, komm, werd* ...o ejemplos de asimilación *drauf, kommste, hamwa*

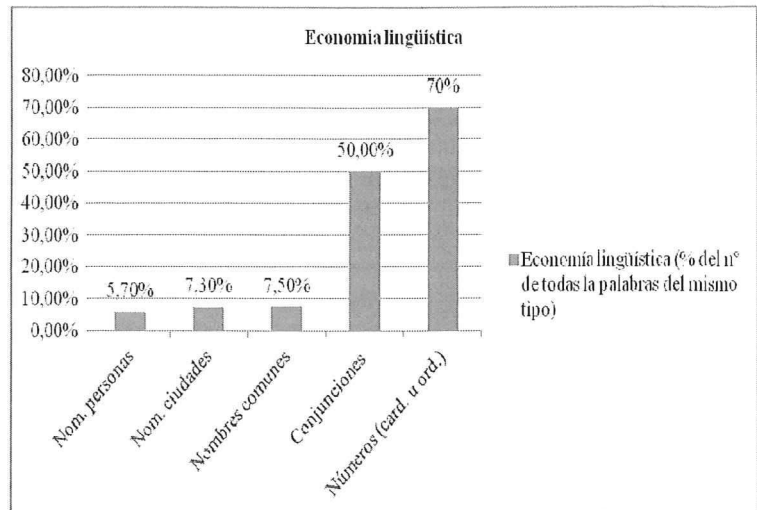
2. 3. Economía lingüística

En un medio como Twitter, donde la longitud del texto está restringida a sólo 140 caracteres, hace que los usuarios utilicen formas reducidas para no traspasar el límite. Conjunciones y nombres propios son los tipos de palabras que más se reducen en el corpus analizado: *q* o sus variantes *k, ke* (en vez de “que”), *pq* o su variante *xq* (en vez de “porque”) o *tb* (en vez de “también”); *Isa* (en vez de Isabel), *Lau* (en vez de Laura), *Dany* (en vez de Daniel), *VLC* (en vez de Valencia), *BCN*

(en vez de Barcelona) o Man (en vez de Manchester). Pero sobre todo destaca la reducción de caracteres utilizando cifras en vez de letras, p. ej. “83 pleno del Comité de las Regiones.” *Puto telefono... ha sonado como 6 veces seguidas!*”

Otra forma de ahorrar caracteres es utilizando funciones típicas de la propia forma de comunicación, como son los hashtags y URL², que es un enlace destino en internet, también conocido como *link*, que puede invocar un documento, por ejemplo una página web, otro recurso, o a una posición determinada en una página web. Estos dos fenómenos aparecen en casi el 50% de todos los tuits.

No obstante, en general, nos sorprendió no haber encontrado un número mayor de ejemplos y mayor variedad de formas reducidas en Twitter.

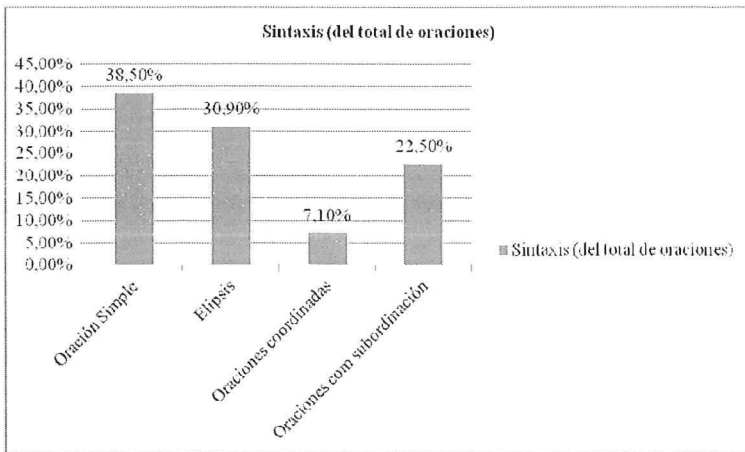


2. 4. Sintaxis

El tipo de oración más frecuente en el corpus es la oración simple (casi un 40 % de todas las oraciones). También el porcentaje de elipsis, es decir, construcciones sin verbo finito explícito, es considerablemente alto: un 31%. Estos datos confirman que la comunicación en Twitter se caracteriza por estructuras sintácticas sencillas, usando más oraciones simples que compuestas por subordinación, lo que la sitúa

² Sigla en inglés de *uniform resource locator*.

más cerca de la lengua hablada³ y por consiguiente del polo de la proximidad en el continuo de Koch/Oesterreicher. En este mismo sentido Ágel/Hennig (2006, pp. 26 y ss.) argumentan que la lengua de proximidad se caracteriza sintácticamente por estructuras *agregativas* en vez de *integrativas* (*'Aggregation vs. Integration'*), que están más organizadas (semántica y pragmáticamente) de forma coherente que (estructuralmente) de forma cohesiva. Ejemplos de estructuras agregativas son las estructuras de *operador-escopo* anteriormente comentadas y estructuras como la siguiente “*ahí sólo veo los que he recibido. No salen los que envié.*” (estructura agregativa) versus “*No salen los que envié, porque ahí sólo veo los que he recibido.*” (estructura integrativa).



2. 5. Características grafo-estilísticas

Con la aparición de los nuevos media, aparecieron en diferentes formas de comunicación, como en el chat, e-mail o en los mensajes de texto de teléfonos móviles, una serie de combinaciones de caracteres, llamados smileys o emoticonos que, girando la cabeza 90° para la izquierda para poder reconocerlos, sirven para compensar la falta de gestos, mímica o emociones en la comunicación. Otro recurso grafo-estilístico muy recurrente por parte de los usuarios es la iteración de letras, muy común en los textos de los comics y que sirve para enfatizar lo dicho. Si la

³ “(...) la sintaxis de la lengua hablada suele estar mucho menos estructurada que la de la lengua escrita (...) la lengua hablada suele contener menos subordinación” (BROWN/YULE, p. 35).

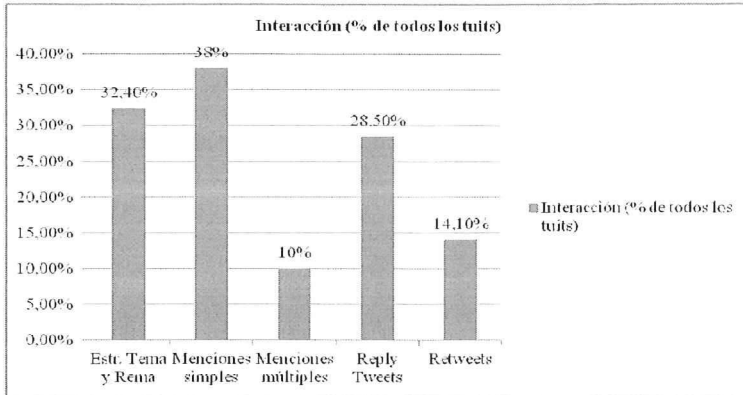
iteración está en letras mayúsculas sirve, a la vez, para dar a entender que el que escribe está chillando. Dado el número limitado de caracteres disponibles en un tuit, creíamos que la utilización de los recursos grafo-estilísticos iba a ser considerable, sin embargo, tan sólo el 1,8% del total de palabras son en realidad smileys, y sólo en el 1,5% de la totalidad de las palabras contienen iteración.

2. 6. Interacción

A pesar de ser “mini-blog” una de las denominaciones de Twitter, en varios puntos muestra más similitudes con el chat que con el blog. Razón de ello son el Timeline (TL), en el cual los tuits del autor y de las personas que sigue van apareciendo y, a la vez, desapareciendo, consonante a la actividad de los usuarios, cuanto mayor sea esta, más activo es el TL, y en consecuencia, con mayor rapidez van bajando los tuits. A partir del TL vemos qué usuarios están online⁴, lo que nos permite interactuar con ellos a partir de la opción de reply, o retwittear un tuit que nos parezca interesante a todos nuestros seguidores.

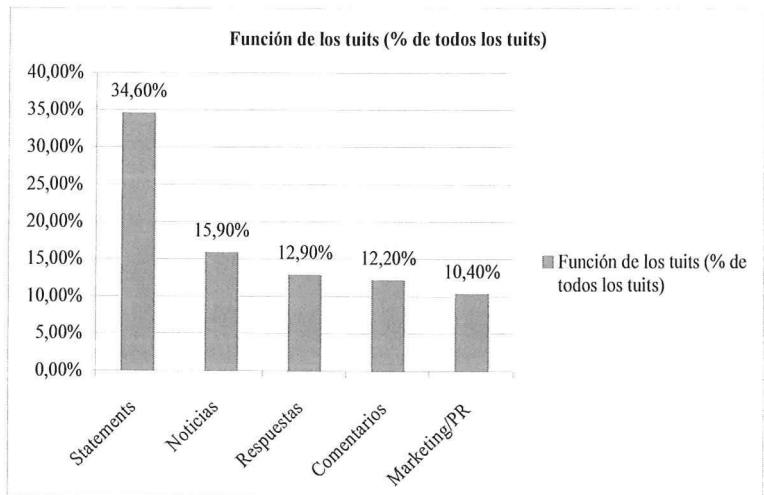
El 38% de todos los tuits tiene una mención simple, es decir, al mencionar el nombre de un usuario nos dirigimos a él con nuestro tuit, o queremos que nuestro tuit, al mencionarlo, aparezca en su TL. En el 10% de todos los tuits hay una mención múltiple, es decir, nos dirigimos a más de un usuario o queremos que nuestro tuit aparezca en sus respectivos TL para que puedan leerlo. El 28,5% de los tuits son respuestas directas – *reply-tweet* – al tuit de otro usuario, creándose así una conversación. El 14,1% de los tuits son *retweets*, es decir, compartir un mensaje de un usuario que nos parece interesante con nuestros followers.

⁴ Los tuits muestran cuando fueron publicados, pueden ser p.ej. de hace sólo segundos, minutos, horas o días.



2. 7. Función de los tuits

Según su función, destacamos aquí cinco tipos de tuits. Los más frecuentes, con un 34,6% de todos los tuits, son los statements o declaraciones de los usuarios sobre su estado de ánimo, sobre lo que están haciendo o sobre cualquier otro motivo de índole personal. Este aspecto funcional se asemeja mucho a la de los blogs del tipo diario personal. Casi un 16% de los tuits son noticias, o incluyen un link o un hashtag de una noticia que para el usuario es interesante. Un 13% son respuestas a tuits de otros usuarios y un 12,2% son comentarios del usuario sobre diversos temas: tuits de otros usuarios, una noticia, un tema actual Un 10,4% de los tuits son marketing o publicidad sobre un site, una empresa ... del propio usuario, que incluye casi siempre su respectivo link.



3. Breve comparación entre tuits de hombre y mujeres

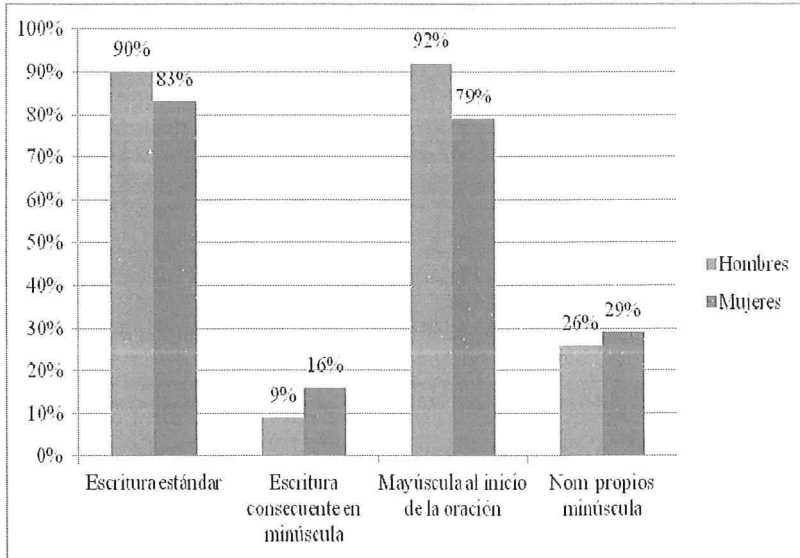
Sin ser el objetivo del proyecto y sin poder llegar a conclusiones generales, dado que el número de usuarios analizados en el corpus es restringido, queremos dejar aquí constancia de algunas diferencias lingüísticas entre hombres y mujeres en el uso de Twitter.

Los resultados relativos a la ortografía demuestran claramente que las mujeres transgreden más las normas, siendo los hombres más conservadores. Llamamos la atención los altos porcentajes en la falta de uso del primer signo de interrogación (77% de todos los casos posibles), ausencia del punto (59%), ausencia de la coma (34%) y la ausencia de acento gráfico (33%). También es significativo el porcentaje de los nombres propios escritos en minúscula, 29% de todos los ejemplos.

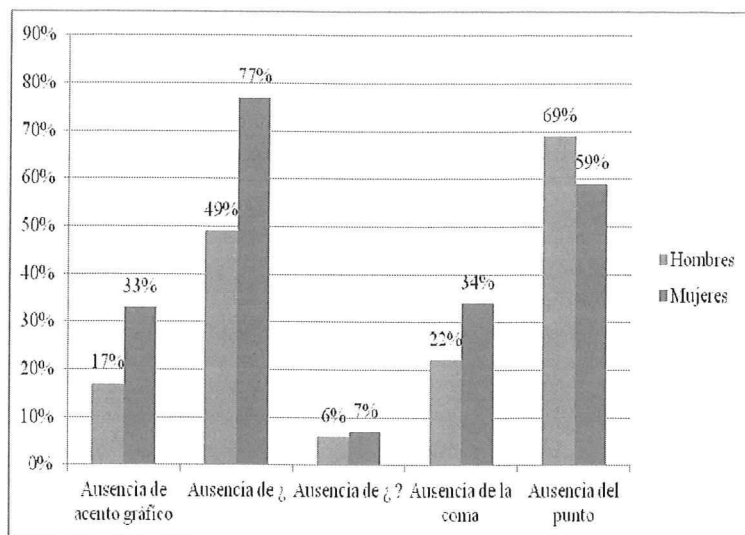
En el uso de abreviaturas y de estrategias grafo-estilísticas los porcentajes de los diferentes casos son muy similares en ambos casos. Sin embargo, sí es destacable que opciones técnicas de la propia plataforma como son los hashtags y las URL. son más utilizadas por hombres (en el 64% de los tuits) que por mujeres (33%).

En las opciones de interacción que ofrece Twitter los resultados muestran unos porcentajes muy similares.

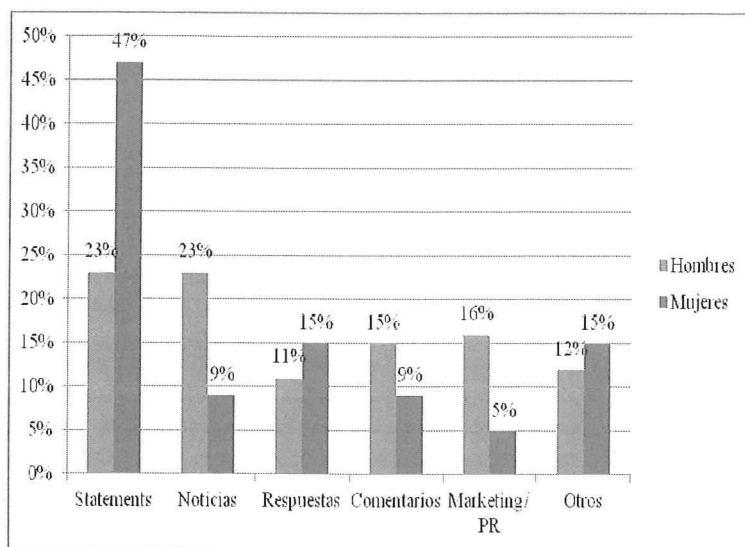
Donde existe una clara diferencia entre ambos sexos es en la función que cumplen los tuits en la comunicación. Las mujeres utilizan Twitter de forma mayoritaria (47% de todos los tuits) para publicar estados de ánimo, sobre lo que están haciendo o sobre cualquier otro motivo de índole personal. Los hombres también publican tuits para los anteriores fines, no obstante, en un porcentaje menor (23%) y publican con el mismo número de porcentaje de tuits dando o comentando noticias via hashtags o URLs.



Ortografía I



Ortografía II



Función

4. Conclusiones

Como hemos visto aquí, la comunicación en Twitter revela muchas características de la lengua de proximidad (KOCH/OESTERREICHER) y, en concreto, de su prototipo: el habla. No obstante, no estamos ante una nueva lengua o un nuevo lenguaje o código (= *netspeak*, *weblish*, *Cyberdeutsch*) como afirman muchos lingüistas (p. ej. Crystal 2001) cuando analizan la lengua usada en las nuevas forma de comunicación a través de medios digitales. El lenguaje utilizado en Twitter no llega a tener el grado de *proximidad* de una conversación hablada (comunicación face-to-face).

El lenguaje utilizado en nuestro corpus por las mujeres tiende a ser más *próximo* y menos normativo que el utilizado por los hombres.

Bibliografía

- ÁGEL, VILMOS/HENNIG, MATHILDE, „Theorie des Nähe- und Distanzsprechens“. En: ÁGEL, VILMOS/HENNIG, MATHILDE (EDS.), *Grammatik aus Nähe und Distanz*. Tübingen: Niemeyer, 2006, p. 3-31 .
- BROWN, GILLIAN/YULE, GEORGE, *Análisis del discurso*. Visor Libros: Madrid. DL 1993.
- CRYSTAL, DAVID, *Language and the Internet*. Cambridge: Cambridge University Press, 2001.
- FIEHLER, REINHARD/BARDEN, BIRGIT/ELSTERMANN, MECHTHILD/KRAFT, BARBARA, *Eigenschaften Gesprochener Sprache*. Tübingen: Gunter Narr Verlag, 2004.
- HENNIG, MATHILDE, *Grammatik der Gesprochenen Sprache in Theorie und Praxis*. Kassel: kassel university press, 2006.
- KOCH, PETER/OESTERREICHER, WULF, „Sprache der Nähe – Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte“. En: *Romanistisches Jahrbuch* 36, 1985, p. 15-43.
- REAL ACADEMIA ESPAÑOLA. *Nueva Gramática Española. Manual*. Madrid: Espasa, 2010.
- SIEBERG, BERND, „O modelo da ‘fala de proximidade e de distância’ e a sua aplicação à língua portuguesa“. En: FRANCO, MARIO/SIEBERG, BERND (EDS.), *Proximidade e Distância. Estudos sobre a Língua e a Cultura*. Lisboa: Universidade Católica Editora, 2011, p. 51-69.