

# Gender Bias in Language Comprehension; Exploring the Role of Gender Bias in Adjectives of the Danish Language

Lasse Hyldig Hansen, 201908623@post.au.dk/lassehyldigh@gmail.com  
Kristian Severin Mengel-Niemann, 201904957@post.au.dk/kmnseverin@gmail.com  
University of Aarhus, School of Communication and Culture

## Abstract

*Humans depend on top-down processes in their understanding of language. Previous studies have shown how gender biases are apparent in human communication and have suggested that these come from previous experiences that lay the grounds for a heuristical approach to language comprehension. In this paper, we suggest that there is a gender bias in some of the most common Danish adjectives. Furthermore, we indicate that this gender bias can be used to alter the way in which Danish speakers understand seemingly gender-neutral, syntactical subjects. This is done through statistical analyses made on the basis of results from online questionnaires. We discuss our results with a starting point in literature about social role theory and gender bias and use this as grounds for proposing that Danish speakers are heuristically limited by gender stereotypes when assessing the gender of a gender-neutral, grammatical subject in written Danish.*

**Keywords:** gender bias, heuristics, top-down processes, Danish adjectives

## Top-Down Processes and Prediction for Efficiency in Language

Humans are social beings who depend on social cognition in language acquisition (Fitch et al., 2010). To make language processing more efficient, we rely on top-down processes. These processes are the integration of existing knowledge with perceptual information (Ferreira & Chantavarin, 2018). The topic of top-down processes as a key stone for fast and comprehensive understanding of language has been heavily debated: for instance, some argue that given our innate understanding of language processes, making use of a priori perceptual knowledge would be a waste of cognitive load (Jackendoff, 2002; Van Petten & Luka, 2012). However, newer research seems to indicate that top-down processes in fact play a crucial role in language processing as the most efficient way for us to process language in everyday life: for instance Kleinschmidt & Jaeger (2015) argue that we make use of previously perceived stimuli that we in an automatized manner use to predict how language is going to be communicated to us. They argue that this is done all the way down to the level of physical attributes of phonetic communication. These predictions, made by our speech perception system, come from experience with previous situations and being able to generalize these adaptations to other situations. This means that we have to make inferences about what messages other people might be conveying based on predictions. Kuperberg & Jaeger (2016)

describe this probabilistic way of using prediction in language comprehension. They use the example of garden path sentences (sentences that are grammatically correct but constructed so that the readers' initial interpretation of them most likely will be wrong) as evidence for top-down processing as a mean to gain contextual predictability in language processing. An example of a garden path sentence that is structured in an ambiguous way so that we initially construe it as grammatically incorrect is given in the following:

(1a) After the student moved the chair...

(1b) ...broke

Initially (1a) makes the reader interpret the sentence as though it is the chair that is being moved. However, with the inclusion of (1b) it is in fact the student that is moving. This is a clear example of how top-down processes of language can bias our initial understanding of meaning.

When this ambiguity is removed, these processes allow for more fluent communication as predictability of context may lead to faster perception of words and meaning. Arnon & Snider (2010) showed that combinations of word strings that we encounter more often are read faster. This would allow for more fluent communication as we would have an easier time reading words that are more often encountered and thereby more available for our immediate understanding.

## **Dichotomization Leading to Stereotypical Bias**

These predictions do not seem to be exclusive to language processing. They seem to be an attribute to our understanding of a diverse world in general (Qian, Jaeger, & Aslin, 2012 and Kuperberg & Jaeger, 2016).

This way of predicting and categorizing the perceptual world is a very efficient way for humans to construe the world. It makes us more efficient, also regarding language processing. We utilize previously encountered stimuli and are able to infer the intended meaning of the language communicated. This ranges from the very lowest part of language processing of words, to higher socially communicative situations. A very simple example of a low-level language prediction is the phenomenon known as the word superiority effect, in which letters are better understood in the context of words than as individual letters (Anderson, 2014, p. 48).

However, predicting and categorizing language stimuli in such a big range comes with its costs, because our cognitive predictions do not change proportionally to social conventions such as gender equality. This is also why top down processes and predictions in language processing can come with complications; they can potentially lead to stereotypical bias in our understanding of language, as society develops faster than prediction efficiency in communication.

This was suggested already before cognitive theories were included in linguistics. Whorf (1956, p. 252) exceeded his time and gave his take on how culture, language, and consciousness coexist and influence one another:

“And every language is a vast pattern-system, different from others, in which are culturally ordained the forms and categories by which the personality communicates, but also analyzes nature, notices or neglects types of relationship and phenomena, channels his reasoning, and builds the house of his consciousness.”

If culture shapes predictions of what words are associated with certain things, it may lead to a certain bias in our language. If certain attributes are assigned females and not males, or vice versa, it may seem that these predictions form a gender-biased language, possibly shaped by a gender-biased culture.

The roots of gender stereotypes are hard to designate, however according to the social role theory (Koenig & Eagly, 2014), they may originate from society's gender norms. Social role theory arose as a tool to understand differences in social behavior to shed light on whether gender behavior differs (Eagly, 1987, see also Eagly & Wood, 1991). Here Eagly described how gender differences in society derive from social dogmatism, namely beliefs about how a female and male should act.

Throughout time, women and men have held different occupations and roles in society. In general, men have held occupations that require leadership skills and physical attributes such as strength and speed and have therefore been described with such features (Eagly et al., 2000). Women on the other hand, have been described with more caring/warm traits (Cuddy et al., 2004 and Eagly & Wood, 2002). These findings could be explanatory of why we potentially see a difference in which words are associated with which genders.

Another issue in language processing and gender bias is that English speakers, when using personal pronouns, seem to have a bias towards using male personal pronouns such as ‘he’ when referring to what should be a gender-neutral description (Moulton et al., 1978). Furthermore, it has been explored by Carreiras, Garnham, Oakhill and Cain (1996) how general knowledge about the stereotypical gender of participants in a text influences comprehension. They looked at reading times for sentences following another sentence that introduced a stereotypical male/female character: moreover, they looked at sentences in Spanish where the gender of the constituent could be disambiguated by the gender of a preceding article *la/el* (which is not the case in English). Here, they found that reading times increased when there was a mismatch between the gender of the article and the gender of the constituent in the text. This suggests that the stereotypical gender of the person described in a text affects the ease with which the text is understood. Furthermore, it is proposed that this happens as early as it becomes possible for the reader, and that it facilitates understanding of the meaning of a sentence. This might be because dogmatism of genders stem from the difference in occupations held by men and women throughout time, and because traits that are used to describe people have not changed as much as our ever-changing societal norms (Wood & Eagly, 2002). Newer research has suggested otherwise (Ahlgreen, 2019). Possible explanations for the contradicting results could be that the two studies were done in different languages with

different cultures, and that they are done with 23 years apart. Both these explanations entail that language is influential on cognition.

## The Influence of Dogmatisms

We suggest that these societal dogmatisms exert an influence on our perception of adjectives and the perception of the gender of the syntactical subject described with adjectives in the Danish language. To investigate this, we used a twofold experimental approach. On the basis of gender stereotypical heuristics, we firstly hypothesized that people in the Danish language have a gender-biased prediction towards adjectives. This prediction could, potentially on the basis of dogmatisms in our society and minimization of prediction error in top-down processes, hold an implicit gender bias. Secondly, we hypothesized that adjectives from the Danish language can affect our interpretation of a sentence containing a gender-neutral subject (e.g. *'The person'*). Thereby also altering the understanding of the neutral subject's unspecified gender accordingly with the average 'gender charge' of the sentence.

We build our hypotheses on the assumption that adjectives hold some kind of implicit gender charge, with females being described with more collaborative and communal adjectives and men being described with more agentic adjectives. As the genders are assigned different qualities, different adjectives will be used more frequently to describe one gender over the other. Thus, creating prejudicial connotations to certain adjectives. We were inspired by Madera, Hebl, and Martin (2009), who suggested that the gender of the person described in letters of recommendation could significantly predict different types of adjectives. Meaning that the use of communal adjectives would be used more often in letters of recommendation about women. This might be a product of our societal beliefs about gender roles described above; therefore, we took to investigate if this bias exists in the Danish language as well. We proposed that, when asking people if they rate adjectives of either gender as either descriptive of a male or of a female, we would reveal some kind of gender bias in the Danish language, possibly with origins from societal dogmatisms. This was done by finding the mean ratings and standard deviations of 72 Danish adjectives rated on a scale from -5 to +5 with -5 being the most feminine an adjective could get and +5 being the most masculine an adjective could get. The words were rated by a group of Danish speakers. Moreover, we also investigated whether these ratings could significantly predict the categorization of gender-neutral, grammatical subjects of sentences. To test this second hypothesis, we used the ratings from the first part of the experiment to form sentences that we had a new group of participants read. These participants were then asked to categorize the neutral subject of a sentence as either a male or a female.

## Methodological Approach

### *Participants 1*

One-hundred and fifty-seven participants (53.09% female; age:  $M = 24$ ,  $SD = 9.2$ ) rated 72 Danish adjectives. All of the participants were fluent Danish speakers.

### *Procedure 1*

We asked the participants to rate 72 Danish adjectives on a scale from -5 to +5. A rating of -5 was characterized as having a *very feminine charge* and representing the most female-charged a word could get. A rating of +5 was characterized as having a *very masculine charge* and representing the most male-charged a word could get. A rating of 0 was characterized as being of *neutral charge*. The participants rated the adjectives online through a Google Surveys questionnaire. We chose the 72 different adjectives presented in the questionnaire ourselves, using a list of the most frequently used adjectives of the Danish language (Basby, 2017) as a starting point. From that list we picked 24 adjectives that we expected would be associated with feminine features, and 24 adjectives that we expected would be associated with masculine features. Lastly, we picked 24 adjectives that we expected to be gender neutral.

Strong \*

-5 (Very female)

-4

-3

-2

-1

0 (Neutral)

1

2

3

4

5 (Very male)

Figure 1. An example of the stimuli participants were presented with in the first experiment

The participants read a short introduction in Google Surveys asking them to rate the following 72 adjectives from -5 to 5 with the possibility of choosing every integer in between. They were then told to rate adjectives they associated as very female with -5, neutral adjectives with 0, and very male adjectives with 5. They were then presented with one word at a time until they had rated all 72 adjectives. All 157 participants rated all 72 adjectives.

## *Participants 2*

We had a new group of 150 participants, who were all participants who had not participated in the first experiment (74% female; age:  $M = 20.04$ ,  $SD = 7.88$ ) categorizing 44 different sentences. All of our participants were fluent Danish speakers. 85% of them were high school students of psychology from Munkensdam Gymnasium (Kolding, Denmark). However, in order to get a more balanced gender distribution, we had to outsource our questions to male people. The participants were all assigned to answer the same questions purposely, thereby creating a repeated measures design.

## *Procedure 2*

The stimuli presented to our participants in our second study were 42 Danish written sentences. In all sentences, the subject was one of four gender-neutral constituents (approximate translation from Danish):

*S1: the person*

*S2: the figure*

*S3: the shadow*

*S4: the silhouette*

The syntactical subject of each sentence was associated with three adjectives from the first part of our experiment. An example of a random sentence with a mean rating of  $-1.84$  for all three adjectives in a sentence (approximate translation from Danish):

*S1: The person was beautiful, sexual, and scared*

The participants were then asked to answer whether they thought the gender-neutral subject in the sentence was a man or a woman. Participants were forced to answer this dichotomously and could not choose not to answer. The participants answered whether they thought the subjects in the sentences were men/women through an online questionnaire.

We had nine different conditions of sentences that we tested. Words which on average were rated between  $-0.35$  and  $0.35$  were characterized as being neutral adjectives. Words which on average were rated above  $0.5$  were characterized as having a male gender charge. We will henceforth refer to the mean gender charge of a sentence as GC. Words which on average were rated with a gender charge below  $-0.5$  were characterized as having a female GC. We made an arbitrary cut-off and excluded words that had a rating between  $0.5$  and  $-0.35$  or words that had a rating between  $0.35$  and  $0.5$  in order to increase the distance between the three categories.

We composed  $3 \times 3$  sentences where the three adjectives were solely characterized as being mostly female ( $GC < -0.5$ ), male ( $GC > 0.5$ ), or neutral ( $GC > -0.35$  and  $< 0.35$ ), thus creating a total of

nine sentences. Furthermore, we created six conditions, in which the adjectives were varied in terms of charge:

**C1:** 2 neutral adjectives: ( $GC > -0.35$  and  $< 0.35$ ), 1 male, adjective ( $GC > 0.5$ )

**C2:** 2 neutral adjectives: ( $GC > -0.35$  and  $< 0.35$ ), 1 female adjective ( $GC < -0.5$ )

**C3:** 2 male adjectives ( $GC > 0.5$ ), 1 female adjective ( $GC < -0.5$ )

**C4:** 2 male adjectives ( $GC > 0.5$ ), 1 neutral adjective ( $GC > -0.35$  and  $< 0.35$ )

**C5:** 2 female adjectives ( $GC < -0.5$ ), 1 male adjective ( $GC > 0.5$ )

**C6:** 2 female adjectives ( $GC < -0.5$ ), 1 neutral adjective ( $GC > -0.35$  and  $< 0.35$ )

The experimental design was created in this way such that the GC of our different sentences were as representative and diverse as possible.

Besides answering the above-described sentences by assigning the subject of these sentences a gender, we created some distraction tasks within the questionnaire. This was done to prevent our participants from recognizing that the experimental design sought to predict gender from adjectives. To do this we created sentences that used the same syntactic structure with an introduction of a gender neutral, Danish, syntactical subject described by three adjectives (in these sentences the adjectives were randomly chosen). The deflection task contained two different conditions. Participants were either asked to decide if the subject in question was a hero/villain or a member of the Danish parliament/in prison. Two examples of our deflective conditions are given below: C1 is an example of a sentence that was structured to create an understanding of a subject as being in prison/parliament and C2 is an example of a sentence that was structured to create an understanding of subject as being a hero/villain (approximate translation from Danish):

**C1:** The silhouette was manipulative, rhetoric and unethical

**C2:** The silhouette was heavy, big and strong

Our participants were exposed to 10 deflecting sentences with the *C1* structure *and* exposed to 8 deflecting sentences structured as *C2*.

## Data Analysis

The purpose of our data analysis is two-fold. Firstly, to demonstrate that people have a shared implicit gender charge. Secondly, to demonstrate how implicit gender charge determines how we perceive seemingly gender-neutral, syntactical subjects (a reminder for the reader that the mean

ratings of the words are referred to as “GC’s”). It is necessary to go through statistical operations to make groundings for our hypotheses.

To begin with, we will present how we determined how each adjective score was on a gender scale from -5 to 5. The purpose of this is to place different adjectives on a gender scale as objectively as possible. In this way, we will be approaching a GC that represents the participant’s implicit GC as closely as possible. This will include an analysis of the level of agreement between our individual participants. This is done by looking at a statistical measure called the standard error of the mean that explains how much people’s answers deviate from the average GC found in the experiment. Another statistical measure used to explain how much the individual participant’s answers deviated from the GC was doing a mixed-effects model. A model that is used to explain the relationship between the average GC of each word and the ratings of each participant. The benefit from doing a mixed-effects model is that you can detect general tendencies from participants. Some might be more conservative in their ratings and some might be more polarized. However, by ruling out individual biases we were able to get a more statistically clear measure of the GC. In this way the model takes into account the individual participants and their respective biases. (For more information about the technicalities of our model, see appendix). Overall, this means that we are not just claiming that people agree on the GC’s found, but we statistically validate the level of agreement of them.

Furthermore, the purpose of our data analysis is to show that by altering the mean GC in a sentence (the mean of 3 GC’s) that describes a gender-neutral subject, we can see if a lower or higher GC will affect whether the syntactical subject is deemed a male or a female by our participants. This was done by getting the different participants to answer whether they deemed the syntactical subject a male or a female when being presented with different sentences with different mean GC’s. Hereafter, a statistical model was fit to explain the relationship between the sentence GC and the gender chosen for the presumable gender-neutral, syntactical subject. We hypothesized that when the average GC of a sentence was  $< 0$ , the syntactical subject would be deemed female more often and vice versa. Again, we wanted this model to be able to account for individual differences to take into account that some people might have skewed responses answering one gender more often than the other. Only the sentences where the participants were asked to deem the subject a man/woman were used for analysis. Choices of the gender of the persons in the sentences were recorded in the questionnaire through Google Surveys. Data analysis and statistical models were in R, a programming language for statistical computing and graphics. (For more information on our analyses, see appendix).

### *Results 1*

As explained in the previous section, the goal of this result section is to determine that people have a shared implicit GC. Furthermore, to investigate the relationship between this shared implicit GC and people’s individual ratings of the words. Simply put, do people agree on the rating scale of the adjectives provided in the survey?



The first goal of this results section was achieved by taking the means of all of the ratings of each adjective. This gave us the shared implicit GC for each adjective. When we did this, we also made statistical calculations to determine how much the implicit gender charge for each adjective deviated between the individual participants. These implicit GC's and the overall variation in the ratings between the participants were calculated for all of the 72 adjectives selected for our study. A visual representation of the adjectives and the calculated statistics are given in Figure 2. The figure should be interpreted as the x-axis representing the GC's and the y-axis as having each word on it. This gives a clear overview of different adjectives and their gender charge in descending order.

When investigating the relationship between the individual participant's ratings and the GC's determined, we found that if people's ratings of a word decreased so would the average GC of a word and vice versa. This means that there was a tendency for people to agree on the ratings of the different words, thus creating a stronger ground for claiming that people agree on these ratings and that the adjectives in themselves do indeed contain an implicit gender-bias. We did this by running a statistical model that can account for the fact that individual people's ratings might be conservative or polarized. This was done as we assumed that people might have some individual biases towards some adjectives. Such biases are visualized in Figure 3 where the GC of the adjectives are plotted on the x-axis against each participant's ratings on the y-axis. Each line/color represents each person's baseline of ratings. For this reason, some lines are more horizontal (meaning they are more conservative in their ratings), and some are more diagonal (meaning their ratings are more polarized). We ran a statistical model that looked as such: (GC ~ Participants Rating + (1|Subject):  $\beta = .99$ , SE = 0.01,  $t = 85.266$ ,  $p < .001$ ). Our statistical model shows that our participants generally agreed on the ratings of the adjectives, and therefore that the mean GC is a good measure for implicit gender-charge across all participants. (For details about a likelihood ratio test that further validate these findings, see appendix).

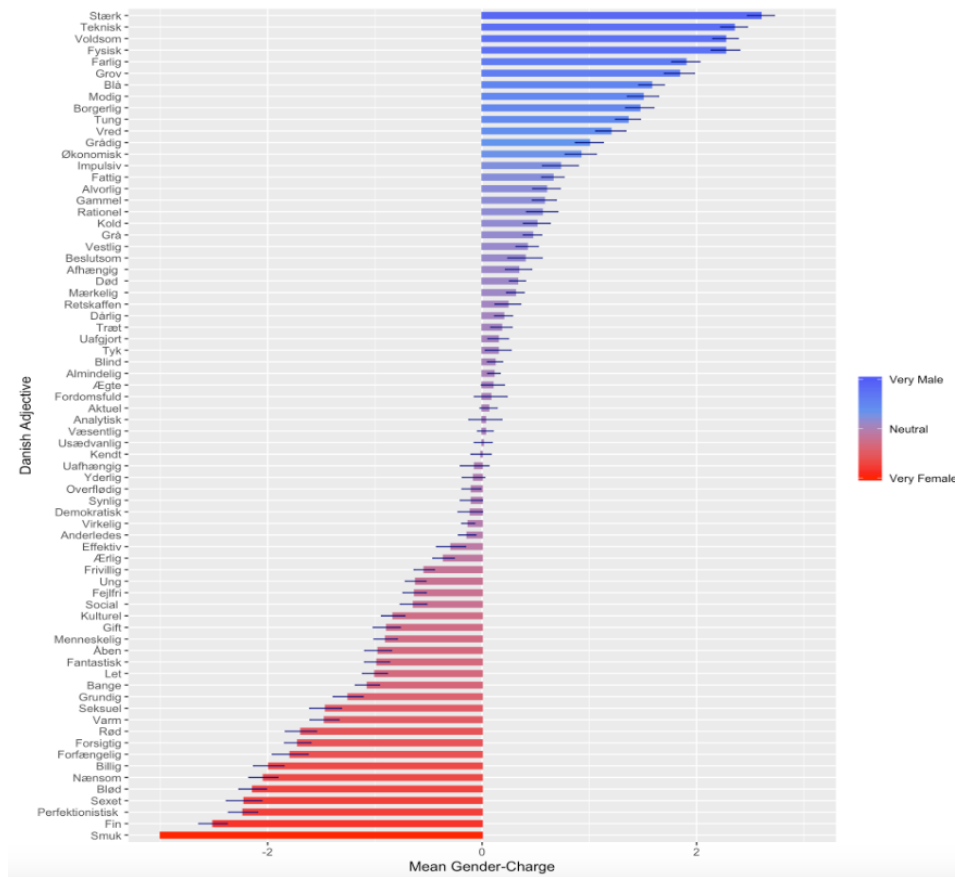


Figure 2. The mean ratings of the 72 Danish adjectives on a scale from -5 (most female) to 5 (most male). Error bars are added with the standard error of the mean.

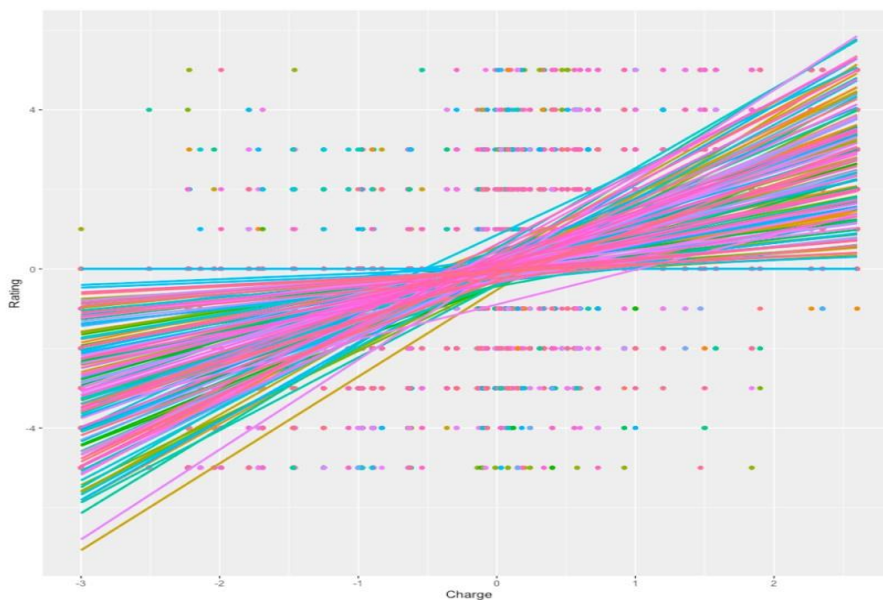


Figure 3. The average gender charge of each adjective predicted by the rating of each word by each of the participants. Different colors represent the ratings for each single participant. It shows that some people are more conservative in their ratings, and some are more polarized.

## Results 2

The second results section revolves around how we use GC's found in "Results 1" to explain which gender people think the gender-neutral subject of a sentence is when we alter the mean GC of a sentence (i.e. when we replace adjectives with certain values of GC with other adjectives that have different scores). Therefore, we looked at the results from the second questionnaire and started by giving sentences an overall sentence GC. A sentence GC was created by adding 3 adjectives to, what we believed to be, a gender-neutral subject, and the average GC was calculated by taking the mean of 3 GC's in one sentence. As we had hoped, the GC's of the sentences followed our participants' choices of the gender of the person being described in a sentence. The effect of our average GC as a function of the sentence charge was statistically significant. The statistical model looked as such: (Binary Outcome ~ Sentence Charge + (1|Stimuli):  $\beta = 1.64$ ,  $SE = 0.16$ ,  $z = 10.16$ ,  $p < .001$ ) suggesting that the gender charge of the sentence had an overall effect on which gender our participants characterized the gender-neutral subject as. That is, when the sentence gender charge was  $> 0$ , people would significantly more often determine the gender-neutral subject of this sentence as a male and vice versa.

Furthermore, our analysis showed that the more negative the sentence gender charge becomes, the more likely it is that the participants will categorize the gender-neutral, syntactical subject as a *woman*. Figure 4 is a visualization of these results, where the average gender charge of a sentence is plotted on the x-axis against the choice each participant made of the gender-neutral, syntactical subject as being either male (1) or female (0). Additionally, we investigated whether the gender of the participant making the choice would influence the choice. Here we found that the gender of the participants themselves did not influence which gender they thought the syntactical subjects in our sentences were. (For more details about the calculations made to get these results, see appendix).

The results of the second part of our experiment were obtained through a forced-answer format. This might have influenced the overwhelmingly harmonious results we got. Furthermore, we constructed the sentences from an assumption that the four chosen, Danish, grammatical subjects were gender-neutral (*personen (the person)*, *skyggen (the shadow)*, *silhuetten (the silhouette)*, *figuren (the figure)* (translations in parentheses)). Thereby our experimental settings do not take into account any implicit GC these words might have.

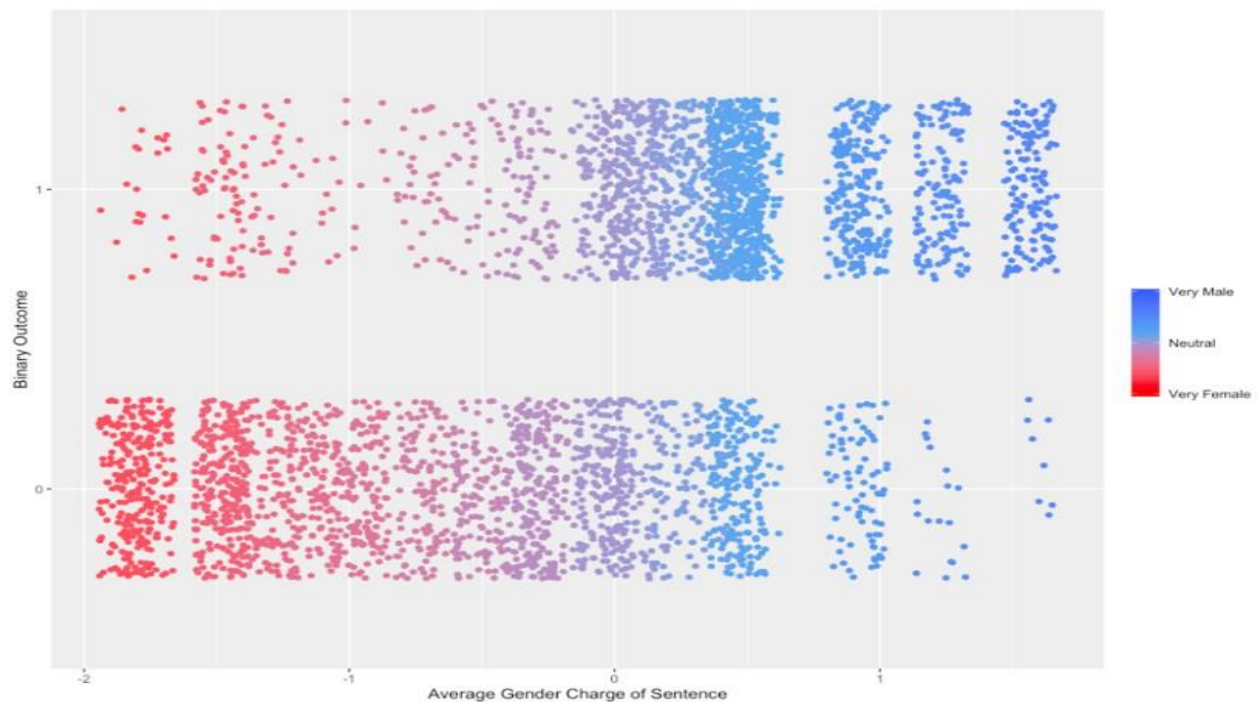


Figure 4. The distribution of answers to whether a neutral subject of a sentence is categorized as male or female given the average GC of a sentence

## Discussion

### *Prejudice Mediated by Danish Adjectives*

The experiments in this study were designed to investigate if implicit gender charge in Danish adjectives bias the way people think about gender-neutral, grammatical subjects. In the first study we explored if some of the most common adjectives in the Danish language (e.g. “beautiful”, “strong”, “ordinary”) contained a bias towards either gender. Participants in our study were simply presented with the adjective and asked to rate its gender charge. This provided us with an average gender charge for each adjective as well as a standard deviation and standard error of the mean for each of them. Furthermore, it allowed us insight into the relationship between the means of each word and every participant’s rating of this word. We hypothesized that people would have a pre-understood interpretation of the adjectives presented. This interpretation being that some words contain an implicit gender charge, and that some do not contain a gender charge. We believed that this interpretation of some adjectives as being more descriptive of one or the other gender may stem from societal dogmatism and general beliefs about genders in our society. Participants were presented with a number of the most used adjectives in the Danish language that they had to rate on a most-female to most-male scale. This allowed us to look at general tendencies and look at individual adjectives that were more charged than others.

The adjectives and their average gender charge (GC) were then used to create sentences for our second experiment. This was presented for a new group of participants to read and answer. The

sentences all had, what we believed to be, a gender-neutral, syntactical subject that was described by the charged adjectives from the first part of the experiment. We used the average GC of a word to create an average GC of a sentence. We believed that this average sentence GC of the three adjectives used to describe the gender-neutral subject of the sentence, should be able to alter our participants' understanding of the gender of the syntactical subject. We hypothesized that these mental representations of the neutral subject would be influenced by the different connotations we have with different adjectives that may stem from societal dogmatism, and that the average sentence GC could significantly predict what gender our participants would ascribe the grammatical subject of the sentence.

The findings of this study provided initial support for our first hypothesis with fairly unilateral ratings of the Danish adjectives. There was a significant tendency between the average GC of each adjective acquired from our participants and the individual participants' ratings of each adjective. This suggests that some adjectives in the Danish language have a prejudiced gender bias and contain a male, female, or neutral charge. This was validated in our results, as there was a linear relationship between GC and people's ratings of the adjectives. Moreover, our study provided support for our second hypothesis as we found that the average gender charge of a sentence had a significant effect on people's choice of characterizing a gender-neutral subject of a sentence. Therefore, our results could suggest that adjectives can alter people's predestined understanding of a gender-neutral subject as being of a certain gender. In general, people were unanimously inclined to ascribe the same genders to the same grammatical subjects when described with our chosen adjectives. We found no significant effect of the gender of our subjects as being able to significantly predict the gender-neutral subject. This could be because we lacked statistical power,  $M = 26\%$ . However, it could also be because we in general had very unanimous answers to the second experiment. This would leave the difference in gender of our participants with very little variance in the data left to explain.

### *Methodological Limitations and Implications*

One thing that could have affected our participants in their ratings of words is the way we approached the dichotomous division of the GC where the positive numbers were attributed to men and the negative numbers were attributed to women. This might have led people to think that we favored men over women, and people could have a bias towards positive numbers being better than negative ones.

Furthermore, the results of the second part of our experiment were obtained through a forced-answer format. This might have influenced the overwhelmingly harmonious results we got. Additionally, we constructed the sentences from an assumption that the four chosen, Danish, grammatical subjects were gender-neutral (*personen, skyggen, silhuetten, skikkelsen*). Thereby our experimental settings do not take into account any implicit GC these words might have.

The fact that the results in the second part of the experiment were obtained through a forced-answer format means that we have no way of knowing whether people have formed an understanding of a person as being of any type of gender before being asked about it. People might only access societal

dogmatism and stereotypes when forced to take a position on the matter. However, Lassonde & O'Brien (2013) did a study where results could suggest that people actually do construct gender representations of characters very early in their thought process. In their experiments, participants were presented with small passages of text describing an occupational character with gender-neutral language (e.g. a firefighter instead of fireman). Then they would go on to read target sentences that could contain a personal pronoun that contradicted the stereotypical expectation that a firefighter would be a man (e.g. refer back to the firefighter as *she*). They found that reading times for the target sentences significantly increased when the characters described in them were women. Furthermore, Blumentritt & Heredia (2005) presents the idea that social stereotypes could be viewed as cognitive structures (*schemas*) that hold networks of abstract information about traits, attributes, and expected behaviors of a member of a particular societal group. These two studies with different methodological approaches present the idea that we have created biased notions as how to describe different genders and that when this bias is contradicted in a sentence, we use longer time to process the sentence.

Our findings add to this idea as our participants agreed more or less unanimously with which genders the gender-neutral subjects were. Our paper suggests a robust, scientific ground for our hypothesis that Danish adjectives can manipulate the way people understand gender-neutral, grammatical subjects. This is done on the basis of the literature reviewed and on our study with significant results both with fairly unanimous categorizations of the GC in adjectives and a fairly high agreement in people ascribing the same gender to the grammatical subjects in our sentences in the different conditions.

### *Heuristics Dictating our Communication; are we Restricted or Effective?*

Furthermore, our study suggests that the gender-charged adjectives did invoke some sort of predestined gender bias in our participants. Our findings propose that Danish speakers rely on heuristics to predict what gender certain adjectives are likely to describe and thus indicate that Danish speakers communicate with an expectation of set, implicit meanings about genders in adjectives. This leads to an indication that heuristics shape the way Danish speakers process adjectives.

If we take a standing point in top-down processes being essential for overall sentence comprehension, like we see in the word superiority effect, this heuristic shortcut would have some kind of effect on our decision-making when it comes to making fast generalizations of the gender-neutral subject in a sentence. This is out of the scope of our study, and we are therefore not able to make inferences about how top-down processes influence language processing.

However, the fact that this a gender-neutral subject could be predicted to such a degree as shown in our study may speak on behalf of societal beliefs being able to, in some way, alter our understanding of how we view different genders. It seems as if we look at a man as *strong*, *economic*, and *physical* and at a woman as *soft*, *sexy*, and *careful*. It should also be mentioned that this might not be a product of individuals purposely following dogmatism as it is hard to designate at which time in a person's life this bias occurs. Therefore, it makes it hard to determine if this bias

occurs on the basis of a person following dogmatism or if this is a matter of keeping prediction error low because we have previous experiences with people who use adjectives surrounding certain genders. Therefore, it might be an adaptation from previous generations and something that would change in accordance with societal roles.

Our results go hand in hand with Eagly & Wood's (2002) biosocial model of sex differences, in which they compared patterns of gender behavior with description differences of gender in language. People seem to describe different genders with different words. Our experiments infer that it is very plausible that Danish speakers are biased in which adjectives they use to describe certain genders in written language. Furthermore, our results also make good explanatory sense when looking at Eagly & Wood's (2002) anthropological review of gender differences. Here men have through time held occupations that demand leadership skills and physical advantages. Women on the other hand are described with more caring traits as they were required to be nursing and expected to raise children. Our results align with this anthropological explanation considering that words such as (*Strong, Technical and Physical*) were some of the words with highest values of GC, and words such as (*Thorough, Warm and Careful*) were some of the words with the lowest value of GC.

Our methodological approach prevents us from explaining this phenomenon from an anthropological/cognitive viewpoint, in terms of; why might this be? However, we can comment on the fact that the conclusions are somewhat in agreement.

### *Further Research*

If our study were to be built upon, the next thing to investigate would be to explore the design in an experiment with a non-forced answer format. Our investigation has experiment limitations when it comes to retrieving answers from our participants. We have no way of knowing if people had actually formed a gender-image of the subjects in our sentences before we asked them to decide if the sentences described a woman or a man. Instead we are limited to investigate gender-stereotypes. Our experimental findings are therefore also of limited significance in regard to concluding anything about language processing through gender bias. They can only suggest that language is processed with a preliminary, heuristical reliance on gender bias.

Furthermore, our experiments were constructed such that there were no real answers. Even though most people agreed on what genders the gender-neutral subjects were, all of the sentences were constructed for the purpose of these experiments. A possible new take on these experiments could be to construct them such that they utilized sentences from Danish literature (e.g. novels, newspaper articles, magazines). These texts could give a more nuanced look into how adjectives are used to describe different genders from a public viewpoint in literature.

Moreover, our experiments were solely focused on written language. Further investigation could involve experiments where it was tested if people have the same gender-biased approach to Danish adjectives in spoken Danish. In such experiments, researchers could take a starting point in Jaeger & Aslin's (2012) proposition that predictions are not exclusive to language processes because they

seem to be an attribute to our understanding of a diverse world in general. Researchers could look to see if predictions are universal to all of our perception systems, and if that means that our understanding of every sensory input we have is grounded in some sort of categorization process.

In this process, it could be investigated how we access heuristics that act as cognitive shortcuts for communicative processes.

## Conclusion

Danish adjectives contain an implicit gender-charge that match gender stereotypes. Our study indicates that 72 of the most common adjectives in the Danish language (e.g. “beautiful”, “strong”, “ordinary”) contain a bias towards gender. Furthermore, it allows an insight in how these adjectives can alter the interpretation of the gender of a neutral, syntactical subject in a sentence. The results provide a supportive view for the social role theory and previous findings on how our perception of different genders alter our understanding of language.

## References

- Ahlgren, L. M. (2019). Om kønnet sprogs indvirken på stereotype kønsopfattelser. *Journal of Language Works - Sprogvidenskabeligt Studentertidsskrift*, 4(2), 5–22.
- Anderson, J. R. (2000). *Cognitive psychology and its implications*, 5th ed. (pp. xvi, 531). Worth Publishers.
- Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language*, 62(1), 67–82. <https://doi.org/10.1016/j.jml.2009.09.005>
- Basby, A. (n.d.). *Talende Adjektivliste*. Retrieved 28 May 2020, from <http://basby.dk/adjektivliste/adj.html>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using **lme4**. *Journal of Statistical Software*, 67(1). <https://doi.org/10.18637/jss.v067.i01>
- Blumentritt, T. L., & Heredia, R. R. (2005). Stereotype Processing and Nonliteral Language. In *Figurative language comprehension: Social and cultural influences*. (pp. 261–281). Lawrence Erlbaum Associates Publishers.
- Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2004). When Professionals Become Mothers, Warmth Doesn't Cut the Ice. *Journal of Social Issues*, 60(4), 701–718. <https://doi.org/10.1111/j.0022-4537.2004.00381.x>
- Eagly, A. H. (1987). Reporting sex differences. *American Psychologist*, 42(7), 756–757. <https://doi.org/10.1037/0003-066X.42.7.755>
- Eagly, A. H., Wood, W., & Diekmann, A. B. (2000). Social role theory of sex differences and similarities: A current appraisal. In *The developmental social psychology of gender*. (pp. 123–174). Lawrence Erlbaum Associates Publishers.
- Ferreira, F., & Chantavarin, S. (2018). Integration and Prediction in Language Processing: A Synthesis of Old and New. *Current Directions in Psychological Science*, 27(6), 443–448. <https://doi.org/10.1177/0963721418794491>
- Fitch, W. T., Huber, L., & Bugnyar, T. (2010). Social Cognition and the Evolution of Language: Constructing Cognitive Phylogenies. *Neuron*, 65(6), 795–814. <https://doi.org/10.1016/j.neuron.2010.03.011>



- Frank, S. L. (2013). Uncertainty reduction as a measure of cognitive load in sentence comprehension. *Topics in Cognitive Science*, 5(3), 475–494. <https://doi.org/10.1111/tops.12025>
- Kleinschmidt, D. F., & Jaeger, T. F. (2015). Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review*, 122(2), 148–203. <https://doi.org/10.1037/a0038695>
- Koenig, A. M., & Eagly, A. H. (2014). Evidence for the social role theory of stereotype content: Observations of groups' roles shape stereotypes. *Journal of Personality and Social Psychology*, 107(3), 371–392. <https://doi.org/10.1037/a0037215>
- Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, 31(1), 32–59. <https://doi.org/10.1080/23273798.2015.1102299>
- Lassonde, K. A., & O'Brien, E. J. (2013). Occupational stereotypes: Activation of male bias in a gender-neutral world: Occupational stereotypes. *Journal of Applied Social Psychology*, 43(2), 387–396. <https://doi.org/10.1111/j.1559-1816.2013.01008.x>
- Lippmann, W. (1965). *Public opinion*. Free Press; /z-wcorg/.
- Madera, J. M., Hebl, M. R., & Martin, R. C. (2009). Gender and letters of recommendation for academia: Agentic and communal differences. *Journal of Applied Psychology*, 94(6), 1591–1599. <https://doi.org/10.1037/a0016539>
- Moulton, J., Robinson, G. M., & Elias, C. (1978). Sex bias in language use: 'Neutral' pronouns that aren't. *American Psychologist*, 33(11), 1032–1036. <https://doi.org/10.1037/0003-066X.33.11.1032>
- Qian, T., Jaeger, T. F., & Aslin, R. N. (2012). Learning to Represent a Multi-Context Environment: More than Detecting Changes. *Frontiers in Psychology*, 3. <https://doi.org/10.3389/fpsyg.2012.00228>
- R Core Team (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
- Rosch, E. H. (1973). Natural categories. *Cognitive Psychology*, 4(3), 328–350. [https://doi.org/10.1016/0010-0285\(73\)90017-0](https://doi.org/10.1016/0010-0285(73)90017-0)
- Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology*, 83(2), 176–190. <https://doi.org/10.1016/j.ijpsycho.2011.09.015>
- Whorf, B. L. (1956). *Language, thought, and reality: Selected writings*. Cambridge, Mass.
- Wood, W., & Eagly, A. H. (2002). A cross-cultural analysis of the behavior of women and men: Implications for the origins of sex differences. *Psychological Bulletin*, 128(5), 699–727. <https://doi.org/10.1037/0033-2909.128.5.699>

## Appendix

The statistical tests and visualization of data can be found in the GitHub-repository below. It furthermore contains information about packages and analysis tools used:

<https://github.com/Lassehhansen/Cognition-Communication/blob/master/R%20Work/An%20analysis%20of%20Gender%20Charge%20in%20danish%20adjectives.Rmd>