Perpetuating gender stereotypes via the internet? an analysis of the women's presence in Spanish online newspapers

Mateos de Cabo, Ruth and Gimeno, Ricardo and Martínez, Miryam and López, Luis

Universidad CEU San Pablo, Banco de España, Universidad Rey Juan Carlos

2011

Perpetuating Gender Stereotypes via the Internet? An Analysis of the Women's Presence in

Spanish Online Newspapers

Ruth Mateos de Cabo, Ricardo Gimeno, Miryam Martínez, and Luis López

University CEU San Pablo

Author Note

Ruth Mateos de Cabo, Department of Management & Business Administration, University

CEU San Pablo; Miryam Martínez, Department of Management & Business Administration,

University CEU San Pablo; Ricardo Gimeno, Department of Quantitative Methods,

University Pontificia Comillas; Luis López, Department of Computing, University Rey Juan

Carlos.

Correspondence concerning this article should be addressed to Ruth Mateos de Cabo,

Department of Management & Business Administration, University CEU San Pablo, C/

Julián Romea, 23, 28003, Madrid, Spain. E-mail: matcab@ceu.es

Abstract:

In the present research we investigate possible predictors of the presence of women in Spanish online newspapers using an automatic content analysis over a three-month period. Results of the analysis reveal that Spanish online newspapers continue depicting women in a stereotyped manner. Women are still linked to traditionally 'female' sections, and they appear more frequently in news considered less important in terms of extension and publication day. The gender of the reporter also matters since female journalists tend to include more women in the news they report than their male peers.

Keywords: newspapers, gender representation, stereotypes, Internet, content analysis.

Perpetuating Gender Stereotypes via the Internet? An Analysis of the Women's Presence in

Spanish Online Newspapers

The media tend to depict women and members of minorities in a stereotyped manner

(Swert & Hooghe, 2010). Research indicates that representation of women in the news

remains heavily biased. In this sense, the Global Media Monitoring Project for the 2010

shows that although women all over the world have gradually acquired more leading

positions in the political, cultural and scientific domains since 1995 (the year when the first

report was launched), they still are very heavily underrepresented in the news that, besides,

depict a similar gender representation as it did decades ago (Gallagher, 2010). Craft and

Wanta (2004) found that women are depicted in the news in a strongly stereotyped way, and

limited to traditional 'female' topics, while they are largely absent in topics like law, finance,

economy or foreign affairs.

Armstrong (2004) argued that by presenting a reflection of society where women are

underrepresented or misrepresented, newspapers can affect gender perceptions and reinforce

a lower public status for women. Nevertheless, media are not only in the origin of the

problem but they are part of the solution too, since, given their influence in Western societies,

they might be a suitable vehicle for social change, that could help to eliminate stereotypes.

This is the reason why the study of gender representation in the media is relevant, since

determining possible causes of disparities in gender portrayals will contribute to offer a more

accurate reflection of the society to readers and, as a consequence, to reduce stereotypes and

wrong perceptions regarding the lack of newsworthy of women for public events.

Accordingly, we decided to study the coverage of women in the news stories, since

they are the major source of information, facts, ideas and opinion, and they are also a critical

information vehicle when it comes to gender social roles (Lobo & Cabecinhas, 2011). From

all possible media, we choose the Internet since it presents a series of characteristics that

make it especially interesting for its study: It is a meta-media; It inherits forms of previous media, adding exclusive characteristics (i.e. fast evolution, relevance, free of geographical barriers, continuous presence); Its consumption is already higher in many countries than that of television and radio. Besides, since its advent has changed the collection, distribution, and editing of media content, along with access and consumption (Schultz & Sheffer, 2007), we considered an interesting issue to examine if it has resulted in any gender-related changes in media content. In this sense, Van Zoonen (2003) argued that this new medium of journalism may play a fundamental role in transmitting feminine not stereotyped images. Besides, the Internet allows automating content analysis reducing its cost considerably in terms of time, money and failures.

Therefore, the primary purpose of this study is to check if Spanish online newspapers reflect the great advancement in gender diversity that has experienced the Spanish society or if, on the contrary, they continue reinforcing masculine stereotypes. In order to address this, we looked for possible predictors of women's presence in Spanish web news, including section, journalist's gender and news' relevance (extension and publication day). This was accomplished by conducting a content analysis of all news by four of the main Spanish online newspapers[1] over a three-month period (March-May 2006).

This study is the first, to the best of our knowledge, to accomplish a content analysis of the news on the Internet with an automatic procedure. Given the possibilities that offered us a media such as the Internet, we implemented diverse computer programs. In particular, for the extraction of the news we implemented automatic robots, known as *crawlers* that are able to navigate on the Internet in a similar way to a conventional browser. Additionally, to analyze the text of the news items we implemented *word counting algorithms*, that identify the presence of forenames classifying them by gender (we also used it to identify the gender of the journalist). Besides, unlike most of previous studies, we conducted multivariate

analyses through the estimation of two different models of qualitative dependent variable[2] (Probit and Poisson regression models), in order to isolate the individual effect of each one of the different variables considered.

Women in the News

Spain has experienced important political, economic and social changes in the last 20 years. The proportion of women in the workplace had risen from 28.8% in 1986 to 47.5% in 2006 (Valls-Fernández & Martínez-Vicente, 2007). According to the World Economic Forum 2009 Gender Gap report (Hausmann, Tyson, & Zahidi, 2009), Spain ranks 9th out of 134 countries regarding political empowerment (6th for women in ministerial positions with 44% female representation, and 11th in women in parliament with 36% of female members), and in terms of education it occupies first position for level of enrolment in secondary and tertiary education (96% and 76% of all women respectively).

While gender diversity has been gaining ground very rapidly in Spanish society, gender balance of the media is changing rather slowly. In fact, if we review gender media representation research in Spain, we discover that women are practically invisible in the world that the Spanish media reflect. Fagoaga and Secanella (1984) found a women's representation on the five newspapers analyzed of just a 7 %. Gallego (1998) showed that although the percentage of women's mention in written press, had increased by late nineties, were not reaching 12% of the total. Franquet et al. (2007) in their analysis of eight electronic media in 2003, found a percentage of women's mentions of just 17%, and that men were mentioned by their name almost two times more than females were.

The situation is similar worldwide. Zoch and Turk (1998) found males were quoted more than four times as often as females in international stories of three southern U.S. daily newspapers from 1986 to 1996, and, in overall, coverage seven in ten sources were males. Cann and Mohr (2001), in their content analysis of news stories broadcasted by five

Australian television networks, found that male outnumbered woman by more than 6 to 1, and that they were the majority in every area except female sport, while in some areas (disaster, technology, and male sport) female experts were entirely absent. Armstrong (2004) found that the overall frequency of male mentions was nearly three times that of female mentions within news stories from 18 U.S. newspapers. The analysis of television news programs by Desmond and Danilewicz (2010) revealed that males predominated as experts sources. Swert and Hooghe (2010) using a dataset of news from Belgium broadcasts for the years 2003–5 found that only in just a 30% of the news item covering 'male' topics (institutional politics, defense, international relations, finance, economy and justice) could be found at least one on-screen female news source. Finally, Gallagher (2010) showed that in 2010 women were just 24% of the news subject (contrasting with the 17% and 18% for 1995 and 2005), and that the voices heard and the faces seen in the news media remind dominated by men.

Hypothesis

Despite the fact that throughout society gender roles had changed dramatically during the 1960s and 1970s, the media still employ stereotyped gender images which are gradually becoming outdated (Carter & Steiner, 2004; Fraser, 1995; Van Zoonen, 1994). According to Lens-Ríos, Rodgers, Thorson, and Yoon (2005), one explanation for the underrepresentation of women is that, when journalist cover stories deemed newsworthy, they reproduce hegemonic cultural norms privileging men. In this study, we focus our attention in three main determinants for the underrepresentation of women in the news stories.

There seems to be a segregation according to which woman are traditionally associated with sections labeled traditionally as 'female' sections such as health or family matters, education, entertainment, culture or social policy (Craft & Wanta, 2004; Lens-Ríos et al., 2005; Rodgers & Thorson, 2003; Van Zoonen, 1998). Research by Kahn (1994)

indicated that discussion of 'female' issues (childcare, education, women's rights and health care) is more elaborate in stories with women protagonists, while men portrayed in newspapers were more likely to be linked with 'male' issues, such as foreign policy, defense spending, and economic concerns. Lens-Ríos et al. (2005) found that female appear proportionally less often than men in stereotyped male sections of the newspapers (business and sports) and more often in female sections such as entertainment. The authors suggest that socialization and journalist education may contribute to differences and stereotyped new coverage. The marginalization of women in sport news has been widely documented (Eastman & Billings, 2000; Pedersen, 2003; Rodgers & Thorson, 2000). Given this fact, we can hypothesize that women will be mentioned more often in stereotypically female sections of the newspaper. Therefore, our first hypothesis is:

H-1: Women are more frequently mentioned in stereotypically female sections (people, society and culture sections in our study).

Another important factor that can be explaining the presence of female in the news is the gender of the reporter. The assumption behind the existence of possible gender differences in reporting has to do with the fact that female reporters bring with them a different value pattern, because of their gender socialization, resulting in more diverse sources in the news and topics receiving attention (Price & Wulff, 2005; Rodgers & Thorson, 2003). This approach corresponds to the gender model of socialization that contends that men and women have different values and priorities because they socialize differently (Dodd-McCue & Wright, 1996). However, Weaver and Wilhoit (1992) did not find any evidence of female reporters having different professional values than their male colleagues or that they preferred different topics. This can be explained using the job model according to which women are assumed to perform the same as men to the extent that their organizational experiences are similar (Aven, Parker, & McEvoy, 1993). In this sense, as female and male

reporters will be socialized into the predominant value pattern that is present in the newsroom they will behave very much the same in terms of sourcing and framing the news (Rodgers & Thorson, 2003; Ross, 2007). Lavie and Lehman (2003) suggested that lack of gender differences may mean that women entering journalism self-select themselves for the profession and so have similarities with their male colleagues.

Swert and Hooghe (2010) point out that empirical studies offer inconclusive results in the question of whether the presence of female reporters actually makes a difference. While some studies see clear differences between female and male reporters on their selection of news sources (Armstrong, 2004; Cann & Mohr, 2001; Rodgers &Thorson, 2003; Sutcliffe, Lee, & Soderlund, 2005; Zoch & Turk, 1998), and prominences given to those sources (Liebler & Smith, 1997); Other studies do not find that female journalist differ from their male colleges in their operating practices (Creedon, 1993; Gallagher, 2006; Lavie & Lehman, 2003; Mills, 1997). Given this inconclusive results, and in order to find which model prevails in the Spanish case (the gender model or the job model) we formulate the following hypothesis:

H-2: Women reporters tend to include in the news they sign more women than their male colleagues.

To have into account the fact that female reporters are routinely assigned to women's issues and 'softer' news stories (Desmond & Danilewicz, 2010; Liebler & Smith, 1997) which could have lead to more selection of female sources, we control by type of section.

Our final hypothesis has to do with the fact that women tend to be sidelined to some extent in the news. In fact, studies on newspapers and television news show that when women are interviewed the average amount of time or space they are allotted is significantly smaller than for male interviewees (Van Zoonen, 1994). In this sense, in a content analysis, Zoch and Turk (1998) found that 69.2% of front-page sources were men, compared to 17.3%

women and when they did appear they were quoted less often and given shorter quotes. Armstrong (2005) in her study found that male sources were placed more prominently (in the headline or first paragraph) in the stories than females, and that they were more likely to appear on the front-page of the newspapers. Additionally, Cann and Mohr (2001) found that male-reported stories tended to be placed higher in the news cast and were slightly longer in duration. Story placement is a significant issue, because the first few stories of a newscast parallel the front page of a daily newspaper, where stories are placed in descending order according to their perceived importance. Swert and Hooghe (2010) also hypothesized that the news items on which women figure will be deemed as less important, although they didn't find empirical confirmation of it.

As we focus our study on the Internet press, and given that in newspapers more important stories are generally longer than less important ones (Zoch & Turk, 1998), we state the following hypothesis:

H-3: Women are more frequently mentioned in smaller news in terms of extension.

Additionally, Fagoaga and Secanella (1987) indicated that women's news are considered less relevant, and left out of the working days, going on to the periphery (weekends and especially Sundays) as 'feature' news. To take into account the importance of the news in terms of publication time we posit this final hypothesis:

H-4: Women tend to appear more often in news published in days of the week where the informative flows decrease such as Sundays.

Method

Sample selection

We conducted a content analysis[3] of four of the main Spanish generalist online newspapers: elmundo.es[4], elpais.es, abc.es, and libertaddigital.es, according with unique visitor's information provided by Nielsen NetRatings (2006) (as shown in table 1), over a

three-month period (March-May, 2006). *Terra Actualidad* (third on the list) was excluded

from the study because it does not have a news library to consult the historical files[5].

All the studied digital newspapers organized its contents in sections and services

offered to the readers, except for abc.es that also included some channels. *El Mundo* was the

one that accumulated the higher number of sections, whereas, *ABC* was the one containing

the higher number of services. *Libertad Digital* was the one with less sections and services

(Table 2).

Regarding the different sections covered by the online newspapers analyzed, both

ABC and *El País* included the following sections: science, culture, sports, economics, people,

international relations, national and society, whereas *El Mundo* did not file in its newspaper

library the economics and the sport sections, and *Libertad Digital* did not file the culture and

people sections. Although science and people sections have different names depending on the

newspaper, we grouped them according to their content. Table 3 shows the distribution of

recovered news items by sections and newspapers.

Data extraction

For each of the specified media, we extracted on a daily basis all news items from the

above mentioned sections. For each item, we recovered the following information: headline,

reporter, full text, date, newspaper and section. Additionally, we proceeded to analyze the

news item to obtain the following information: length of the news item (in number of words),

number of male and female first names in the body of the news item, gender of the reporter,

and the presence of a male or female first name in the headline.

In order to make the analysis manageable we decided to collect and analyze data

automatically, through a web crawler[6] that behaves like a human user using a web browser

(Eliassi-Rad & Shavlik, 2003; Kushmerick, 2000a, 2000b). This software enabled us to

connect to the online media, recover the text associated to each item, analyze it and store the results into a relational database in an appropriate way to be later processed.

Then, from this database, we accomplished an automatic identification of male and female first names, through a text analysis tool. In order to do this, we implemented a word counting algorithm based on two dictionaries, one of 5,900 female names and another one containing 8,000 male names, that included most usual first names from major languages (i.e., Spanish, English, French, German, Italian, Arabic, Chinese and Japanese). The word counting algorithm automatically compared each word on the text with the dictionaries to detect if it were a male or female name or none of them[7]. From this search, we identified more than 175.000 names. The same methodology was used to identify the gender of the journalist.

As a reliability check, we compared the results obtained from this tool with those obtained from human experts in a random sub-sample of 100 news items. The results are presented in Table 4 and reveal the high fidelity and accuracy of the automatic word counting developed for this study. This allowed us to conclude that the automatic count of the women mentioned in the news provides a data set reliable enough to perform a statistical analysis.

Variable Definition

The dependent variable is the proportion of female first names among total first names mentioned in each news item.

To account for the effect of the section as an independent variable (Hypothesis 1), we created dummy variables for the following sections: science, culture, sports, economics, people, international relations, and society, being national section the one used as reference. Sections traditionally considered as feminine are culture, people (celebrity and royalty) and society (Bueno Abad, 1996). Traditionally male sections are sport and economics, whereas we considered neutral the remaining ones (science, international and national sections).

The gender of the reporter is also a relevant variable according to Hypothesis 2. To measure this variable we introduced two dummy variables: news item bylined by a female journalist, and news item bylined by a mixed female/male team, being the category of reference the case where the reporter is a man or where the author is not identified.

Hypothesis 3 and Hypothesis 4 state that news items in which women are mentioned are considered less important both in terms of length and time than those in which men appear. We took this into account by including two different variables: the extension of the article (number of words) to measure length and the day of publication to measure time, using Monday as the reference day.

We also controlled by women's presence in the headline of the news item, number of mentioned first names, month of publication (in order to avoid distortions produced by a special hot issue of a specific month), the percentage of women on the staff, the cost per employee and the number of employees of each online newspaper.

A clear indicator of the women's presence in the news is the fact that they appear in the headline. Often, when women are mentioned in the headline, it is because they play a leadership role of have some institutional position, turning, this way, into spokespersons for their organizations. Therefore, it is expected that the variable woman mentioned in the headline will be highly significant when explaining their mention in the text of the news items.

Finally, the way media shows social reality is the result of the interaction of institutional forces: owners, control of the publishing structure, professionals' or organization's ideology (Fagoaga & Secanella, 1987). In order to consider the heterogeneity of the different online newspapers analyzed, we included specific variables for each one, in particular: percentage of women in the staff[8], number of personnel, as well as cost for employee[9].

Data analysis and results

Our analysis covered a period of three months, widely overcoming the twelve day size recommended by Stempel (1952) and the two weeks recommended for local stories by Riffe, Aust, & Lacy (1993). This sample size allowed avoiding any possible bias in the flows of news that could prevail in shorter periods of time and to get a sufficient sample to yield reliable estimates of all type of stories. In this sample only 31.677 names, out of a total of 175.965 identified names via the automatic check up, corresponded to women, which account for just an 18%.

Differences in the attention received by women in the news depend on a great number of elements (Shoemaker & Reese, 1996). Nevertheless, many of these elements are closely related (i.e., women journalists report more often on female sections as society or people), preventing that a descriptive analysis could provide reliable conclusions. Therefore, we proceed by specifying econometric models that allow isolating the influence of each explanatory variable in *caeteris paribus* conditions (i.e., controlling for the rest of variables entered in the model).

Since the gender of the name of the persons mentioned in a news can be consider a dummy variable that takes the value 0 (man) or 1 (woman), we used two different models of discrete dependent variable to explain the presence of women on the news of the web sites analyzed. This way, the proportion of women mentioned in a given news item might be modeled by a binomial variable, where p_i is the probability of each person mentioned in the news item to be a woman, that can be determined by the variables previously indicated according with a *grouped Probit* model:

$$p_i = \Phi(\mathbf{X}_i\boldsymbol{\beta}) \tag{1}$$

Where $\mathbf{X}_i$ is a vector of independent variables of each news item influencing the above mentioned probability, $\boldsymbol{\beta}$ is the vector of coefficients to estimate, and Φ is a standard normal distribution.

Table 5 shows the estimation results obtained by maximum likelihood[10]. The first model includes all the variables. In order to obtain the second model, the variables with higher p value in a LR test[11] were eliminated one by one until only significant variables remained. According to both Wald and LR tests, both models are able to explain the proportion of women mentioned in the news.

Alternatively, we the dependent variable can be modeled by a *Poisson* regression model. In this model, each observation is the outcome of a random variable with *Poisson* distribution with parameter λ_i,

$$\Pr[Y = y_i] = \frac{(\lambda_i)^{y_i}}{y_i!} e^{-\lambda_i} \qquad y_i = 0,1,2,\ldots \qquad (2)$$

Being y_i, the number of women mentioned in the news item, and λ_i the expected number of women mentioned in the news that would depend on $\mathbf{X}_i$, according with the following expression,

$$\lambda_i = n_i \cdot e^{\mathbf{X}_i \boldsymbol{\beta}} \qquad (3)$$

Where, n_i is the number of people mentioned in the news item.

Results of these models (see table 6) are close to those obtained for the grouped Probit model (see table 5). The estimated models show that there are three clear determinants of the presence of women on the news: section, gender of the reporter, and the importance of the news item in terms of length and publication day.

The hypothesis that women are being mentioned much more often in stereotypically female sections seems to be confirmed. As can be seen in Figure 1, there is a greater presence

of women[12] in the sections of people (33%), society (28%) and culture (24%), by this order.

By contrast, in national, economics and international sections (18%), there are no significant

differences. This way, women are overwhelmingly depicted in items linked to leisure

activities, to aspects of intervention and social commitment, or as professional of the world of

culture and art. It is also important to emphasize that the probability of being mentioned

diminishes drastically when the news appears in the section of sports (about 7%). This

unequal distribution by sections shows that this form of segregation and stereotyping remains

in the online press, determining that women continue being more frequently associated with

sections concerning feminine and soft topics.

We also found evidence that female reporters tend to mention more women in the

news they report when compared to their male counterparts (Hypothesis 2), even after

controlling by section (Figure 2). This is also true when the news item is reported by a mixed

female/male team. This way, the probability for a person mentioned in a news item to be a

woman increases around 4%-5% when it is bylined by a woman. This impact is greater than

the one reported by Swert and Hooghe (2010). In this sense, the gender model seems here to

be empirically supported, suggesting that female reporters offer a more plural and diverse

approach to the news in terms of gender representation possibly as the result of women´s

socialization since childhood.

The fact that women are sidelined in the news of more importance, both in terms of

length and publication day is confirmed (Hypothesis 2 and 3 respectively). This way, the

news items in which women are mentioned have a lower profile in terms of length (see

Figure 3). For instance, the probability that the name mentioned belongs to a woman in the

news of 100 words is 1.5% higher than one in a news item of median extension (400 words).

Besides, in both models (Figure 4) Sunday turns out to be a significant variable. This

result corroborates the hypothesis that women are traditionally linked to the least important

news that are usually relegated to Sunday editions, when the informative flows diminish (Sundays are the day of the week where less news items has been obtained) and the 'hot issues' are substituted with topics or general events of lesser importance.

Regarding the control variables, it is also observed that the fact that a woman appears in the headline of the news, exercises a considerable effect on the probability that the names that appear in the body of the news are women. Nevertheless, the number of occasions in which the woman appears as protagonist of the news is very low (5%). The number of names mentioned also has a negative impact. In this sense, the more number of persons are mentioned in the news items the lower is the proportion of women.

In relation to specific variables of each media, the presence of women in positions of responsibility of the media influences positively the female presence in the news. The cost per employee has also a positive impact on the women's presence on the news, whereas the number of employees has a negative influence. Nevertheless, since only four media have been analyzed, we have to be cautious with these results.

Finally, we observed differences per months, which reflect that punctual events that can take place in certain months can contaminate the results. This way, women's proportion is higher in March than in April and May, which could have been influenced by events as the death of two famous Spanish singers (Rocío Durcal and Rocío Jurado), two events that were an object of a wide informative coverage in multiple sections.

Summary and Discussion

The analysis of the women's presence in the Internet newspapers constitutes a key element to confirm if online press continues excluding and demeaning the value of women in Spanish society, perpetuating in this way stereotypes and gender inequality or, on the contrary, it takes advantage of the communal nature of the Internet to favor the public appearance of women previously marginalized.

In order to test this, we accomplished a content analysis of the news items filed in the libraries of four of the main Spanish online newspapers from March to May 2006, finding that female names represent only an 18% out of the total. One possible explanation for this acute underrepresentation is the one offered by Kanervo and Kanervo (1995), who argue that women do not hold positions that generally serve as spokespersons for the media. However, since in Spanish society gender diversity has been gaining ground very rapidly, additional causes must be explored. In fact, in this research we identified some important predictors of the underrepresentation of women in online newspapers.

It is confirmed that the higher women's representation corresponds to stereotypically female sections such as people, society and culture. It seems that in online newspapers the problem of stereotyping is perpetuated. This unequal gender distribution for sections supposes the existence of a gender segregation that will condition that women's news continue being more frequent in the sections associated to 'soft' topics.

It turns out also that female journalists tend to include more women in the news items they bylined than their male counterparts even after controlling by the section. We also observed this effect when the news items were reported by a mixed female/male team. This result seems to confirm the gender model approach that posits that female journalist, as a result of their different socialization, draw upon a wider variety of viewpoints than men when reporting stories.

Finally, it is also supported that women appear more frequently in news considered less important in terms of length and publication day. This way, Sunday is the day of the week where we found a higher women's proportion in the news. This point out that the news in which women appear are considered suitable to reflect topics or general events of lower relevance that therefore are displaced to days, like Sundays, when the informative flows

decrease. News items in which women are mentioned have a lower profile in terms of length, too. In this sense, it is verified that the news in which women appear are more 'feature' news.

Given these results, we can argue that by portraying women linked to feminine sections, to less important news items and more frequently mentioned by female reporters, Spanish online newspapers continue to be a factor that favors the formation of stereotypes and gender segregation as conventional newspapers do. If we want that the online newspapers reflect the progressive incorporation of the woman to the public arena that has been taking place in Spanish society during the last three decades, it would be desirable to take advantage of the opportunities that offers a more innovative and interactive media as the Internet to represent a more plural and diverse reality. The loss of the potential of this media to transform and transmit new journalistic practices to the conventional media and to change the way readers perceive the role of women in society is a genuine detriment to journalism, readers and the whole society.

Finally, based on the results, we offer some recommendations for actions that would improve the amount of coverage that women receive in the Spanish online newspapers. So, one way to reduce gender stereotyping in news content, would be the inclusion in traditionally 'male' sections of expert women in order to achieve more gender-balanced news. It would be also desirable an increase in the number of female reporters on the news desk, and the number of mixed teams in terms of gender. Besides, female journalist could have a more active role in the creation and diffusion of information on the Internet, for example, through the use of forums, blogs, chat rooms and message boards. Finally, the use of effective gender-equity policies in the newsroom could help to make women more visible not only in feature news but also in specific events.

Limitations and Recommendations for Future Research

The results of this study are limited by its scope: It involved only news in four of the main digital newspapers of national scope during March, April and May, 2006. Obviously, it is the need to expand the study to other online newspapers. A limitation derived from the automation of the content analysis is the definition of the dependent variable. It would be desirable to accomplish an automatic identification of persons mentioned not only through first names but also by mean of other identifying attributes (i.e. pronouns like he/she, titles like Mr/Mrs). Nevertheless, this would imply the use of much more complex computer programs that are out of the scope of this research. Future research could also include semantic analysis, in order to determine the role that play the women mentioned in the news.

Finally, we consider that it is vital an ongoing monitoring of gender content in the news on the Internet that allow tracking the progress in this field and a research that continue to interrogate practices that minimize women in news.

References

Armstrong, C. (2004). The Influence of Reporter Gender on Source Selection in Newspaper Stories. *Journalism and Mass Communication Quarterly*, 81(1), 139–54.

Aven, F.F, Parker, B. & McEvoy, G.M. (1993). Gender and attitudinal commitment to organizations: A meta-analysis. *Journal of Business Research*, 26, 63-73.

Bueno Abad, J.R. (1996). *Estudio longitudinal de la presencia de la mujer en los medios de comunicación de prensa escrita*. Valencia: Nau Llibres.

Cann, D.J. & Mohr, P.B. (2001). Journalist and source gender in Australian television news. *Journal of Broadcasting & Electronic Media*, 45, 162–185.

Carter, C. & Steiner, L. (2004). Mapping the Contested Terrain of Media and Gender Research. In Carter C. & Steiner L. (eds.) *Critical Readings: Media and Gender* (pp. 11–36). Maidenhead: Open University Press.

Craft, S. & Wanta W. (2004). Women in the Newsroom: Influences of Female Editors and Reporters on the News Agenda. *Journalism and Mass Communication Quarterly*, 81(1), 124–38.

Creedon, P.J.(ed.) (1993). Women in Mass Communication (2nd ed.). Newbury Park, CA: Sage.

Desmond, R. & Danilewicz, A. (2010). Women are on, but not in, the news: Gender roles in local television news. *Sex Roles*, 62, 822-829.

Dodd-McCue, D., & Wright, G.B. (1996). Men, women, and attitudinal commitment: The effects of workplace experiences and socialization. *Human Relations,* 49, 1065-1091.

Eastman, S.T., & Billings, A.C. (2000). Sportscasting and sports reporting: The power of gender bias. *Journal of Sport and Social Issues*, 24, 192–213.

Eliassi-Rad, T. & Shavlik, J. (2003). A System for Building Intelligent Agents that Learn to Retreieve and Extract Information. *User Modeling and User-Adapted Interaction*, 13(1-2), 35-88.

Fagoaga, C. & Secanella, P. (1984). *Umbral de presencia de las mujeres en la prensa española*. Madrid: Serie Estudios.

Franquet, R., Luzón V., & Ramajo N. (2007). News in the Principal On-line Mass Media. Studying Gender Representation. *Zer. Revista de Estudios de Comunicación*, 12(22), 267-282.

Fraser, N. (1995). Pragmatism, Feminism, and the Linguistic Turn. In S. Benhabib, J. Butler, D. Cornell & N. Fraser (Eds.), *Feminist Contentions* (pp. 157–171). New York: Routledge.

Gallagher, M. (2006). *Who Makes the News?* London: World Association for Christian Communication.

Gallagher, M. (2010). *Who makes the news? Global Media Monitoring Project 2010*. Toronto: Word Association for Christian Communication.

Gallego, J. (1998). *Génere i informació*. Barcelona: Associació de Dones Periodistas.

Gunter, B. (2000). *Media research methods*. London: Sage.

Hausmann, R., Tyson, D. L. & Zahidi, S. (2009). *The global gender gap report*. Geneva: Word Economic Forum.

Kahn, K.F. (1994). The Distorted Mirror: Press Coverage of Women Candidates for Statewide Office. *Journal of Politics*, 56, 154-173.

Kanervo, D. & Kanvervo, E. (1995). Small Town CEO's Views on Newspaper Coverage. *Newspaper Research Journal*, 16, 29-41.

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). Thousand Oaks, CA: Sage.

Kushmerick, N. (2000a). Wrapper induction: Efficiency and expressiveness. *Artificial Intelligence*, 118(1-2), 15-68.

Kushmerick, N. (2000b). Wrapper verification. *World Wide Web Journal*, 3(2), 79-94.

Lavie, A., & Lehman-Wilzig, S. (2003). Whose news? Does gender determine the editorial product? *European Journal of Communication*, 18, 5–29.

Len-Ríos, M.E., Rodgers, S., Thorson, E. & Yoon D. (2005). Representation of Women in News and Photos: Comparing Content to Perceptions. *Journal of Communication*, 55(1), 152–168.

Liebler, C. & Smith S. (1997). Tracking Gender Differences: A Comparative Analysis of Network Correspondents and their Sources. *Journal of Broadcasting & Electronic Media*, 41(1), 58–68.

Lobel, T.E., Rothman G., Abramovizt E., & Ziva M. (1999). Self-Perception and Deceptive Behavior: The Uniqueness of Feminine Males. *Sex Roles*, 41(7–8), 577–587.

Lobo, P. & Cabecinhas, R. (2011). The negotiation of meanings in the evening news: Towards an understanding of gender disadvantages in the access to the public debate. *International Communication Gazette*, 72, 339-358.

Mills, K. (1997). What Difference do Women Journalists Make? In P. Norris (Ed.) *Women, Media and Politics* (pp. 41–55). Oxford: Oxford University Press.

NIELSEN NETRATING (2006). *Panel De Hogar*. Data nov.05-apr.06.

Pedersen, P.M. (2003). Examining stereotypical written and photographic reporting on the sports page: An analysis of newspaper coverage of interscholastic athletics. *Women in Sport & Physical Activity Journal*, 12, 67–75.

Price, C. & Wulff S. (2005). Does Sex Make a Difference? Job Satisfaction of Television Network News Correspondents. *Women's Studies in Communication*, 28(2), 207–234.

Riffe, D., Aust, C.F., & Lacy, S.R. (1993). The effectiveness of random, consecutive day and constructed week sampling in newspaper content analysis. *Journalism Quarterly*, 70, 133–139.

Rodgers, S., & Thorson, E. (2000). "Fixing" stereotypes in news photos: A synergistic approach with the Los Angeles Times. *Visual Communication Quarterly*, 55(7), 8–11.

Rodgers, S. & Thorson, E. (2003). A Socialization Perspective on Male and Female Reporting. *International Journal of Communication*, 53(4), 658–675.

Ross, K. (2007). The Journalist, the Housewife, the Citizen and the Press: Women and Men as Sources in Local News Narratives. *Journalism*, 8(4), 449–473.

Royo-Vela, M., Aldas-Manzano, J., Küster, I. & Vila, N. (2008). Adaptation of marketing

    Activities to Cultural and Social Context: Gender Role Portrayals and Sexism in

    Spanish commercials. *Sex Roles*, 58, 379-390.

Schultz, B., & Sheffer, M. L. (2007). Sports journalists who blog cling to traditional values.

    Newspaper Research Journal, 28(4), 62–76.

Shoemaker, P. & Reese S. (1996). *Mediating the Message: Theories of Influence on Mass

    Media Content* (2nd ed.). New York: Longman.

Spears, G., Seydegart, K. & Gallagher, M. (2000). *Who Makes the News? The Global Media

    Monitoring Project 2000*. London: Word Association for Christian Communication. .

Sutcliffe, J., Lee, M., & Soderlund, W. (2005). Women and crisis reporting: Coverage of

    television news coverage of political crises in the Caribbean. *Harvard International

    Journal of Press/Politics,* 10, 99–124.

Swert, de K. & Hooghe, M. (2010). When do women get a voice? Explaining the presence of

    female news sources in Belgian news broadcast (2003-5). *European Journal of

    Communication*, 25(1), 69-84.

Valls-Fernández, F. & Martínez-Vicente, J.M. (2007). Gender Stereotypes in Spanish

    Television Commercials. *Sex Roles*, 56, 691-699.

Van Zoonen, L. (1994). *Feminist Media Studies*. London: Sage.

Van Zoonen, L. (2003). Gendering the Internet: Claims, controversies and cultures. *European

    Journal of Communication*, 17(3), 5–23.

Weaver, D. & Cleveland W.G. (1992). *The American Journalist: A Portrait of US News People and their Works*. Bloomington: Indiana University Press.

Zoch, L.M. & Turk J.V. (1998). Women Making News: Gender as a Variable in Source Selection and Use. *Journalism and Mass Communication Quarterly*, 75(4), 762–75.

Footnotes

[1] abc.es, elmundo.es, elpais.es and libertaddigital.es

[2] These estimations allowed us to quantify the effect of different explanatory variables. For example, some studies find positive relations between female reporters and the mentions to women in the news they cover; at the same time these mentions are more frequent in sections of society (Franquet, Luzón & Ramajo, 2007). It might be argued, therefore, that the positive relation found is due to the fact that women journalists are more frequently assigned to traditionally female sections (such as society). For this reason it is not simple to identify the individual effect of each one of these factors.

[3] Content analysis is an unobtrusive or non-reactive method used by social scientists that has been applied to nearly every form of communication, such as newspapers, television and radio broadcasts, speeches or literature (Gunter, 2000; Krippendorff, 2004).

[4] elmundo.es has three editions (morning, evening and night), but we have worked always with the morning edition.

[5] We have worked with historical files because we need to guarantee the continuity in the extraction of news each day.

[6] We built a specific crawler for each media, based on the XULRunner API, whereas an additional piece of software and the behavior of the crawler were defined using Java Standard Edition and XPath technology respectively.

[7] Most of the detected names were Spanish names. Spanish first names clearly identify gender; even for those names with the same root they have a different suffix (i.e. *Antonio* for male, *Antonia* for female). In the case of composite names, the order signals the gender (i.e. *María-José* for female and *José-María* for male). Therefore, in case of consecutive names, we considered them a single person, with gender identified by the first name.

[8] Information obtained from the Presidency of Government webpage:

http://www.la-moncloa.es/ServiciosdePrensa/Agendadelacomunicacion. In the case of Libertaddigital.com information comes directly from their webpage.

[9] Information obtained from SABI (*Informa* and *Bureau Van Dijk*. http://www.bvdep.com/SABI.html)

[10] A robust variance-covariance matrix is used in order to correct for heterokedasticity, as well as correlation between persons cited in the same news.

[11] Likelihood ratio test is reported instead of the Wald test, because on these models the variance of the estimated parameter would be higher, the higher is the beta estimated value. This causes that Wald statistics is biased and tends to accept the null hypothesis even when it should be rejected.

[12] These effects are computed as departures from a base case of a news with 400 words, where 5 persons are cited, appeared on Monday, of March in the National section, whose author it is not identified as a woman, published by a media with 19 employees, with a cost per employee of 56000 € and a 10% of female staff.

Table 1

Most popular Spanish generalist Digital Newspapers (average value November05-April06)

Ranking	Newspaper	Number of unique visitors
1	El Mundo	2572
2	El País	1664
3	Terra Actualidad	1385
4	ABC	595
5	Libertad Digital	548
6	20 Minutos	535
7	Cadena SER	519
8	Periodista Digital	358
9	La Verdad	346
10	El Periódico	335
11	La Vanguardia	334
12	El Correo	322
13	La Razón	261
14	La Voz de Galicia	246
15	MSN News	234

From Nielsen Netratings, 2006.

Table 2

Number of sections, services and channels offered by digital newspapers

	Sections	Services	Channels
ABC	11	23	17
El Mundo	23	14	
El País	14	19	
Libertad Digital	10	10	

Table 3

Number of news items analyzed by sections and newspapers.

	National	International	Society	Culture	Economics	Sport	Science	People
ABC	3356	1444	2092	1350	1514	4396	167	1003
El Mundo	740	781	772	754	--	--	918	976
El País	1264	1194	1004	907	1101	1071	1648	815
Libertad Digital	748	813	831	--	828	1073	693	--

Table 4

Automatic vs. Manual identification of mentioned people

	#Men	#Women	%men	%women
Automatic	428	132	76.4%	23.6%
Manual	444	140	76.0%	24.0%

Table 5

Grouped Probit model on the probability of being a woman the person mentioned in the news

item

Variable	Model I (All variables)			Model II (Only significant variables)		
	β	LR	p-value	β	LR	p-value
Constant	-1.683			-1.670		
Number of mentioned persons	-0.002	12.58***	0.000	-0.002	12.31***	0.000
Number of words in the news (log)	-0.018	6.12**	0.013	-0.018	6.39**	0.011
Woman mentioned in the title	0.834	4809.92***	0.000	0.833	4819.30***	0.000
News bylined by						
a woman	0.159	82.95***	0.000	0.159	83.56***	0.000
a mixed female/male team	0.187	15.06***	0.000	0.186	14.97***	0.000
Day of the week						
Tuesday	0.002	0.03	0.872			
Wednesday	0.021	2.42	0.119			
Thursday	0.021	2.32	0.127			
Friday	0.023	2.76*	0.097			
Saturday	0.018	1.70	0.193			
Sunday	0.040	8.27***	0.004	0.026	5.71**	0.017
Month						
April	-0.082	82.11***	0.000	-0.083	84.09***	0.000
May	-0.056	39.97***	0.000	-0.058	42.03***	0.000
Section						
Science	0.081	18.35***	0.000	0.082	21.48***	0.000
Culture	0.207	265.84***	0.000	0.208	337.77***	0.000
Sports	-0.555	1721.57***	0.000	-0.556	2234.78***	0.000
Economics	-0.003	0.02	0.877			
People	0.462	941.32***	0.000	0.463	1119.17***	0.000
International	-0.003	0.04	0.832			
Society	0.320	596.80***	0.000	0.321	752.35***	0.000
Media related variables						
% of women in the *Staff*	1.300	6.95***	0.008	1.309	7.11***	0.008
Cost for employee (th.eur.)	0.014	13.43***	0.000	0.014	13.67***	0.000
Number of personnel	-0.002	12.65***	0.000	-0.002	12.89***	0.000
Number of observations	175965			175965		
Pseudo *R2*	9.9%			9.9%		
Wald's test	15000***			14999***		
LR test	16348***			16343***		

***p < .01, **p < .05, *p < .10. Maximum Likelihood estimations with STATA. For each of

the variables, a likelihood ratio test between the model that includes this variable and other

one in that this variable in question has been omitted has been computed.

Table 6

Poisson model on the number of women mentioned in the news item

Variable	Model I (All variables)			Model I (Only significant variables)		
	β	LR	p value	β	LR	p value
Constant	-2.920			-2.915		
Number of mentioned persons	-0.002	11.38***	0.001	-0.002	11.54***	0.001
Number of words in the news (log)	-0.025	4.98**	0.026	-0.024	4.97**	0.026
Woman mentioned in the title	0.879	3112.72***	0.000	0.879	3123.47***	0.000
News bylined by						
a woman	0.199	58.88***	0.000	0.199	58.98***	0.000
a mixed female/male team	0.235	10.92***	0.001	0.235	10.88***	0.001
Day of the week						
Tuesday	0.001	0.00	1.000			
Wednesday	0.036	3.02*	0.082	0.0354	4.12**	0.042
Thursday	0.033	2.36	0.124	0.0324	3.11*	0.078
Friday	0.033	2.50	0.114	0.0332	3.33*	0.068
Saturday	0.032	2.12	0.145	0.0314	2.76*	0.097
Sunday	0.062	7.96***	0.005	0.0616	10.38***	0.001
Month						
April	-0.108	61.05***	0.000	-0.108	61.06***	0.000
May	-0.071	27.13***	0.000	-0.071	27.08***	0.000
Sections						
Science	0.114	13.74***	0.000	0.111	14.53***	0.000
Culture	0.303	231.70***	0.000	0.299	291.62***	0.000
Sports	-0.970	1549.63***	0.000	-0.973	2019.75***	0.000
Economics	0.004	0.02	0.893			
People	0.560	648.17***	0.000	0.557	791.13***	0.000
International	0.008	0.13	0.723			
Society	0.442	481.33***	0.000	0.439	606.51***	0.000
Media related variables						
% of women in the *Staff*	2.085	8.03***	0.005	2.084	8.08***	0.004
Cost for employee (th.eur.)	0.022	14.03***	0.000	0.022	14.11***	0.000
Number of personnel	-0.003	12.41***	0.000	-0.003	12.46***	0.000
Number of observations	29960			29960		
Pseudo R^2	15.1%			15.1%		
Wald's test	9038***			9023***		
LR test	12564***			12568***		

***p < .01, **p < .05, *p < .10. Maximum Likelihood estimations with STATA. For each of the variables, a likelihood ratio test between the model that includes this variable and other one in that this variable in question has been omitted has been computed.
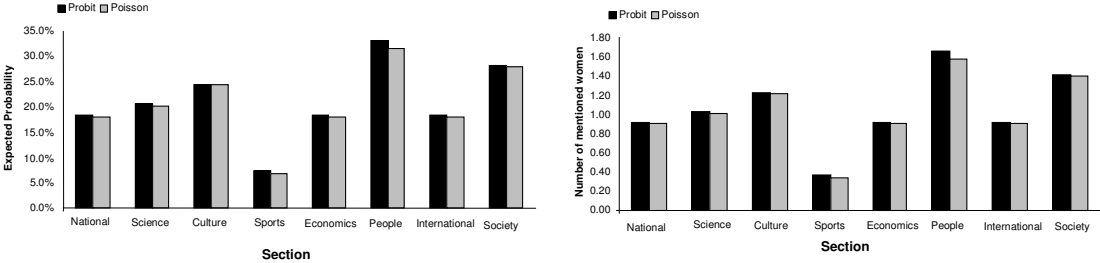
Figure 1

Sensibility to the section

Figure 2

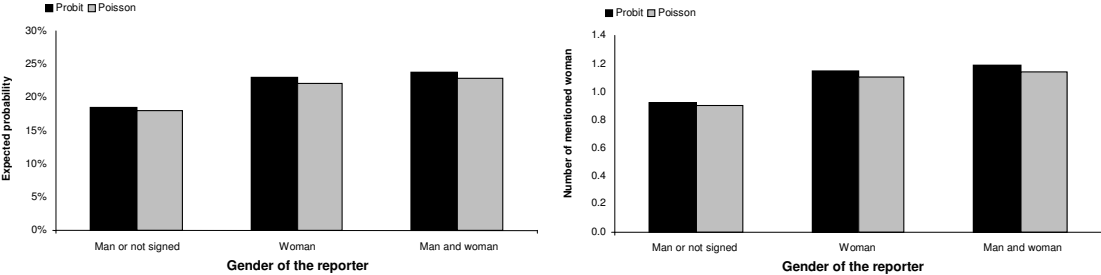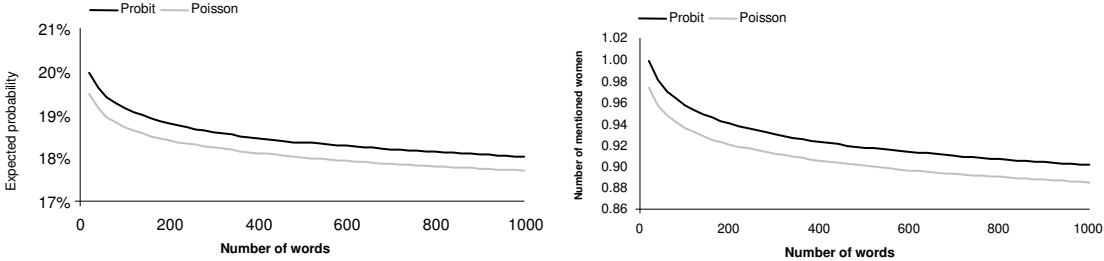*Sensibility to the gender of the reporter of the news item*

Figure 3

Sensibility to the length of the news

Figure 4

Sensibility to the day of the week