ORIGINAL ARTICLE

"She" and "He" in News Media Messages: Pronoun Use Reflects Gender Biases in Semantic Contexts

**Marie Gustafsson Sendén · Sverker Sikström ·
Torun Lindholm**

Abstract Previous research has shown a male bias in the media. This study tests this statement by examining how the pronouns *She* and *He* are used in a news media context. More specifically, the study tests whether *He* occurs more often and in more positive semantic contexts than *She,* as well as whether *She* is associated with more stereotypically and essential labels than *He* is. Latent semantic analysis (LSA) was applied to 400 000 Reuters' news messages, written in English, published in 1996–1997. LSA is a completely data-driven method, extracting statistics of words from how they are used throughout a corpus. As such, no human coders are involved in the assessment of how pronouns occur in their contexts. The results showed that *He* pronouns were about 9 times more frequent than *She* pronouns. In addition, the semantic contexts of *He* were more positive than the contexts of *She*. Moreover, words associated with *She*-contexts included more words denoting gender, and were more homogeneous than the words associated with *He*-contexts. Altogether, these results indicate that men are represented as the norm in these media. Since these news messages are distributed on a daily basis all over the world, in printed newspapers, and on the internet, it seems likely that this presentation maintains, and reinforces prevalent gender stereotypes, hence contributing to gender inequities.

Keywords Gender stereotypes · Gender · Media news · Latent semantic analysis · Sentiment analysis · Linguistic biases

M. Gustafsson Sendén (✉) · T. Lindholm
Department of Psychology, Stockholm University, 106
91 Stockholm, Sweden
e-mail: mgu@psychology.su.se

S. Sikström
Department of Psychology, Lund University, Lund, Sweden

Introduction

Evaluative communication of gender is, arguably, to a large extent conducted through language. Gendered pronouns (e.g., *she, he*) are excellent markers for gender in text, because they are easy to detect, and occur with a very high frequency in most contexts (Pennebaker et al. 2003; Twenge et al. 2012). Progress in computational methods now allows for quantification of sentiment, which can be used for measuring the valence of the context of gender pronouns (Gustafsson Sendén et al. 2014a). These methods may be used to provide new light on how the mass media subtly contribute to, and communicate gender stereotypes, as well as for devaluing women. Such reliable measurements have largely been missing in quantitative content analyses (Neuendorf 2011).

Analyzing how many pronouns there are in a given context may say something about gender status (Twenge et al. 2012), although evaluations can also be extracted by the semantic contexts in which the pronouns occur (Gustafsson Sendén et al. 2014a,b). Because words are associated with valence (Bradley and Lang 1999), it is possible to automatically assess the valence of the semantic contexts of the pronouns. Hence, the present study adds to the gender literature on how language can be subtly biased when it comes to the level of text production using pronouns. The approach is completely data-driven and the analyses are based on which words are used (or not used) in the semantic contexts around the pronouns. Four indices of gender biases are examined, namely differences in frequency, evaluations, and whether the semantic contexts around the pronouns include descriptive biases, such as essentializing language and homogeneity effects (Linville 1982; Maass et al. 2013).

To achieve this aim, we use a year's production of news from Reuters, which is the largest international news provider in the world, with more than 2,000 journalists in nearly 200 locations around the globe (www.reuters.com). Reuters

provides short media news articles that are reprinted and redistributed in magazines and internet services all over the world, reaching billions of people a day, giving it a critical impact on people's worldview (David et al. 2004; Rose et al. 2002). Because Reuters is a global actor, it is plausible that its news production influences news production in many countries, and also in other languages than English. Of further interest is that Reuters aims explicitly at providing unbiased news, and presentation of social groups, as reflected in their handbook of journalism:

> …A person's gender, race, religion, nationality, sexual orientation or marital status should not be cited unless it is relevant to the story. Even then, consideration must be given to where in the story such information needs to be placed. It is wrong to assume that police, firefighters or soldiers are men…. (Reuters handbook of Journalism 2014, "Writing")

In 2004, Reuters released a dataset for research and development of natural language processing, information retrieval and machine learning systems (David et al. 2004). The dataset consists of a year's worth of news production written in English, from 1996 to 1997 and is considered to be one of the largest datasets of media news (David et al. 2004; Rose et al. 2002).

Pronouns are especially well-suited for research on language biases since they reflect important social categories, for example inclusion and exclusion (We and They), individual or collective identities (I and We) or gender (She and He; Gustafsson Sendén et al. 2014b). Personal pronouns may also be used with less cognitive monitoring than words denoting explicit categories (Pennebaker et al. 2003), which makes them especially interesting in studying psychological biases and gender relations. Twenge and colleagues (2012) have shown that the ratio of male to female pronouns is an indicator of gender relations in US society, such that a lower male-to-female ratio of pronouns in US literature during the 20th century was correlated with women entering education and the labor market in the US. Automatic counting of pronouns is a highly effective way of examining biases, since pronouns are generic markers for gender and include most men and women referred to in articles. Thereby, human coders might be omitted from the most time-consuming parts of the analyses.

The analyses in the current study are conducted by latent semantic analysis (LSA; Landauer 1999; Landauer et al. 2007), which is a multi-dimensional computerized method that deduces the meaning of words by examining how words are used throughout a text corpus. LSA transforms a text corpus into a so-called semantic space by using computational algorithms. The transformation includes computations of the average meaning of how words are used throughout a corpus.

For example, the meaning of She and He is derived from every passage where these words occur, and also every occurrence where they do not occur (Landauer et al. 2007). The LSA is completely data-driven, making the results resilient against influence bias from the researcher. Computerized methods also allow analyses of large amounts of data where manual coding would be highly time-consuming (Neuendorf 2011; Rudy et al. 2010). For example, analyzing 400,000 articles (which is the data set used in this study) in a year would take 400,000 hours, which would correspond to 200 full-time coders working for a year if every article were analyzed for 1 hour (based on 2,000 working hours per year).

To summarize, in this study we will examine gender bias in media news that aims at being neutral. We will use pronouns as markers for gender and LSA for the analyses. This means that gender bias will be assessed at the level of text construction from both evaluative and descriptive aspects by computerized methods. First, we examine how gender pronouns vary with respect to frequency. Second, we analyze the evaluative semantic contexts in which the pronouns occur. Finally, we examine the meaning of He and She from how these pronouns are used throughout the corpus, and we consider whether essentializing language and homogeneity effects are present.

Gender Biases in Media News

Frequency The global media-monitoring project (www.whomakesthenews.org) investigates female representation in the media throughout the world. In 2010, it was found that women are underrepresented in the media throughout the world, to varying degrees and depending on region and news topic. A comparison of gender representation between the time the Reuters news was published, and now, shows that the global percentage of women in the media increased from 17 % in 1995 to 24 % in 2010. The same figures for North America were 21 % in 1994 and 28 % in 2010 (www.whomakesthenews.org 2010). In an earlier paper published in *Sex Roles*, (Collins 2011) indicated that mentions of men and women in the media vary significantly, and there are more mentions of men. The relative ratio of men and women in studies limited to specific cultures or topics varied from 3:1 (Armstrong 2004; Schwartz 2011) to 10:1 (Armstrong 2006; Koivula 1999). For example, Schwartz (2011) studied the representation of men and women in articles about attitudes towards homosexual marriage in four US newspapers, and found a ratio of three men for each woman. This indicated that also in topics challenging heteronormativity, male dominance was found. Matud and colleagues (2011) studied around 3,200 news articles published in 2005 in Spanish newspapers, and found a male dominance in news that was explained by a strong sexist culture in Spain at that time. On the other hand, (Koivula 1999) studied gender biases in one of the most gender-equal

countries in the world (Sweden), and she also found a strong male dominance, although in a stereotypical male topic, sports news. This means that no matter the country or type of story, female underrepresentation is clearly documented (Collins 2011; Rudy et al. 2010, 2011).

The study by (Twenge et al. 2012) showed that pronoun use in US literature was related to gender equality. About 1.2 million books released throughout the 20th century were analyzed. This study correlated the ratio of female to male pronouns with the status of women in the US society, and showed that the ratios of pronouns reflect changes in gender relations. A decrease in the male-to-female-ratio over the years was correlated with women's status such as educational attainment and labor force participation. Although the ratios have decreased, their results showed that even at the best of times, male pronouns outweighed female pronouns by a ratio of 2:1 (Twenge et al. 2012), indicating that text production is still associated with a male norm.

> H1: Based on the consistent finding of a male dominance in previous studies on media news and literature, a higher frequency of *He* than *She* was expected.

The study by (Twenge et al. 2012) also shows how pronouns may be used as proxies for gender, and are especially well-suited for text analysis of large corpora. In comparison, we examine the ratio of female and male pronouns in a different context (news media), and we also include analyses of the evaluative and the semantic contexts in which the pronouns occur.

Evaluations It has been argued that the higher frequency with which men appear in the media implicitly indicates that men are more highly evaluated than women (Matud et al. 2011; Rudy et al. 2010; Schwartz 2011), suggesting for example that men's voices are more important than women's voices (Twenge et al. 2012; Zoch and Turk 1998). In support of that notion, research has shown that men are overrepresented as experts whereas women are overrepresented as ordinary people in both Spanish (Matud et al. 2011) and US newspapers (Armstrong 2004; Zoch and Turk 1998). For example, Zoch and Turk (1998) studied 1,126 stories from the first page of South American newspapers and found that men were more often used as sources but also that men were reported in longer articles and with longer citations. Such findings have also been found in studies of US news by Armstrong (2004) and Freedman et al. (2007) and in studies of Spanish newspapers by Matud et al. (2011). For example Freedman and colleagues studied reporting of a US senate election campaign in the largest daily newspapers in the US. They found that men were more often used as sources than women, and that this gender gap was larger for expert than for non-expert sources (Freedman et al. 2007).

On the other hand, it could be argued that men are also especially frequently reported in the most negative topics, for example in topics related to crime and violence, which would contribute to more negative evaluations. However, a comparison of gender representation in different news topics (Whomakesthenews 2010) showed that topics such as crime and violence have a rather high female representation (21 %), in comparison to topics such as politics (7 %) and the economy (10 %). Because men more often occupy high positions within business and politics (Center for American Woman and Politics 2012; Eagly and Steffen 1984; Kanter 1977; The Catalyst 2013), and because the Reuters News corpus is dominated by such news, we expect that higher positions will be associated with men and also with more positive evaluative contexts.

> H2: *He* will occur in contexts of more positive valence than *She*.

It should be noted that evaluations of *He* and *She* are assessed by examining the semantic contexts in which these pronouns occur. Such analyses are important since they draw attention to the creation of evaluative differences between categories in text production.

Descriptive biases At the level of sentence production, there are linguistic consequences of being underrepresented in the media, such as being described with category labels, for example the *female* athlete or the *female* business director (Stahlberg et al. 2007). Such explicit labeling is in line with norm theory (Kahneman and Miller 1986) in that group comparisons are based on one group being the norm and the other group being "the effect to be explained". Being associated with the norm does not imply any specific labels or explicit categorization, whereas countering the norms does, and such linguistic normativity may also reinforce gender stereotypes (Bruckmüller et al. 2012; Hegarty and Pratto 2001; Miller et al. 1991). This type of communication has also been called essentializing language (Maass et al. 2013) because, for example, it presents women as representatives of their gender more often than men. Essentializing language might further lead to the homogeneity effect (Linville 1982), such that women as a group are described as more similar, than is the case for men as a group. In the media, such processes might be reflected in interviews with men and women. For example, discourse analyses of descriptions of female leaders in the Finnish media have found that interviews with female business leaders tend to focus on gender issues, family relations, and the fact that the leaders are women (Lämsä and Tiensuu 2002) instead of focusing on the specific industry in which the women work. In Swedish sports news, a similar effect was found (Koivula 1999), such that female athletes were asked questions about

their private life and about being a woman, but such questions were not addressed to male athletes. This means that an interview with a female football player might include the same content as an interview with a female swimmer, whereas interviews with men would focus on football and swimming respectively. Also in studies on gender differences in psychological articles, essentializing language when describing women has been found (Hegarty and Buechel 2006), such that female attributes were mentioned more often than male attributes.

To the best of our knowledge, there is one study (Lenton et al. 2009) that previously used LSA to study how gender-marked words are related to traits and roles in a corpus based on fictional and educational literature. The results showed that gender words (e.g., women, she/man, he), showed a stronger semantic similarity to gender-stereotypical professional roles (e.g., dietician, florist, / carpenter, fire-fighter) than traits (e.g., affectionate, caring loyal/aggressive, dominant, individualistic). In the current study, the analyses are broadened from roles and traits, and from literature to analyses of news media and the general semantic contexts in which *He* and *She* appear. Hence, this study focuses on pronouns and assesses the meaning of *She* and *He* by examining all words that are most proximal to *He* and *She* in a semantic space (e.g., the words with the highest cosines to *He* and *She*).

This means that descriptive biases such as essentializing language and homogeneity effects are assessed by investigating the extent of similarity between so called "semantic associates" of *He* and *She* in the semantic space. Semantic associates are words that co-occur with the pronouns throughout the corpus, either directly in the closest context, or indirectly by second order associations, and as such they describe or specify the meaning of a word (Landauer et al. 2007). To illustrate this, the semantic associates of *bad*, are *problem*, *performing*, and *worse*; whereas the semantic associates of *good*, are *pretty, excellent* and *positive* in the semantic space based on Reuter news. Based on theories and studies of essentializing language, the hypotheses were:

> H3: The semantic associates of *She*-contexts will include more gender labels and essentializing words than the semantic associates of *He*-contexts.

> H4: The semantic associates of the *She*-contexts will show a higher semantic similarity than the semantic associates of *He*-contexts.

The homogeneity effect is measured by investigating the semantic similarity between pronouns and their semantic associates. The semantic similarity is measured by the cosine, which can vary from +1 to −1, where a high positive cosine indicates a high similarity. To continue with the examples of *good* and *bad*, the cosines of the semantic associates of *bad*

are .62 (*problem*), .58 (*performing*), and .56 (*worse*), indicating a high similarity between the semantic associates and the target word *bad*.

To summarize the hypothesis, we expect that *He* will occur more frequently than *She* (H1), that *He* will occur in more positive contexts than *She* (H2), that the contexts of *She* will include more gender labels and essentializing words than the semantic contexts of *He* (H3) and, that the semantic associates of *She*-context will be more homogenous than the semantic associates of *He*-contexts (H4).

Method

LSA has been suggested as a theory and a method for deriving semantic meaning from the contexts in which words are used (Landauer and Dumais 1997; Landauer 1999; Landauer et al. 2007). Through computations, LSA derives the meaning of a word from every encounter with it and from the composition of all the passages in which a word does not occur. "…the similarity estimates derived by LSA are not based on simple contiguity frequencies, co-occurrence counts, or correlations in usage, but depend on powerful mathematical analysis that is capable of correctly inferring much deeper relations (thus the phrase *latent semantic*")" (Landauer et al. 1998, p. 261). The exact calculations in LSA are beyond the scope of this article, but suffice to say, LSA relies on singular value decomposition, which is akin to factor analysis and multidimensional scaling (Landauer and Dumais 1997; Landauer et al. 2007). For more detailed descriptions of LSA, we refer to methodological articles on LSA (see for example, Foltz et al. 1998; Landauer 1998; Landauer and Dumais 1997; Landauer et al. 2007) and empirical papers (see for example Campbell and Pennebaker 2003; Gustafsson Sendén et al. 2014a; Lenton et al. 2009). The semantic analyses conducted in this paper were made using SemanticExcel, which is web-based software developed for scientific analysis and statistical evaluation of semantic representation. This software was developed by the second author of this paper and is available on the webpage semanticexcel.com, and has been used for previous publications (see for example: Garcia and Sikstrom 2013; Karlsson et al. 2013, Gustafsson Sendén et al. 2014a). The most common way of using LSA is to compare similarities between words, word lists, or documents (see for example the study by Lenton et al. 2009). However, it is also possible to add other parameters such as valence to the representation, which is what we have done in this study and in previous studies on self- and group-serving biases (Gustafsson Sendén et al. 2014a). The method for introducing valence is explained further below.

Material

To create a semantic space with media news, a corpus with approximately 800,000 (404 MB) Reuters news messages published in 1996 and 1997 was used. When the current project was initiated, Reuters released this data set publicly, and it is one of the richest samples available for research (David et al. 2004; Rose et al. 2002). The dataset includes topics such as business (approximately 55 %), politics (30 %), international relations, war (10 %), and sports (5 %) (http://about.reuters.com/researchandstandards/corpus/statistics/index.asp). The articles are also categorized depending on what regions or countries are focused on in the news. English-speaking countries dominate the dataset, with the result that the US is the country most frequently mentioned (35 %). Approximately 32 % of the news mentions countries within the European Union (10 % includes the UK); Asian countries are mentioned in 25 % of the news (10 % refers to former parts of the British Empire: India, Australia and Hong Kong). The smallest percentages relate to African countries (4 %; South-Africa 2 %), South-American countries (2 %) and Canada (3 %). The coding was made with a combination of manual and computerized coding (for more information about the coding, see also Rose et al. 2002). The INFOMAP software (http://www.infomap-nlp.sourceforge.net) was used to perform the LSA algorithm. The Infomap NLP software performs automatic indexing of words and documents from free-text corpora, using a variant of LSA to enable information retrieval and other applications. It was developed by the Infomap project at Stanford University's "Center for studies of language and information" (Description of infomap algorithm; http://infomap-nlp.sourceforge.net/doc/algorithm.html.).

Because the quality of the semantic space is improved when words used to build the matrix contain meaningful semantic content (Landauer et al. 2007), about 750 high frequent, non-content words (e.g., and, but, who), were removed before reduction of the matrix. From the original corpus the 15,000 most common words were used to build a space with 100 dimensions. Two sentences (retrieved from www.reuters.com) are presented here to show how gendered pronouns might be used in a news context and the type of words included in the construction of a semantic space (excluded words are presented in parenthesis).

"…penalties would remain (in) (the) bill (as) (it) makes (its) way through (the) legislative process. (Butler) *she* said (the) thefts (are) becoming (more) common (as) biofuels increase (in) popularity. (It) (is) (already)…";
"… India (and) (the) United States, NN's tech career started (at) internet software pioneer XX. *He* joined xx

(in) 1992 (and) quickly climbed (the) corporate ladder (with) leading roles (in) (the) office…"

For more examples; see www.reuters.com.

Introducing Valence to the Semantic Space

In order to test hypothesis 2, the valence of the context of each pronoun was computed. Using human coders in such a project would be highly resource- and time-consuming. Moreover, due to the difficulty of achieving valid criteria that could be used by human coders to evaluate media text intended to be objective in nature, one would expect low inter-rater reliability. Hence, to predict valence from the semantic space, we trained the space based on a standardized word list (the Affective Norms for English Words; ANEW; Bradley and Lang 1999), with 1,034 words that were ranked for valence by human participants (on a Likert scale from 1–9). Multiple linear regression was used to create a vector across the 100 dimensions in the semantic space that best fitted the valence of the words in the ANEW word list. Hence, all the words in the ANEW words list were used as the dependent variables in the regressions (ANEW words that denoted gender were excluded from the regression) and the dimensions in the space were used as independent variables. To avoid over-fitting of the valence prediction, we used cross-validation by the "leave-one-out-method" (Picard and Cook 1984) to estimate how many dimensions should be included in the final model. The highest explained variance was achieved with 85 dimensions, such that using more dimensions yielded over-fitting and a less correct prediction model for valence. A single observation from the ANEW sample was used as the validation data, and the remaining ANEW data was used to predict the valence. This was repeated so that each word in the ANEW sample was used once as the validation data. The correlation between the cross-validated sample and ANEW was=.48, which is an estimate of the valence prediction for all the words in the semantic space that were not rated by human agents. The correlation based on all ANEW words in the space and the ANEW list was=.62, which is similar to previous measures of inter-rater congruence by LSA and human raters (Landauer et al. 2007). In the last step, every pronoun occurrence was given a valence by averaging the valence of the 15 words preceding and following each pronoun. This means that each pronoun valence was derived from how it was used in each context. The valences for all the pronouns were then analyzed in SPSS, where the comparisons of the valences of the pronouns were computed by ANOVAs and T-tests.

Extracting Semantic Associates to Deduce Meaning

LSA creates a representation of each word as a vector in a multidimensional semantic space. The words that are

semantically similar to a given word are "semantic associates", and as such, they should represent the meaning of a target word or a target context. As previously described, the closest semantic associates to *bad* are *problem*, *performing* and *worse*, whereas the closest semantic associates to *good*, are *pretty*, *excellent* and *positive*. The semantic associates can be validated by synonym tests (Landauer et al. 2007). In this study, the semantic associates are extracted in relation to the contexts of *He* and *She*. It should be noted that the semantic space is multidimensional and that the semantic associates include both words that co-occur in the space around the pronouns, and words that are used in similar contexts in other articles. Consider, for example, the words *physician* and *doctor*; those words may not necessarily be used in the same 15-word spaces; however, because synonyms are used in similar contexts they also become semantic associates.

Results

Hypotheses 1, 2 and 4 included quantitative analyses with gendered pronouns as the independent variable, and frequency, evaluations and semantic similarity (descriptive bias) as dependent variables. Hypothesis 3, about essentializing language and gender labels, was tested by qualitative interpretation of the semantic associates.

The first hypothesis concerned the relative ratios of *He* and *She*. As predicted, *He* ($N = 81,630$) outnumbered *She* ($N = 9,075$) by a relative ratio of approximately 9:1; that is, nine instances of *He* for each *She*.

To test the second hypothesis, that the valence of the contexts around *He* would be more positive than the contexts around *She*, we computed a univariate ANOVA, with valence as the dependent variable and gender as the independent variable. The results showed that the contexts of *He* ($M = 6.63$, $SD = 1.35$) had a more positive valence than the contexts of *She* ($M = 6.40$, $SD = 1.31$), $F (1, 90,702) = 226.4$, $p < .001$, Cohen's $d = 0.17$. Since the differences in frequencies were large, and the variance heterogeneous (Levene's test: $F (1, 90,702) = 7.50$, $p = .006$), we also computed an ANOVA with a randomly selected subset of *He*-pronouns ($N = 9,700$). These results were similar, such that *He* ($M = 6.64$, $SD = 1.34$) occurred in contexts of more positive valence than *She* ($M = 6.40$, $SD = 1.31$), $F (1, 18,773) = 150.78$, $p < .001$, Cohen's $d = 0.18$. Finally, two t-tests were used to examine whether *He* and *She* differed from the average valence in the semantic space ($M = 6.52$, $SD = 1.38$). The contexts of *He* were more positive than the average in the space $t (81,629) = 22.70$, $p < .001$, Cohen's $d = .08$; and the contexts of *She* were more negative than the average valence in the space $t (9,075) = -8.47$, $p < .001$, Cohen's $d = -.08$.

The third and fourth hypotheses included the descriptive biases, and how *He* and *She* were used throughout the Reuters corpus. The third hypothesis was that the semantic associates of *She* would include more essentializing language and gender labels than would the semantic associates of *He*. The fourth hypothesis was that the semantic associates of *She* would show a higher mean cosine than the semantic associates of *He*. Table 1 includes the semantic associates and the cosines for the *He* and *She* contexts. The third hypothesis was tested by a qualitative analysis of the semantic associates. As can be seen in Table 1, the semantic associates of *She* are gendered pronouns (her, herself), labels related to family relations (mother, husband), and feminine labels (woman, girl). Some words could also be interpreted as passive or responsive (answer, admitted, refused). The semantic associates of *He* include only one gendered pronoun (his), no gender labels, and words related to actions or directives (made, make, adding, give).

To test the fourth hypothesis, about higher homogeneity between the semantic associates of the contexts around *She* than the contexts of *He*, the mean cosines were compared. The cosine for each semantic associate in Table 1 represents the semantic distance between the target word (*He/She*) and the semantic associate. A *t*-test showed that the average cosine for the female semantic associates ($M = .55$, $SD = .054$) was higher than the average of male semantic associates ($M = .50$, $SD = .027$), $t (18) = -2.71$, $p = .014$, Cohens $d = 1.18$, supporting the view that the meaning of *She* is more homogenously presented in Reuter news than *He* is.

Discussion

In this study, gender pronouns in news media were analyzed with regard to frequency, as well as by the evaluative and descriptive semantic contexts in which they occur. This was done using multi-dimensional computerized semantic analyses – LSA (Landauer et al. 2007).

The analyses comparing the relative frequencies of *He* and *She* showed that the male dominance was substantial. For each *She* there were nine instances of *He*. Male dominance of a similar extent has previously been demonstrated in Swedish sports news (i.e., Koivula 1999) but general news coverage in the US typically shows ratios of between 2 to 4 men for each woman (see for example; Armstrong 2004; Schwartz 2011). As we did not control for whether the occurrences referred to the same person or not, the large difference in ratios could either mirror the actual numbers of men and women in the news, or may indicate that men were given more space, for example by occurring in longer articles with more citations. In support of this notion, studies have shown that news articles about women were shorter than articles about men, and that news about men was located in more salient positions on the
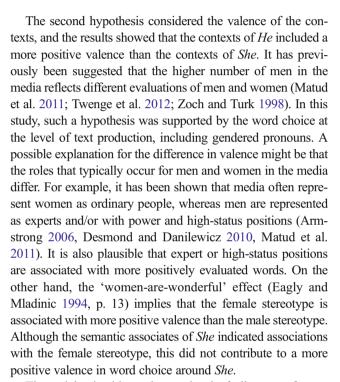
Table 1 The meaning of *He* and *She* as extracted by the semantic associates in the semantic space

Words	1	2	3	4	5	6	7	8	9
He-contexts	He	made	clear	make	adding	quickly	his	who	giving	give
Cosine	.55	. 52	.51	.50	.50	.49	.49	.49	.48	.46
She-contexts	She	her	woman	mother	husband	herself	girl	answer	admitted	refused
Cosine	.66	.62	.57	.56	.56	.54	.51	.50	.50	.50

The nine most nearby semantic associates are presented in order to how close they are to the *He* and *She* contexts. The cosine (could vary from −1 to+1) below each word is an indicator of the distance between the pronoun context and each word. A higher value indicates more association.

newspaper page (Armstrong 2004; Freedman et al. 2007; Matud et al. 2011). If women are given less space and/or fewer citations, it could be argued that the media consider women to be of less interest than men (Zoch and Turk 1998). On the other hand, the ratios could also be compared to women participating in politics and business, which are the main topics in Reuters news. During the time from 1996 to 2013, female representation in higher political and business positions increased, although men were still in the majority. For example, in the US congress, female representation increased from 11.8 % in 1998 to 18.5 % in 2012 (Center for American Woman and Politics 2012), and female government positions worldwide increased from 3 % in 1946 to 18.6 % in 2009 (Inter-parliamentary 2010), which indeed shows a change, but a rather slow change. Among US Fortune 500 companies, female board members increased in number from 9.6 % in 1995 to 16.9 % in 2013 (The Catalyst 2013). Future studies could therefore test whether the pronoun ratio reflects female and male distributions within specific news topics, such as politics, business or sports.

One previous study examined the ratios of gendered pronouns in a large corpus of text. (Twenge et al. 2012) studied English literature (fiction and factual literature) published mainly in the 20th century, and found a male to female ratio of 2:1 at the end of 2,000, with a negligible change during the time period 1996–2010. The large difference between the ratios found in our study and the study of Twenge needs to be addressed in future research. One possible reason could be that the audience of texts also influences the ratio of men and women in the texts, such that the smaller ratio in fiction literature is due to the fact that women read more fiction than men do (Tepper 2000), whereas men consume more news than women do (Benesch 2012). However, it should be remembered that although women read more fiction, and that literature may be female-oriented, the number of men is still double in this type of text (Twenge et al. 2012). No matter what the reason is for the ratio of male to female pronouns in the Reuter news, the size of the ratio must influence the perceptions of whose voices are important, perceptions of gender roles, and also the motivation for women to strive for high positions in business and politics.

The second hypothesis considered the valence of the contexts, and the results showed that the contexts of *He* included a more positive valence than the contexts of *She*. It has previously been suggested that the higher number of men in the media reflects different evaluations of men and women (Matud et al. 2011; Twenge et al. 2012; Zoch and Turk 1998). In this study, such a hypothesis was supported by the word choice at the level of text production, including gendered pronouns. A possible explanation for the difference in valence might be that the roles that typically occur for men and women in the media differ. For example, it has been shown that media often represent women as ordinary people, whereas men are represented as experts and/or with power and high-status positions (Armstrong 2006, Desmond and Danilewicz 2010, Matud et al. 2011). It is also plausible that expert or high-status positions are associated with more positively evaluated words. On the other hand, the 'women-are-wonderful' effect (Eagly and Mladinic 1994, p. 13) implies that the female stereotype is associated with more positive valence than the male stereotype. Although the semantic associates of *She* indicated associations with the female stereotype, this did not contribute to a more positive valence in word choice around *She*.

The articles in this study consisted of all types of news published by the Reuters. Future research should examine whether the pattern found varies in different topics. For example, it might be assumed that men are associated with more positive valence in topics related to politics, the economy and sports, whereas topics about crime and violence for example might reveal an opposite pattern. It would also be interesting to test whether a higher ratio within a topic is related to more positive evaluations for the more common category, so that when women are in the majority, *She* would also be seen in more positive semantic contexts.

Hypothesis 3 included a qualitative analysis of the descriptive content of the words associated with *He* and *She*. The semantic associates extracted by the LSA represent the meaning of a word as it is used throughout the corpora of media news (Foltz et al. 1998; Landauer et al. 2007). The semantic associates of *She* were more gendered (e.g., mother, women and girl) than were the semantic associates of *He*, which supports the notion of essentializing language of women (Maass et al. 2013) and norm theory (Kahneman and Miller

1986) where underrepresented categories are more often labeled in accordance with their category (Kahneman and Miller 1986; Kanter 1977; Stahlberg et al. 2007). During the 20th century, perceptions of women changed such that women are now considered more agentic than previously (Kanter 1977, Rudman and Glick 2001; Twenge et al. 2012). However, prescriptions of female communality are still strong today (Rudman and Glick 2001), and it is argued that despite important societal changes, communal traits for women are still endorsed and this will continue into the foreseeable future (Diekman and Eagly 2000). We also found two differences worth noticing that were not hypothesized. First of all the semantic associates of *She* included more nouns, whereas the semantic associates of *He* included more verbs. Previous research has found that nouns have more inductive potential, more strongly indicate group membership, and are thereby associated with essentializing language (Carnaghi et al. 2008, Graf et al. 2012; Maass et al. 2013). Second, the semantic associates may also reflect gender differences by describing men as active and women as passive. For example, *He*-associates included more verbs than *She*-associates (i.e., *made* and *give*). The verbs among the *She*-associates were also more responsive than assertive (i.e., *response, admit*). Actually, we did not expect such a pattern to occur with only nine associates of the *She*- and *He*-contexts. It is also important to note that these generalizations are based on computerized, data-driven analyses of words that are used (and not used) in the semantic contexts around the pronouns. Altogether, these results further support the interpretation that news media frame their descriptions of men and women differently in many various ways.

In this study, we analyzed the first nine semantic associates (i.e., the nine words that are the most closely located in the semantic space to *She* and *He*). This number was chosen because the first associates are the most interesting and meaningful to examine. The further apart the associates are, the more associated they are also to other words and concepts. We could also have chosen fewer associates but this would have given a weaker comparison of the mean similarity between *He* and *She* and the semantic associates. Besides, using nine associates gave more salient differences than predicted.

A homogeneity effect (i.e., the effect of essentializing language) was tested in Hypothesis 4, where the distances between the semantic associates and the target words *He* and *She* were compared. Homogeneity effects (Linville 1982) refer to how descriptions of different groups differ, where for example individuals in an outgroup are referred to as more similar to each other than are individuals in the ingroup. If news media are associated with men, men might be considered and described as the ingroup and women as the outgroup, resulting in higher homogeneity in descriptions of women. Indeed, the results showed that the semantic associates of *She* were more homogenous to *She* than were the semantic associates of *He*. A side effect of the homogeneity effect is that

men are described as more differentiated. Such differences were also present in media presentations of Swedish athletes and Finnish business leaders (Koivula 1999; Lämsä and Tiensuu 2002), in that women were asked very similar questions, no matter what sport they performed, or what business they ran. These results are also in alignment with studies of other media, for example prime-time fictional television in the US, where it was found that women were usually shown in family roles, whereas men roles were depicted in various types of work roles, (Lauzen et al. 2008), and when women worked, their occupations were more narrow and traditional than the occupations of men (Signorielli and Kahlenberg 2001). Hence, the assessment of homogeneity effects could be an indicator of how open or permissive a context is to men and women, and the differences could potentially lead to more differentiated role models for men than women.

Limitations and Future Directions

The dataset ranges back to the late 90s, and the relevance of the findings for today could be discussed. When the current project was initiated, this Reuters sample released in 2004 from 1996 to 1997 was the largest, and one of the richest samples available for research (David et al. 2004). The quality of the corpora and the coding is also well documented (Rose et al. 2002). As computerized methods are developed, and written documents become more and more available, this dataset could also be of importance as a starting point for future comparisons.

It is possible that the ratios of gendered pronouns have changed, since more women are now in positions associated with power and media representation. For example, global studies of the media (www.whomakesthenews.org) show a change from 1:5 to 1:4 women per men, and female governmental representation has changed from 1:10 to 1:5 in the same period (UN 2009). When it comes to evaluative and descriptive differences, it is an open question whether they also have diminished or whether they are still the same. Some research indicates that although frequencies change, the relative difference in evaluations and stereotypes remains the same (Ridgeway 2001).

This is, as far as we know, the first study implementing computational methods for studying gender bias beyond frequency. These types of analyses have been enabled by development in computerized methods (Landauer et al. 2007, Pitts and Nussbaum 2006) and by corpora that have been made available for research. The corpus we used was based on English news. Although this might indicate that the results are limited to English-speaking countries, we suggest that the results are also valid for other countries because the Reuters news is also translated and distributed outside the English-speaking cultures. Because content analyses are time consuming, there have been few studies comparing different sorts of written material and cultural aspects (Neuendorf 2011). Analyzing pronouns with computerized methods may facilitate

future research that compares various discourses. Future research might also investigate whether author gender is related to the evaluative contexts of *She* and *He*. Studies have shown that female authors more often than male authors use women as sources in the media (Armstrong 2004). The question is whether authors are biased in reporting their own gender in more positive semantic contexts.

The partial effect sizes of the valence differences in the current study were small. Field studies, as compared to laboratory experiments, often have smaller effect sizes (Evans 1985; McClelland and Judd 1993), since they include many sources of variance. McClelland and Judd (1993) estimated that field experiments are 80 % less efficient than experiments in a laboratory setting. Reuters news includes both positive topics such as sports and business, and negative topics such as war and crime, and possibly, the topics are more important predictors of valence than pronouns are; hence large effect sizes were not expected in this material. However, people are soaked in news from different media throughout the day, and given this massive exposure and constant repetition, small and subtle evaluative differences are likely to have implications for how people construe their social world (Collins 2011; Geschke et al. 2010).

Conclusion

The current study constitutes the first attempt to investigate gender biases at the level of text production in media news by examining pronouns and their semantic contexts. We believe that this is promising approach. While it has previously been claimed that the media presents men more favorably than women, this study showed that biases come about in various ways; by sheer numbers, in valence of contexts and by word choice. *He* occurred 9 times more often than *She*, in more positive contexts, and with words in accord with gender stereotypes. Because the differences are subtle, repeated on a daily basis, and spread in various forms of news media, the gender representations in news media are likely to maintain and reinforce gender inequity.

References

Armstrong, C. L. (2004). The influence of reporter gender on source selection in newspaper stories. *Journalism and Mass Communication Quarterly, 81*, 139–154. doi:10.1177/107769900408100110.

Armstrong, C. L. (2006). Story genre influences whether women are sources. *Newspaper Research Journal, 27*, 66–81.

Benesch, C. (2012). An empirical analysis of the gender gap in news consumption. *Journal of Media Economics, 25*, 147–167. doi:10.1080/08997764.2012.700976.

Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Stimuli, instruction manual and affective ratings*. Gainesville: The NIMH center for the study of emotion and attention, University of Florida.

Bruckmüller, S., Hegarty, P., & Abele, A. E. (2012). Framing gender differences: Linguistic normativity affects perceptions of power and gender stereotypes. *European Journal of Social Psychology, 42*, 210–218. doi:10.1002/ejsp.858.

Campbell, R. S., & Pennebaker, J. W. (2003). The secret life of pronouns: Flexibility in writing style and physical health. *Psychological Science, 14*, 60–65. doi:10.1111/1467-9280.01419.

Carnaghi, A., Maass, A., Gresta, S., Bianchi, M., Cadinu, M., & Arcuri, L. (2008). Nomina sunt omina: On the inductive potential of nouns and adjectives in person perception. *Journal of Personality and Social Psychology, 94*, 839–859. doi:10.1037/0022-3514.94.5.839.

Center for American Woman and Politics. (2012). Fact sheet. Retrieved from http://www.cawp.rutgers.edu/fast_facts/levels_of_office/documents/elective12.pdf

Collins, R. L. (2011). Content analysis of gender roles in media: Where are we now and where should we go? *Sex Roles, 64*, 290–298. doi:10.1007/s11199-010-9929-5.

David, D. L., Yiming, Y., Tony, G. R., & Fan, L. (2004). RCV1: A new benchmark collection for text categorization research. *Journal of Machine Learning Research, 5*, 361–397.

Desmond, R., & Danilewicz, A. (2010). Women are on, but not in, the news: Gender roles in local television news. *Sex Roles, 62*, 822–829. doi:10.1007/s11199-009-9686-5.

Diekman, A. B., & Eagly, A. H. (2000). Stereotypes as dynamic constructs: Women and men of the past, present, and future. *Personality and Social Psychology Bulletin, 26*, 1171–1188. doi:10.1177/0146167200262001.

Eagly, A. H., & Mladinic, A. (1994). Are people prejudiced against women? Some answers from research on attitudes, gender stereotypes, and judgments of competence. *European Review of Social Psychology, 5*, 1–35. doi:10.1080/14792779543000002.

Eagly, A. H., & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology, 46*, 735–754. doi:10.1037/0022-3514.46.4.735.

Evans, M. G. (1985). A Monte-Carlo study of the effects of correlated method variance in moderated multiple-regression analysis. *Organizational Behavior and Human Decision Processes, 36*, 305–323. doi:10.1016/0749-5978(85)90002-0.

Foltz, W., Kintsch, W., & Landauer, K. (1998). The measurement of textual coherence with latent semantic analysis. *Discourse Processes, 25*, 285–307. doi:10.1080/01638539809545029.

Freedman, E., Fico, F., & Love, B. (2007). Male and female sources in newspaper coverage of male and female candidates in US Senate Races in 2004. *Journal of Women Politics & Policy, 29*, 57–76. doi:10.1300/J501v29n01_04.

Garcia, D., & Sikstrom, S. (2013). A collective theory of Happiness: Words related to the word "Happiness" in Swedish online newspapers. *Cyberpsycholology, Behavior and Social Networking, 469–472* doi: 10.1089/cyber.2012.0535.

Geschke, D., Sassenberg, K., Ruhrmann, G., & Sommer, D. (2010). Effects of linguistic abstractness in the mass media: How newspaper articles shape readers' attitudes toward migrants. *Journal of Media Psychology: Theories Methods and Applications, 22*, 99–104. doi:10.1027/1864-1105/a000014.

Graf, S., Bilewicz, M., Finell, E., & Geschke, D. (2012). Nouns cut slices: Effects of linguistic forms on intergroup bias. *Journal of Language and Social Psychology, 32*, 62–83. doi:10.1177/0261927x12463209.

Gustafsson Sendén, M., Lindholm, T., & Sikstrom, S. (2014a). Biases in news media as reflected by personal pronouns in evaluative contexts. *Social Psychology, 45*, 103–111.

Gustafsson Sendén, M., Lindholm, T., & Sikstrom, S. (2014b). Selection bias in choice of words: Evaluations of "I" and "We" differ between contexts, but "They" are always worse. *Journal of Language and Social Psychology, 33*, 49–67. doi:10.1177/0261927X13495856.

Hegarty, P., & Buechel, C. (2006). Androcentric reporting of gender differences in APA journals: 1965–2004. *Review of General Psychology, 10*, 377–389. doi:10.1037/1089-2680.10.4.377.

Hegarty, P., & Pratto, F. (2001). The effects of social category norms and stereotypes on explanations for intergroup differences. *Journal of Personality and Social Psychology, 80*, 723–735. doi:10.1037//0022-3514.80.5.723.

Inter-parliamentary Union (2010). *Is parliament open to women? An appraisal*. Retrieved from http://www.ipu.org/english/surveys.htm

Kahneman, D., & Miller, D. T. (1986). Norm theory - Comparing reality to its alternatives. *Psychological Review, 93*, 136–153. doi:10.1037/0033-295x.93.2.136.

Kanter, R. M. (1977). *Men and women of the corporation*. New York: Basic Books.

Karlsson, K., Sikström, S., & Willander, J. (2013). The semantic representation of event information depends on the cue modality: An instance of meaning-based retrieval. *PLoS ONE, 8*, e73378. doi:10.1371/journal.pone.0073378.

Koivula, N. (1999). Gender stereotyping in televised media sport coverage. *Sex Roles, 41*, 589–604. doi:10.1023/a:1018899522353.

Landauer, T. K. (1998). Learning and representing verbal meaning: The latent semantic analysis theory. *Current Directions in Psychological Science, 7*, 161–164. doi:10.1111/1467-8721.ep10836862.

Landauer, T. K. (1999). Latent semantic analysis: A theory of the psychology of language and mind. *Discourse Processes, 27*, 303–310. doi:10.1080/01638539909545065.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104*, 211–240. doi:10.1037/0033-295X.104.2.211.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes, 25*, 259–284. doi:10.1080/01638539809545028.

Landauer, T. K., McNamara, D. S., Dennis, S., & Kintsch, W. (2007). *Handbook of latent semantic analysis*. Mahwah: Lawrence Erlbaum.

Lauzen, M. M., Dozier, D. M., & Horan, N. (2008). Constructing gender stereotypes through social roles in prime-time television. *Journal of Broadcasting & Electronic Media, 52*, 200–214. doi:10.1080/08838150801991971.

Lenton, A., Sedikides, C., & Bruder, M. (2009). A latent semantic analysis of gender stereotype-consistency and narrowness in American English. *Sex Roles, 60*, 269–278. doi:10.1007/s11199-008-9534-z.

Linville, P. W. (1982). The complexity-extremity effect and age-based stereotyping. *Journal of Personality and Social Psychology, 42*, 193–211. doi:10.1037/0022-3514.42.2.193.

Lämsä, A., & Tiensuu, T. (2002). Representations of the woman leader in Finnish business media articles. *Business Ethics: A European Review, 11*, 363–374. doi:10.1111/1467-8608.00296.

Maass, A., Suitner, C., & Merkel, A. (2013). Does political correctness make (social) sense? In J. P. Forgas, O. Vincze, & J. László (Eds.), *Social Cognition and communication* (pp. 331–346). New York: Psychology Press.

Matud, M. P., Rodríguez, C., & Espinosa, I. (2011). Gender in Spanish daily newspapers. *Sex Roles, 64*, 253–264. doi:10.1007/s11199-010-9874-3.

McClelland, G. H., & Judd, C. M. (1993). Statistical difficulties of detecting interactions and moderator effects. *Psychological Bulletin, 114*, 376–390. doi:10.1037/0033-2909.114.2.376.

Miller, D. T., Taylor, B., & Buck, M. L. (1991). Gender gaps: Who needs to be explained? *Journal of Personality and Social Psychology, 61*, 5–12. doi:10.1037/0022-3514.61.1.5.

Neuendorf, K. A. (2011). Content analysis—A methodological primer for gender research. *Sex Roles, 64*, 276–289. doi:10.1007/s11199-010-9893-0.

Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology, 54*, 547–577. doi:10.1146/annurev.psych.54.101601.145041.

Picard, R. R., & Cook, R. D. (1984). Cross-validation of regression models. *Journal of the American Statistical Association, 79*, 575–583. doi:10.2307/2288403.

Pitts, M. J., & Nussbaum, J. F. (2006). Integrating the past and paving the future - Examining current trends and extending boundaries of language and social psychology research. *Journal of Language and Social Psychology, 25*, 197–202. doi:10.1177/0261927x06289421.

Reuters handbook of journalism. (2014). Writing: Discriminatory language. Retrieved from http://handbook.reuters.com/?title=W#discriminatory_language

Ridgeway, C. L. (2001). Gender, status, and leadership. *Journal of Social Issues, 57*, 637–655. doi:10.1111/0022-4537.00233.

Rose, T., Stevenson, M., & Whitehead, M. (2002). *The Reuters Corpus Volume 1 - from yesterday's news to tomorrow's language resources*. Paper presented at the Proceedings of the 3rd International Conference on Language Resources and Evaluation.

Rudman, L. A., & Glick, P. (2001). Prescriptive gender stereotypes and backlash toward agentic women. *Journal of Social Issues, 57*, 743–762. doi:10.1111/0022-4537.00239.

Rudy, R. M., Popova, L., & Linz, D. G. (2010). The context of current content analysis of gender roles: An introduction to a special issue. *Sex Roles, 62*, 705–720. doi:10.1007/s11199-010-9807-1.

Rudy, R. M., Popova, L., & Linz, D. G. (2011). Contributions to the content analysis of gender roles: An introduction to a special issue. *Sex Roles, 64*, 151–159. doi:10.1007/s11199-011-9937-0.

Schwartz, J. (2011). Whose voices are heard? Gender, sexual orientation, and newspaper sources. *Sex Roles, 64*, 265–275. doi:10.1007/s11199-010-9825-z.

Signorielli, N., & Kahlenberg, S. (2001). Television's world of work in the nineties. *Journal of Broadcasting & Electronic Media, 45*, 4–22. doi:10.1207/S15506878jobem4501_2.

Stahlberg, D., Braun, F., Irmen, L., & Sczesny, S. (2007). Representation of the sexes in language. In K. Fiedler (Ed.), *Social Communication* (pp. 163–187). New York: Psychology Press.

Tepper, S. J. (2000). Fiction reading in America: Explaining the gender gap. *Poetics, 27*, 255–275. doi:10.1016/S0304-422X(00)00003-6.

Catalyst, T. (2013). *Statistical overview of women in the workplace*.

Twenge, J. M., Campbell, W. K., & Gentile, B. (2012). Male and female pronoun use in US books reflects women's status, 1900–2008. *Sex Roles, 67*, 488–493. doi:10.1007/s11199-012-0194-7.

UN (2009). *Is parliament open to women? An appraisal conference for chairpersons and members of parliamentary bodies dealing with gender equality*. Geneva: Inter-Parliamentary Union.

Whomakesthenews (2010). Who makes the news? The global media monitoring project. Retrieved from http://www.whomakesthenews.org

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*, 762–775. doi:10.1177/107769909807500410.