Analyzing gender bias in children's television shows

Name: Madina Habib

Student number: 5784182

Supervisors: Joris Veerbeek, Dr. Max Kemman, Dr. Mirko Schaefer

Date: 09-08-2021

Abstract

In this research gender bias was analyzed in two of the most watched children's television shows in the Netherlands, *Sesamstraat* and *Het Klokhuis,* due to the impressionability of the target audience of these shows. Automated speech recognition has been used to generate a textual corpus of 20 episodes per year per show in the time period from 2011 until 2020.

In this research, bias is defined as the existence of prejudice towards certain (groups) of people. The (groups of) people this paper focuses on, are distinct in gender, male and female. A model calculating word distances (Word2Vec) was used to calculate the distance between words from six LIWC categories and two lists containing male or female related words.

Using linear regression, the change over time in gender bias in each LIWC category was calculated per year to show in what aspects gender bias has changed in these two television shows over the last decade.

Aside from the LIWC categories 'Groom' and 'School' in which a slight decrease in the slight bias towards women was found, no trend was found in the other categories. This could mean several things, for example 1) there is no trend in gender bias in these shows over the last decade, or 2) the corpus was of insufficient size to get proper results.

The limitations of this research make the research valuable as a steppingstone towards further research into the field of gender bias related to children's television shows.

Keywords: gender bias, Word2Vec, word distances, children's television shows

Analyzing gender bias in children's television shows

The first and longest running educational children's television show that is broadcasted worldwide is Sesame Street. With the success and proven results of Sesame Street, more educational shows for children have since been produced. Many children's shows aim to be educational, be it on socialization, teaching aspects on culture, teaching skills directly related to school, or a combination of these things. Teaching through television shows is a great way to reach children who do not have the privilege of being able to be (properly) taught these things by their parents or caregivers (Cooney, 1967).

However, besides the direct learning goals of these shows, the makers of these educational television shows can unconsciously form the show in a way where their own unconscious bias seeps through into the show. Being biased is human nature, it is impossible to have no biases, which is why an impressionable audience like children must be protected from (in)directly being taught these biases. In this research, two major Dutch television shows (*Sesamstraat* and *Het Klokhuis*) have been analyzed on how gender bias has changed in these shows in the last decade using a model to calculate word distances (Word2Vec) based on automatically generated subtitles for these television shows. An elaborate description of how this analysis has taken place can be found in the 'Methods' section. While these two shows are far from the complete television consumption done by children, they are two of the most watched and accessible children's shows on Dutch television, which is why it is interesting to research bias in these two shows specifically.

One of the causes of bias is social categorization. Research has found that people categorize each other based on their (outward) characteristics, judging, and forming opinions based on what group they think a person belongs to, while favoring the group they themselves feel like they belong to (Wilder, 1986). One of the consequences of social categorization is the association of certain characteristics with certain groups of people.

These characteristics can be innocent in the form of categorizing a person as male or female, or categorizing a person as an old person or a young person. This social categorization can be beneficial in certain situations, for example when looking for a certain product in a store, approaching a person who looks like they work at the store. However, social categorization can also lead to bias when the association is an overgeneralization of a group of people, leading to all members of this group to be treated differently. This can cause damage to these groups of people in many ways, it can lead to discrimination, racism, and sexism amongst other harmful consequences.

So, when a children's educational show is made, what if unconscious social categorization by the makers of the show slips through into the show? A single event where this happens does not necessarily have to be picked up on by the viewer, however, consistent exposure to a certain bias can lead to the bias being learned by the person watching the show through inductive generalization (Yu, 2021).

Inductive generalization is when something is learned about something, which is then extended onto other things from the same (semantic) category (Yu, 2021). For example, a young child reaches into a flame and burns his hand, through inductive generalization he learns that all flames will burn him, because of his previous experience with a flame. In the case of gender bias in television shows, this could for example go as follows: A young girl is watching her favorite tv show, in this show she hears one of the main characters say that his sister is bad at math because she is blonde. The young girl watching the show could then inductively generalize this as all blondes being bad at math, because the sister of the main character is also bad at math and the given explanation was of her being blonde. Of course, it could also be the case that the sister is portrayed as being bad at math, because she actually is bad at math. So, when does this become bias? Again, an association that gets consistent exposure resulting into an overgeneralization of a group of people which could result in different treatment of that group, is when it becomes bias.

An example of gender bias that can lead to damaging consequences in children, is little boys frequently being exposed to the notion that "boys don't cry". Research by Vogel et al (2011) suggests that notions like "boys don't cry" and conforming to these notions by men, is linked to men having unfavorable attitudes towards seeking (professional) help when struggling with mental health issues. Combining these findings with the fact that over three quarters of suicides are by men (Office for National Statistics, 2017) and suicide being the biggest cause of death for men under 35 (Office for National Statistics, 2017), while men are less likely to go to therapy (NHS, 2017) raises serious concerns.

Another example of unwanted consequences due to gender bias, is the shift from psychology being a field dominated by men to psychology being a field dominated by women. This shift has caused a decrease in the so-called prestige of the field, while the quality of work has not decreased (Philipson, 1993).

Besides these general damaging consequences of gender bias, the Geena Davis Institute in collaboration with the J. Walter Thompson Company (2016) has researched movies and television shows and the positive influence it can have on women. This research has found that 61% of women reported that female role models in shows and movies have played an influential role in their lives and 58% reported that these female role models from shows and movies have inspired them to be more ambitious and assertive. From this same research, it was found that while women make up 40% of the global workforce, in television and movies this percentage was only 25%. The same for women in political positions, in television and movies women are seen in political positions less than 10%, while the number for the real political positions is 24%. So, if women can be influenced by the character traits of female role models in media, then the same is possible for children.

So, while television shows are definitely not the only potentially contributing factor to the creation and existence of these kinds of biases, television shows certainly do play a role. Especially shows that are deemed educational and are often used in the school curriculum as well.

Literature review

Discrimination in the workplace, unequal pay, unequal opportunities, underestimation and being undervalued. These are only some of the negative effects (unconscious) gender bias has in society. To understand how the previously mentioned effects happen because of gender bias, it is first important to understand what gender bias is exactly.

In this research, bias is defined as the existence of prejudice towards certain (groups) of people. The (groups of) people this paper focuses on, are distinct in gender. Though there are many gender identities, this research will only distinguish between male and female when discussing gender. Thus, gender bias, in this research, is defined as the existence of prejudice towards males or females. This bias does not have to be on purpose, biases are often ingrained in one's way of thinking, unconsciously being used in day-to-day life. This can also be called implicit bias. There is also conscious bias, which is when a person is aware of their opinion on a group of people and lets their opinion consciously change their behavior (Gaddes et al, 2021). For example,
people who believe a woman should not be president and thus they do not vote for the female presidential candidate because of this opinion.

An example of implicit bias that has been investigated by Cao and Banaji (2016), is the finding that most people associate the word 'doctor' with men more often than women, and the word 'nurse' is associated more often with women than with men. Cao and Banaji (2016) also found the same stereotypic association when it comes to the words 'scientist' which was mostly associated with men, and 'artist', which was mostly associated with women. Many more examples of implicit biases like these can be thought of that are present in society.

Understanding the existence of such biases, raises the question of where these prejudices that form the bias, come from. The director of the Prejudice and Intergroup Relations Lab, Professor Devine (1989) has found that implicit bias is formed due to repeated exposure to stereotypes and is activated when a person is in the presence of someone or something that the person associates with the stereotyped group. Reusing the previously mentioned example of the word 'artist', the gender ratio for male and female artists in the United States of America where this research was conducted, is 54% male artists to 46% female artists (National Endowment for the Arts, 2008). How is it then possible that the participants in the research are associate the word artist more with women than with men? It all comes back to exposure (Devine, 1989). One of the many ways people are exposed to stereotypes is through media.

Previous research in social, behavioral, and developmental science has already shown that media such as television shows aimed at children influence their consumption (Smith et al, 2019), their stereotype endorsement (Wille et al, 2018), and their behavior (Kostyrka-Allchorne, Cooper & Simpson, 2017).

There are many forms of media: radio, television, newspapers, et cetera. All these forms of media have one main goal, which is to inform. Inform listeners, watchers, or readers of current events all over the world. Depending on the type of media, it could amongst other uses, also be used for entertainment or educational purposes. One of the most widely known educational media that has been produced, is Sesame Street. Sesame Street is an educational children's television show aimed at young children from 3 to 7 years old and their parents and was created with the idea that it could be used to prepare children from less advantaged backgrounds for school, by bringing their knowledge up to the level of their advantaged peers

(Cooney, 1967). Sesame Street has been created to bridge this gap between disadvantaged and advantaged children in the areas of basic language skills, concepts of space and time, logical concepts, mathematical concepts, and reasoning skills (Cooney, 1967).

When Sesame Street was first broadcasted, a randomized control trial showed that Sesame Street had an immediate impact on literacy and numeracy in children in the ages of three and four years old (Kearney & Levine, 2019). Kearney and Levine (2019) investigated the long-term effects of Sesame Street on children who were six years old or younger when Sesame Street first aired. Kearney and Levine combined this data with data on broadcast exposure of Sesame Street along with the grades in which a student should be based on their age. From this research they found that students who had more broadcast exposure to Sesame Street between the ages of zero and six years old, were 14% more likely to be in the school grade they were supposed to be based on their age (Kearney & Levine, 2019). Their research also found that Sesame Street improved school performance, in particular for boys.

Sesame Street is constantly being changed and developed in order to stay up to date with the changing views and important events of the country it is being aired in (Gardner, 2021). In the last 10 years, more attention has been put on the subject of gender bias in society. This has been done in the form of scientific research, like gender bias in hiring (José González, Cortina, Rodríguez, 2019), viral social media posts creating awareness to existing problems related to unequal treatment based on gender (Radionova, 2017), and gender bias being mentioned in news outlets (Young, 2019). This attention to the existing problem of gender bias, has caused it to be more relevant to non-researchers as well. The previously mentioned stereotype endorsement research done by Willie et al (2018), has shown that there is partial support for short-term effects on stereotype endorsement in children who watch shows that have gender stereotypes embedded in them. This research was specifically aimed at gender stereotypes affecting STEM (Science, Technology, Engineering, and Mathematics).

Het Klokhuis is an educational show that teaches children about the world, in the broadest sense possible, and is also often used in schools as an extra teaching tool. From different cultures to finding out how cheese is made, many different subjects come up on the show with a new episode every weekday. If a show like this is biased, consistent exposure to bias is definitely possible.

Researchers at Trimbos Instituut (2014) have investigated the effect that *Het Klokhuis* has on children, using the four-week period in 2013 in which *Het Klokhuis* showed one episode per week on child abuse. Besides dedicating one episode per week to this subject, *Het Klokhuis* also launched a website and made a lesson plan for Dutch primary schools. The findings of this research were that due to the consistent exposure to this subject, children became more open to talking about this subject with their teachers, peers, and parents and children also learned how to recognize child abuse. This research showed that consistent exposure of a subject in children's lives by *Het Klokhuis*, influences the children who are exposed to it, taking into consideration that television was not the only medium through which the children were exposed to the subject (also through a website and in class).

For the reasons discussed above, it would be interesting to analyze gender bias in media targeted to children in the age groups of 0 to 12 years old, produced by the Dutch public broadcast corporation NPO (Nederlandse Publieke Omroep) over the last ten years. The use of the last decade is chosen for several reasons. Research by See Jane (2019) has shown that there has been a dramatic rise in leading female characters in children's shows, from 42% in 2008, to 52% in 2018 with 55.3% screen time and 50.3% speaking time. These female characters are also more likely to be shown as leaders in 2018 than men, which challenges the existing gender stereotypes for women in leadership roles (See Jane, 2019)

Investigating the children's television shows on gender bias will be done with the following research question: "In what aspects has gender bias in children's television shows changed throughout the last 10 years?".

Research by the Nederlands Jeugd Instituut (2015) has shown that Dutch children under the age of 12 have varying hours in their use of media. On average, toddlers consume around one hour of media per day, children under the age of 7 consume around two hours of media per day, and children around 12 years old consume around three hours of media per day. In this age group, most of the media consumption time consists of watching television shows, which is why two television shows produced by the Dutch public broadcast corporation NPO will be used in the analysis.

The NPO itself is assigned to broadcast television and radio that is suitable for all groups in society and is maintained by the Dutch government (Ministerie van Algemene Zaken, 2021). The NPO is accessible to everyone in the Netherlands, with or without cable TV. This would mean that children who live in homes where there is a television, but no cable will still be able to see the shows provided by NPO. Comparing the NPO to a commercial television channel like Disney Channel, Disney Channel aims to entertain, while the goal of the NPO is to inform and educate, which is why shows broadcasted by the NPO are more interesting for this research.

Data

To analyze these television shows, the Netherlands Institute for Sound and Vision in collaboration with CLARIAH Media Suite have generated automated speech recognition (ASR) files based on a selection of children's television shows produced by the NPO. The chosen television shows are *Sesamstraat* and *Het Klokhuis*. *Sesamstraat* is an educational children's television show aimed at young children from 3 to 7 years old and their parents (Peters, 2011). *Het Klokhuis* is an educational television show aimed at children between the ages of 7 and 12 years old. These shows have been chosen, because of the age ranges for their target audience and their educational function as a show. These shows are also often used in classrooms as an educational tool, which broadens the exposure of children to these shows, making these shows more influential than other shows (Peters, 2011; Lekker Veel Klokhuis, n.d.).

Dataset

The dataset consists of ASR files saved as .json files, these are files that are generated by a process that is applied to the digitized audio materials of the show. The process converts the input (the audio signals of a show) into textual representation (CLARIAH Media Suite - Version 3, 2018). The version used for the creation of the files in this dataset, is version 3.

There are some limitations when it comes to using ASR which could have influenced the dataset. First, ASR works best when the audio is of good quality. Due to the audio for this research being produced for television, the audio was of good quality, and this should not have affected the ASR file generation. What could have formed a challenge in the ASR data generation, is that ASR typically uses a fixed vocabulary, if a word that does not exist in the vocabulary is used as input, the system will not recognize it. Accented speech could also be a problem for the system.

The generated ASR files consist of the recognized speech in each episode as text, a start time, timecode, and word times after each word. All ASR files are accompanied by a file containing metadata on the file. Amongst other data, the file contains data such as the main subject of the episode, the presenters, and the production year of the episode.

For *Sesamstraat*, ASR has been applied on 20 randomly chosen episodes per year, from 2011 until 2019. Each episode has a duration of around 25 minutes until 2014. From 2014 and on, every episode has a duration of around 15 minutes. In each episode, the number of sentences varies between 50 and 200 sentences.

For *Het Klokhuis*, ASR has been applied on 20 randomly chosen episodes per year, from 2011 until 2020. Each episode has a duration of around 15 minutes, with between 50 and 100 sentences being spoken per episode.

The research is not done on the whole episode, but only on the ASR generated data of it, which is in the form of text. The total amount of words in the corpus is 771732 and each episode contains on average 1437.12 words. The words per year can be seen in figure 1. Due to 2020 only containing episodes from Het Klokhuis, the number of words in 2020 is far less than the years before. The decline in number of words in 2014 can be explained by the decision to shorten the length of the episodes that was made in 2013.

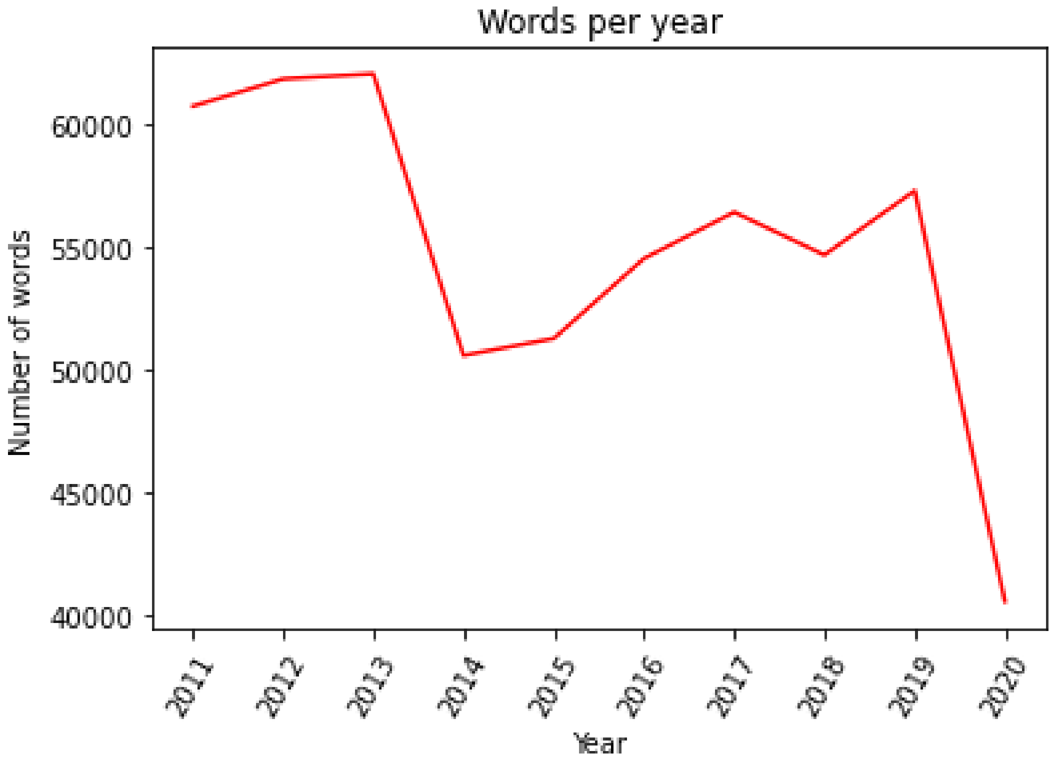

Figure 1 Words per year from 2011 to 2020

Using data from *Sesamstraat* as well as data from *Het Klokhuis* gives a broader range in television consumption by children. These shows together aim to reach an audience of children between 0 and 12 years old that could be affected by any possible bias that could be present in these shows. As previously mentioned, both of these shows are used in education as well and while it does not mean that it gives a complete view of children's television, it does represent an important part of children's television for children who live in the Netherlands.

However, because of the wide age range of the viewers of these shows, the language used in each show differs greatly from each other. For example, *Sesamstraat* uses more simple words while *Het Klokhuis* uses more advanced words that 12-year-olds can understand, but maybe 6-year-olds cannot. This results into a dataset that is diverse in language levels.

To summarize, all years up until 2019 contain 40 episodes total of *Sesamstraat* and *Het Klokhuis*, and 2020 only contains 20 episodes of *Het Klokhuis*, due to the discontinuation of filming new *Sesamstraat* episodes. Due to time constraints for generating the ASR data required for the analysis, twenty random episodes per series were chosen (figure 2). This is limiting to the research, because using all episodes in each year for each show in the analysis could give a more complete image of the bias.

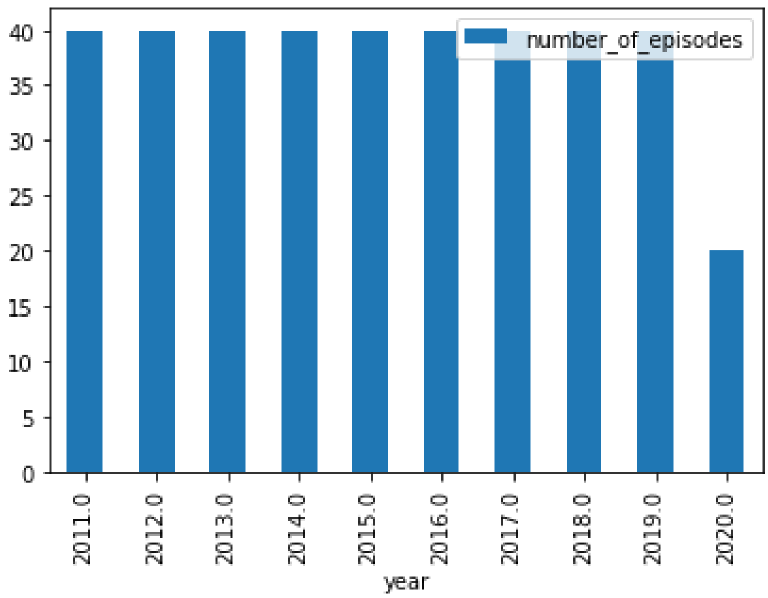

Figure 2 Graph showing the number of episodes per year

Data preparation

First, all ASR files are loaded and put into a Pandas DataFrame. A DataFrame represents a table with rows and columns. Each row consisting of the data in the file, adding an extra column to also add in the filename per row. Then, all ASR files are cleaned up, removing unnecessary characters such as brackets, quotation marks, double spacings, characters indicating white lines, and "words" that consist of only one letter. All words are also made to consist of only lowercase letters. The final DataFrame consists of one column containing the filename, and one column containing the text spoken in each episode. All other data such as start time for each word, and timecodes are omitted.

After loading in and cleaning up all ASR files, the files with metadata are loaded in to a new DataFrame. In the same way as with the ASR files, in this DataFrame an extra column is also created to add in the filename per row. The only information of interest in the metadata files, is the episode year, which is found in the column 'curatedDate'. The year number is extracted from this column for each file, and the ASR DataFrame and the metadata DataFrame are then merged, creating the final DataFrame with the columns 'year', 'text', and the filename set as the index.

Male and female words

To investigate the change in gender bias, it is necessary to be able to recognize words indicating a man or a woman. Two lists have been created, indicating the most common words referencing men and women, and two more lists have been created with the most popular names for men and women in the Netherlands from 1880 until 2007. The words list indicating a woman does not include the word 'ze', because of its ambiguity in meaning. 'Ze' can be used to refer to women as the Dutch word for 'she/her', but it is also often used when referring to multiple people as a form of 'they'. The same can be said for the word 'zij' which has the exact same ambiguity in meaning, but 'zij' is still in the female related words list. This decision to exclude 'ze' but include 'zij' is based on a random sample of 20 episodes over all episodes in which the context relating to the words 'ze' and 'zij' was investigated. In this random sample, all uses of 'ze' were referring to a group of people, and the uses of 'zij' were more often referring to a single person.

The contents of the first list can be seen in table 1.

Table 1 List of female and male words used to create the male and female average vector

Female words	**Male words**
Zij	Hij
Zijzelf	Hijzelf
Haar	Hem
Haarzelf	Hemzelf
Zus	Broer
Zusje	Broertje
Moeder	Vader
Mama	Papa
Ma	Pa
Mammie	Pappie
Mams	Paps
Oma	Opa
Nicht	Neef

The second pair of lists contains the top 100 boy and top 100 girl names in each year, from 1880 until 2011 (Jaarlijkse Nederlandse Voornamen Top 100 van 1880–2011 Naamkunde, 2011). The names lists have been cleaned up to remove duplicates, and the names of the main cast and presenters from *Sesamstraat* and *Het Klokhuis* have been manually added to the corresponding male or female names lists. Finally, the two lists have been added together, creating one list for male related words and one list for female related words.

Data exploration

To explore the data, a count was done on the amount of female related words and the amount of male related words in the corpus, to get an overall view on the data. The count for male related words was done by counting how many times each word in the list of male related words was found in the whole corpus, i.e., in all 380 episodes. The same was done for the female words. Substrings containing the words in the male words list or the words in the female words list have also been counted as a male word or a female word. This is a limitation, because the whole word could have been unrelated to either gender or the whole word could be related to the opposite gender than it was classified as. Take for example 'pa', the Dutch word for dad. The substring 'pa' occurs in multiple Dutch words, one of these words being 'paar'. 'Paar' is the Dutch word for couple and is not a male related word or a female related word, but it could still have been counted as a male related word in this corpus, because the word contains the substring 'pa'. This way of counting has only influenced the data exploration, as it could have given a wrong overall view of the dataset.

From the total amount of words, the male related words count is 18410 and the female related words count is 43278. So, 2.4% of the corpus consists of male related words and 5.6% of the corpus consists of female related words.

By dividing the amount of male related words found by the amount of female related words, the male/female words ratio was calculated. In the whole corpus, the overall male/female words ratio is 0.425. This is roughly around one male related word for every two female related words.

In the figure 3, the male/female related word ratio per year is visualized. As can be seen in figure 3, there are around two to three times as many female related words than male related words in every year.

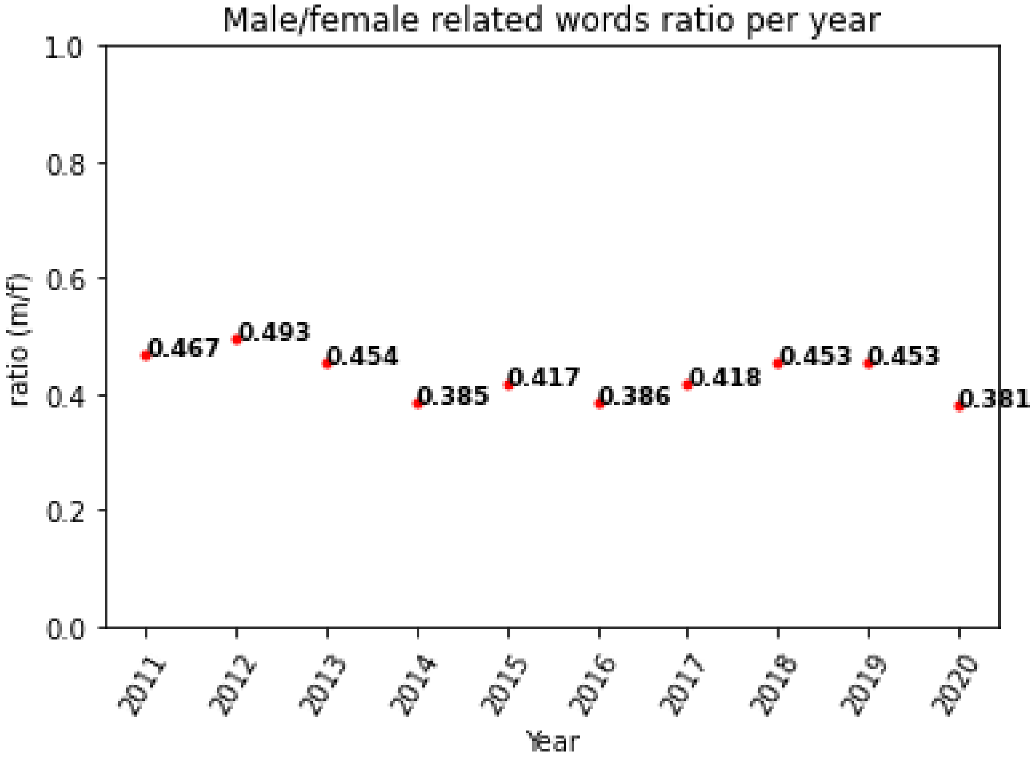

Figure 3 Male/female related words ratio per year

In figure 4, a closer look is taken at male related words that occur a minimum amount of 20 times in the corpus. The minimum number 20 is chosen, because each series has 20 episodes per year, so 20 would be a relevant number to show each word that overall occurs once per episode.

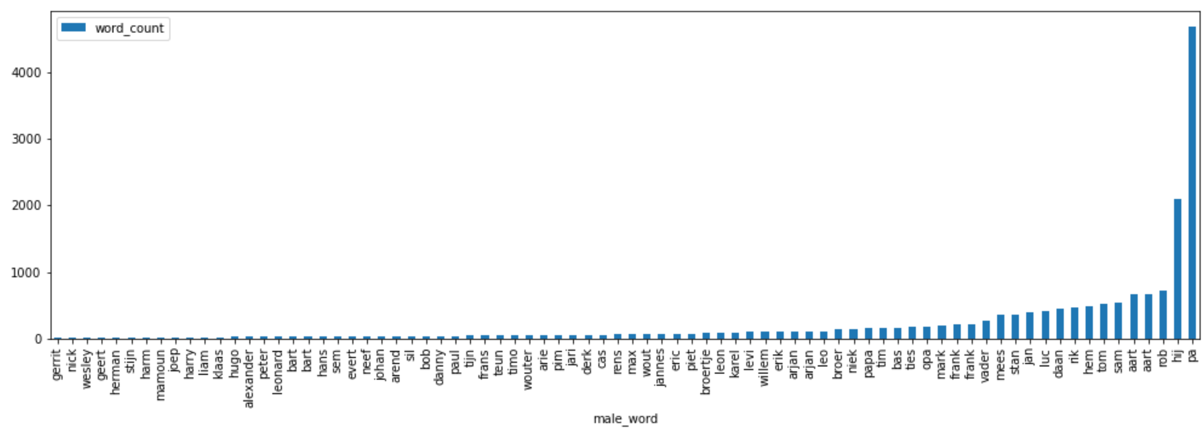

Figure 4 Male related words that occur in the corpus at least 20 times

Figure 4 shows the male related word that is most common in this corpus is 'pa', which is the Dutch word for 'dad'. Given that the series that are being analyzed are aimed at children, this is not a surprising find. The second word that is most often found in the corpus is 'hij', which is a word used when referring to males and can be translated to 'he/him' in English. The most common words after 'pa' and 'hij', are mostly the names of the male actors in the series.

The same has been done for female related words in figure 5, again all words that occur a minimum of 20 times in the corpus are shown.

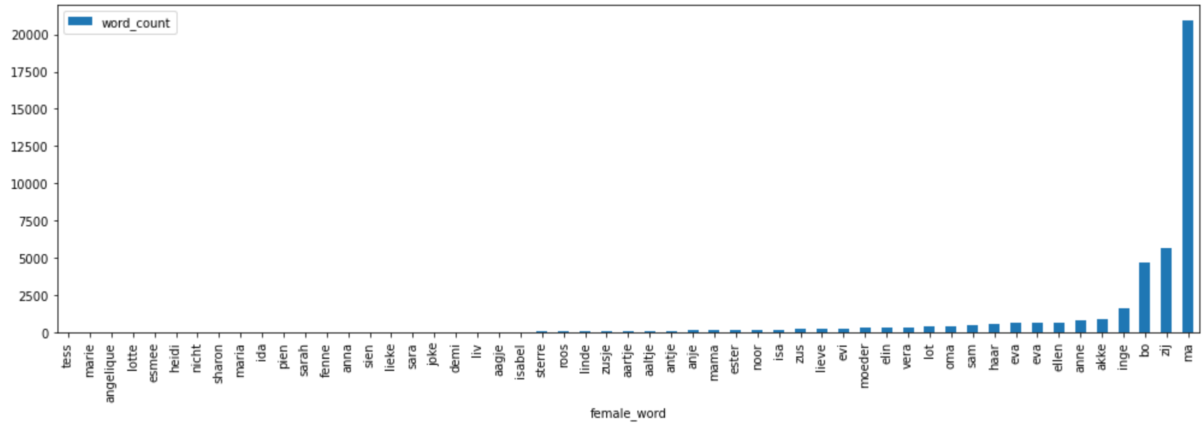

Figure 5 Female related words that occur in the corpus at least 20 times

The most common female related word in the corpus is 'ma', which is the Dutch word for 'mom'. Again, just like with the male related words, it is not surprising that this word occurs so often in the corpus. After 'ma', the most found word in the corpus is 'zij', the female word for 'she/her'. The most common female words after these two, are mostly female names.

While it is unsurprising that the words 'ma' and 'pa' occur most often in the corpus, it is surprising how much more the word 'ma' is used than the word 'pa', almost four times as often. When looking at the male/female word ratios, female words occur two to three times as often, but 'ma' occurs almost four times as often as 'pa', while both are words used to refer to a parent.

Ethical and Legal considerations

The data used in this research does not contain any sensitive data on individuals, companies, or others, which makes using this data for this research ethical. The data has been generated by the Netherlands Institute of Sound and Vision and CLARIAH Media Suite with the proper permissions from the NPO. The dataset used in this research cannot be made publicly available, due to it being the property of the Netherlands Institute of Sound and Vision, CLARIAH Media Suite and NPO. This is a limitation for the reproducibility and verification of this research. However, affiliates can request the dataset at Utrecht University or approach the NPO.

It is important to highlight the fact that this research is based on a subset of episodes of two different children's television shows that are broadcasted in the Netherlands. Due to this, the results from this research should always be interpreted with the size of the dataset in mind. It should also be taken into consideration that this research defines bias as the relatedness of a word to male words or female words, this research does not show the positive or negative gender bias, only the presence of gender bias.

Methodology

In order to be able to answer the research question proposed in the introduction, the research question must first be translated into a question that is answerable by applying data science techniques. In the introduction several aspects of life, like careers and mental health, have been mentioned in which (implicit) gender bias can play a role. Determining gender bias in children's television shows by investigating different categories, like occupation and emotions, and determining the change in gender bias is one way to answer the research question. In this research, bias is defined as the existence of prejudice towards certain (groups) of people. For the calculation of gender bias, the definition of gender bias is the relatedness a word has to male related words or female related words. This calculation can be done by calculating the distance between a male or female related word to a word for a certain occupation or a certain emotion. The greater the distance, the less the word related to a gender is associated with the word, and thus the less the gender is related to the word. This method only shows the presence of bias in the corpus, it does not show whether the bias is positive or negative.

While counting words as done in the 'Data exploration' chapter is one way to show the distribution of male related words and female related words in a corpus, this method lacks the ability to account for relationships between words. An algorithm that can account for those relationships between words and preserve them is Word2Vec, which is what will be used in this research to calculate the change in gender bias over the last decade.

Word2Vec

Word2Vec is a natural language processing algorithm that learns word associations from a large corpus of text. The corpus that is used to train the Word2Vec model, is translated into a large vector space over many dimensions. In that vector space, each word in the model has its own place. Word associations are then calculated by calculating the distance between the vector of the first word in the vector space and the vector of the second word in the vector space (Karani, 2020).

When a word is used as input in the model with the intention to see the most similar words, a value is shown that shows how associated the input word and the output word are. An example of this could be the words 'Utrecht' and 'Moscow' having a small Euclidean distance (more related, means their distance is smaller in the vector space), while 'Apple' and 'Moskou' have a higher Euclidean distance (less related, means their distance is bigger in the vector space). The relationship between the words 'Utrecht' and 'Moscow' is much more obvious (both are cities) than between the words 'Apple' and 'Moscow' (one is a fruit, and one is a city), which is why the cities have a small Euclidean distance. It is also possible that two words that are expected to have the same relatedness to each other, have different

Euclidean distances, like 'Utrecht' and 'Moscow', and 'Rotterdam' and 'Moscow'. Let's say in this example that 'Utrecht' and 'Moscow' have a smaller Euclidean distance than 'Rotterdam' and 'Moscow'. All three words are cities, Rotterdam and Utrecht both being Dutch cities, how is it then possible that 'Utrecht' and 'Moscow' have a smaller Euclidean distance? An explanation for this could be that 'Utrecht' and 'Moscow' appear closer together in the same corpus more often than 'Rotterdam' and 'Moscow', which has resulted in different Euclidean distance values between these two examples. For this reason, it would be interesting to use Word2Vec to investigate the association of male/female related words to words in certain LIWC categories.

There are two ways the Word2Vec model can calculate the association between two words: Skipgram and CBOW. When using a Skipgram Word2Vec model, the input word is used to determine the output of the model. CBOW does the opposite and uses all surrounding words to determine the output. A visual representation of this process can be seen in figure 6.

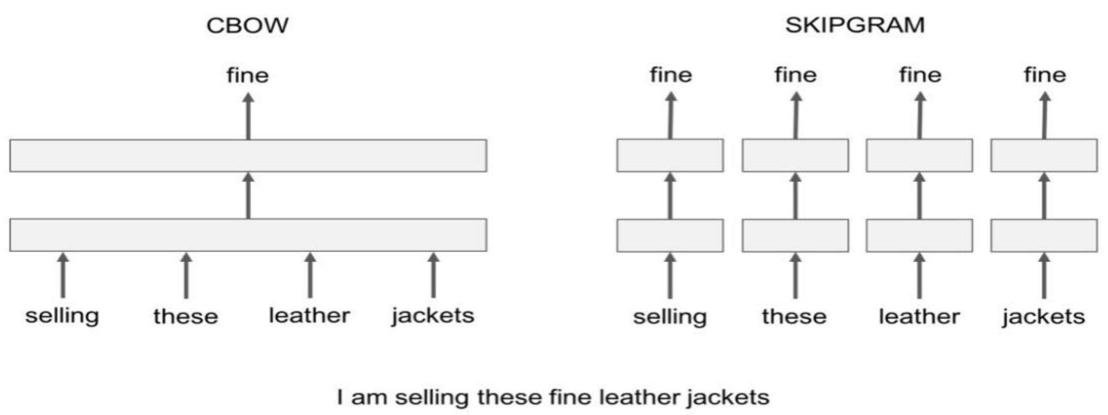

Figure 6 Explaining the difference between CBOW and Skipgram with an example sentence (Fast Text, 2020).

Both methods could be used to calculate the gender bias, but for this research, Word2Vec will be used with Skipgram, because Skipgram is better at in finding semantic relationships, instead of just morphologically similar words like plurals (Mikolov et al, 2013). It is also less sensitive for overfitting because it takes one word at a time.

The embedding size or dimensionality of the vector is determined by the 'size' parameter, which is set to 300. This size is chosen because it's the conventional size to use for corpora of this size.

Word2Vec also contains a parameter 'window size'. The value of 10 is most commonly used with Skipgram and is used to indicate how interchangeable two words are based on their distance to each other. In this research, the value of 10 is chosen for the window size.

Another parameter is the number of workers. Workers indicate the number of parallel ways the model is being trained. The optimal number for this parameter is the number of cores the machine that the code is being run on has, which is four in this case.

The model only trains on words that appear at a minimum of three times in the corpus, this parameter is the 'min_count'. The value of three is chosen, because of the previously mentioned repeated exposure that is important in the creation of implicit bias.

The final parameter is the number of epochs, which is the number of times the model iterates over the corpus. This number is set to equal the standard value, which is five iterations.

In summation, the model parameter settings can be seen in table 2.

Table 2 Parameter settings for the Word2Vec model

Parameter	**Value**
SIZE	300
SG	1 (Skipgram)
WINDOW	10
N_WORKERS	4
MIN_COUNT	3
EPOCHS	5

Training the model

This created model is then trained on the corpus filtered by year. For each year, the corpus is split into sentences and each sentence is then further split into words, which is also known as tokenization. Splitting the corpus into sentences is beneficial when it comes to the computing time of training the model, it takes less time to find related words, because there are already some related words in the sentence itself. The pretrained Spacy library nl_core_news_sm for the Dutch language is trained on data from Dutch newspapers and is used to reduce the training time of the model.

As done in previous research by Wevers (2019) and Garg et al (2018), the Word2Vec model is used to construct an average vector that is created by the previously created list with male related words, such as 'hij' (he), 'vader' (father), and 'man' (man) and the most popular names in the Netherlands for males in the last decade. The same is done to create an average vector for the list containing female related words, such as 'zij'(she), 'haar'(her), and 'vrouw'(woman) and the most popular names in the Netherlands for females in the last decade. These two average vectors represent the gender dimensions (male and female) per television show.

LIWC

To calculate the bias in each category, the Dutch Linguistic Inquiry and Word Count (LIWC) dictionary is used. The LIWC is a list of words separated by category, developed by researchers with interests in social, clinical, health, and cognitive psychology (How does LIWC analyze language?, 2015). These categories vary from 'music' to 'family' to 'jobs'. Each category is defined by a list of words. This list of words, or dictionary, is often used in text analysis and is also used in this research. Gender bias has many consequences, for this reason, the following LIWC categories have been chosen to calculate the gender bias in: *groom, school, occup, posemo, achieve, negemo*.

The category '*groom'* contains words related to appearance and hygiene, like the Dutch words for washing and soap. This category has been chosen, because of the existing stereotype of women having better hygiene than men (Crawford, 2015). However, this category also contains the Dutch word for shaving, which can also be used when speaking for example about men's beards.

The category '*school'* contains words related to school, like professor and exam. This category is chosen because of the gender differences in the enrollment and completion of higher education (Stoet & Geary, 2020).

The category '*occup'* contains words related to occupations, such as the Dutch words for colleagues and bonus. This category is chosen due to the percentual gender difference in the amount of working people in the Netherlands, 63,2% of women versus 72,5% percent of men in 2018 and 61,3% of women versus 71,7% of men in 2011 (Centraal Bureau voor de Statistiek, 2019).

The category '*posemo'* contains words related to positive emotions, like winning and peaceful. The category '*negemo'* contains words related to negative emotions, such as wrong and boring. The final category 'achieve' contains words related to achievements, like fail and win. All three of these categories have been chosen due to their influence on the self-worth of children (Kamins & Dweck, 1999).

Figure 7 shows the occurrence of the words in each category in the corpus per year.

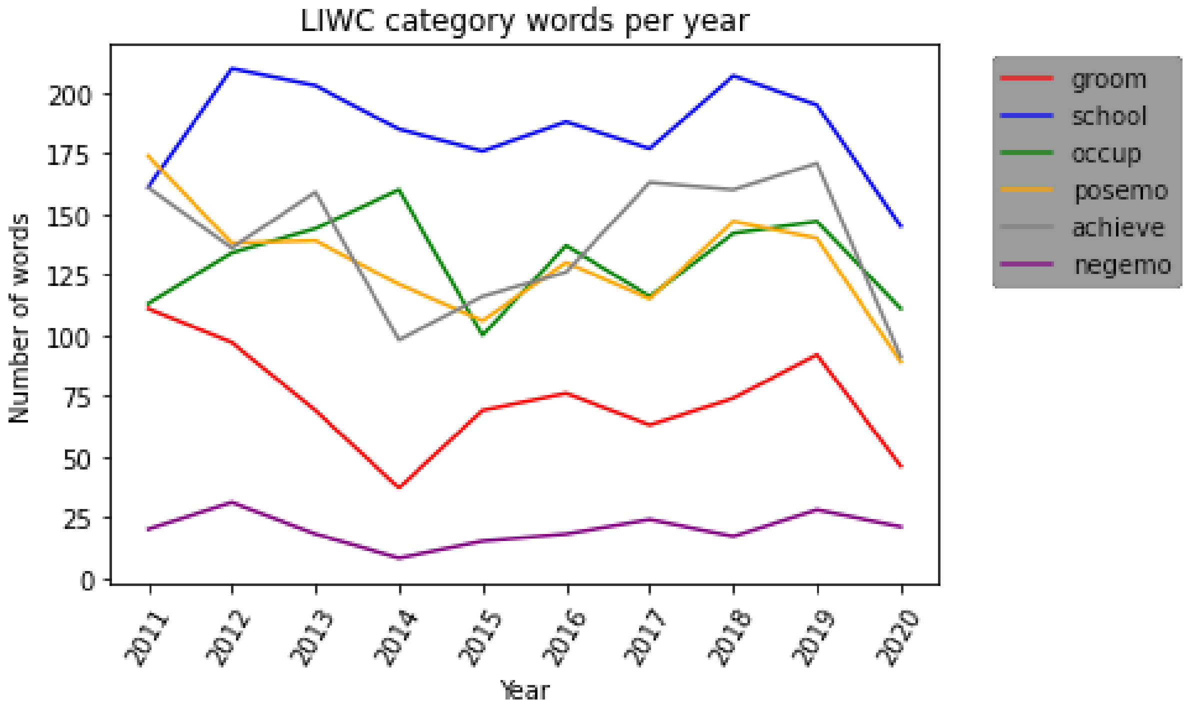

Figure 7 LIWC category word occurrences per year

While the categories *occup, posemo*, and *achieve* are close in number over all years, *groom* and *negemo* differ greatly in number of word occurrences. The words in the category *school* occur the most in the corpus. This already shows that each category is not represented equally in the corpus, which will have consequences for the results.

Calculating bias

To calculate the gender bias for each category, first for each word within an LIWC category the distance of the word to the male average vector was calculated. The same was done for the female average vector. By subtracting the distance to the female average vector for each word from the distance to the male average vector, the bias per word is calculated. Taking the average bias for each category, gives the gender bias per LIWC category per year. This distance calculation is based on the Word-Embedding Association Test (WEAT) as introduced by Caliskan, Bryson and Narayanan (2017), This measure of semantic association was specifically made to show bias in word embedding models. A positive bias value denotes bias towards females, a negative bias value denotes bias towards males.

The model is trained on word level with sentences, but the distance calculations are made on model level. This means that as previously explained, the corpus is split into sentences which are then further split into words. The model calculates the distance between the words in each sentence and creates a vector for each word. It does this over all sentences in the corpus. When calculating the distance between a word and the average vector that represents a gender, this is done on model level. This means that the final vector that is created over all occurrences of a certain word in the corpus, is the vector that is used to calculate the distance between the word and the average vector.

In table 3, the interpretation of the bias values is shown (Wevers, 2019). While Wevers (2019) does not explicitly state the chosen boundaries, in his analysis he shows the shifts in bias over the years using the values in the manner shown in table 3. Bias values between 0 – -0.01 and 0 – 0.01 are deemed small enough to be considered insignificant. A value between -0.01 – -0.2 and between 0.01 – 0.2 can be interpreted as *slight bias*, with *slight* reflecting the strength of the bias that is present. All values < -0.2 and > 0.2 can be interpreted as bias, which means bias is clearly present in the analyzed corpus.

Table 3 Interpretation of bias values and their boundaries.

Gender bias value male words	Interpretation	Gender bias value female words
0 – -0.01	Insignificant	0 – 0.01
-0.01 – -0.2	Slight bias	0.01 – 0.2
< -0.2	Bias	> 0.2

To properly answer the research question, the bias needs to be measured over time. For the final step in the analysis, linear regression is done on each LIWC category and the r2 value is calculated. Linear regression shows a straight line, also known as the trendline, that best fits the data points in a way that the line has the least amount of distance to each datapoint. The r2 value, also known as the coefficient of determination, when combined with the trendline, shows the proportion of datapoints that are on the trendline. It shows how well the line fits on the observed data points. These measures will be used to answer how the gender bias has changed over the last decade. The r2 value and the trendline are calculated on the average value of each category in each year and not on the underlying data. The r2 value shows the relationship between variables, by using the average value in each year, the r2 value shows the relationship between the averages of one variable (the word categories) and the non-average data of the other variable (year).

Results

After running the Word2Vec model on each year, the bias for each category per year was generated, as can be seen in table 4. What this value shows, is how strongly the words in each chosen LIWC category are associated to either a word related to males, or a word related to females.

Table 4 Results for gender bias per LIWC category, from 2011 to 2020

Category	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Groom	0.0496	0.0142	0.0259	-0.0024	0.0008	-0.0387	-0.0092	-0.0074	-0.0067	0.0246
School	0.0572	0.0359	0.0255	0.0174	0.0285	-0.0431	-0.0061	0.0321	-0.0023	0.0107
Occup	0.0255	0.0152	0.0064	0.0142	-0.0109	-0.0138	0.0075	0.0277	-0.0001	0.0065
Posemo	0.0108	0.0159	-0.0076	0.0129	-0.0015	-0.0229	0.0207	0.0257	-0.0043	-0.0032
Achieve	0.0072	0.0018	-0.0089	0.0249	-0.0389	0.0495	-0.0056	0.0309	0.0005	-0.0130
Negemo	0.0140	0.0246	0.0068	-0.0259	0.0086	-0.0496	0.0237	0.0249	-0.0010	0.0130

For easier analysis of the data, the values for bias have been plotted to visually illustrate the datapoints. A datapoint above zero indicates bias towards women, a datapoint below zero indicates bias towards men, and a datapoint at zero indicates no bias. The no bias boundary is emphasized by the pink line in each graph. The minimum and maximum y values have been chosen for better visualization of the results and are based on the boundary for slight bias (table 3).

Figure 8 shows the bias for the LIWC category 'Groom' per year. This category contains words related to grooming and hygiene. The category 'Groom' starts with slight bias towards women in 2011, gets greatly decreased in the years between 2011 and 2020, and ends with slight bias towards women in 2020. With the exception of 2016 in which there is slight bias towards men for the category 'Groom', between the years 2014 and 2019, the bias value is so close to 0 that it is deemed insignificant. The trendline and the r2 value show that there is a very weak decreasing trend over the years with regards to gender bias in the category

'Groom'. What this means for the gender bias in this category, is that the gender bias towards females in the category 'Groom' has shown a decrease over the years, due to the words used in the category 'Groom' being associated less with female related words.

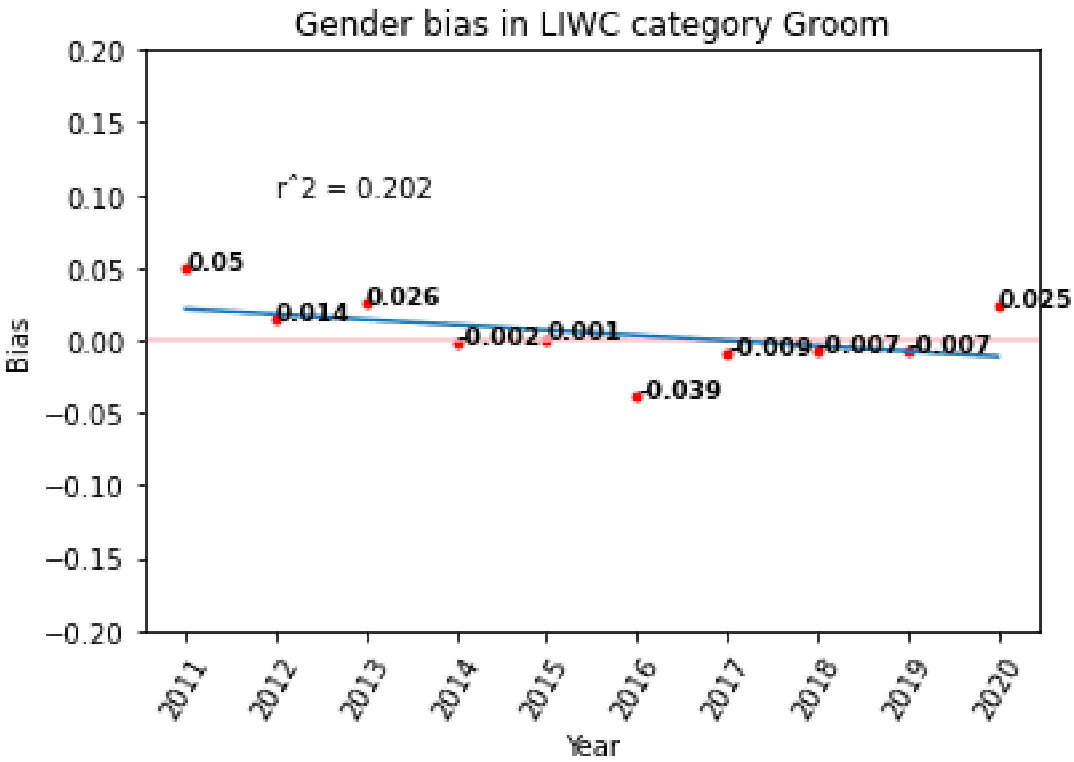

Figure 8 *Visual representation of gender bias throughout the years for the LIWC category 'Groom'*

Figure 9 shows the bias in the LIWC category 'School' per year. This category contains words related to school. In 2011, the category 'School' starts off slightly biased towards women, decreases yearly until 2015 where the category's bias values start to fluctuate between slight bias and insignificant bias. The trendline and the r2 value show a weak decreasing trend for the bias values in the category 'School'. This means that over the last decade, school related words have become less associated with female related words and the bias value is close to being insignificant for this category.

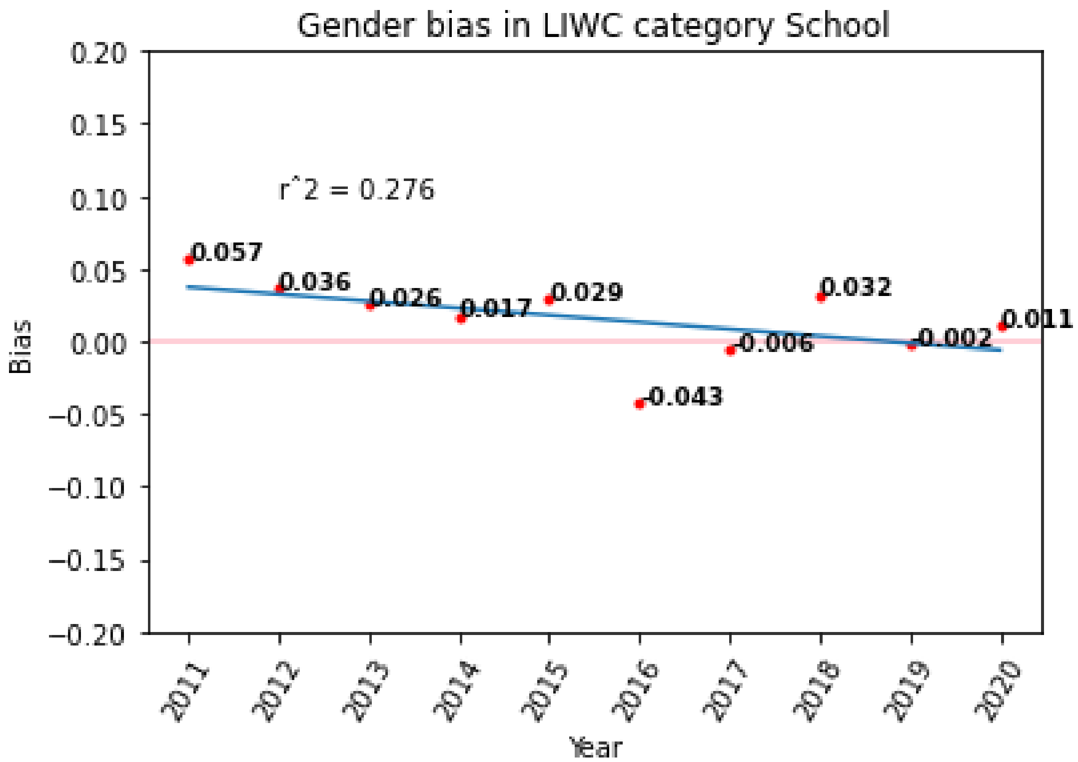

Figure 9 Visual representation of gender bias throughout the years for the LIWC category 'School'

Figure 10 shows the bias in the LIWC category 'Occup', which stands for occupation, per year. This category contains words related to occupations, like work and work activities. In 2011 the category 'Occup' starts with slight bias towards women, decreases and stays around zero for the following years with a small increase in 2018 and ends with insignificant bias in 2020. From the trendline and r2 value, it is visible that there is no trend for this category. For the category 'Occup', the gender bias is mostly close to being insignificant since 2011, meaning there is almost no gender bias present in this category.

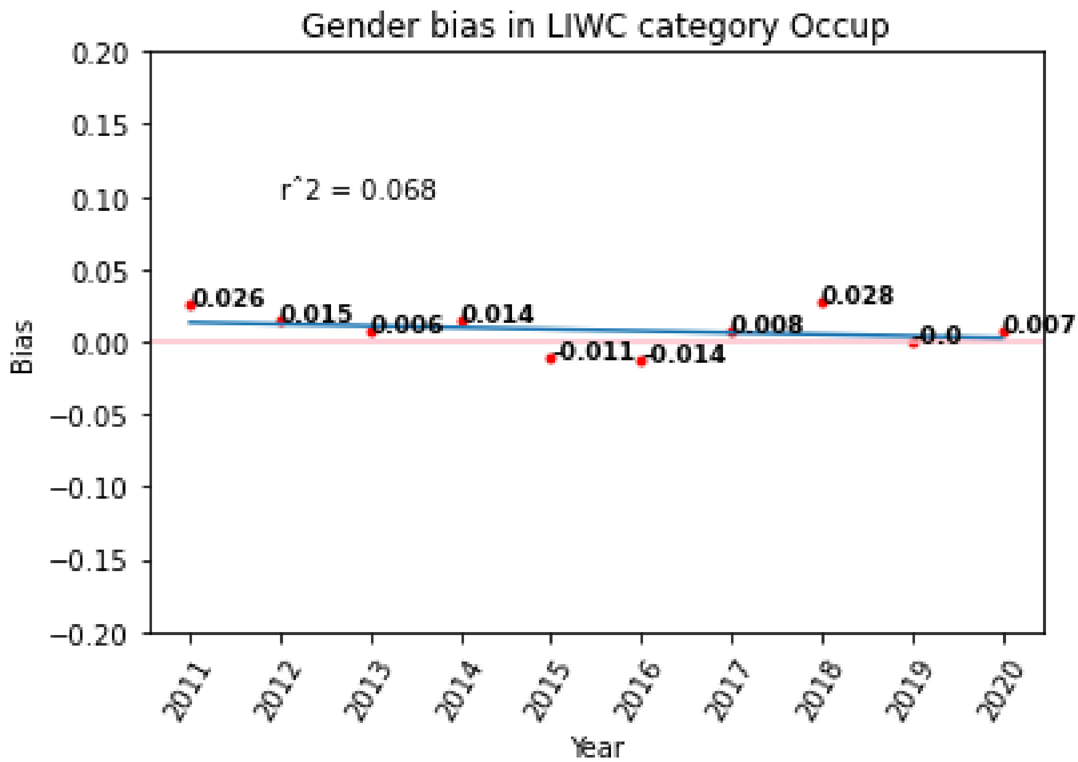

Figure 10 Visual representation of gender bias throughout the years for the LIWC category 'Occup'

Figure 11 shows the LIWC category 'Posemo' per year. This category contains words related to positive emotions, like happiness. In 2011 the category 'Posemo' starts off slightly biased towards women, this slight bias fluctuates and in 2020 it ends with a bias value that can be deemed insignificant. For this category it is also the case that there has been no gender bias trend over the last decade.

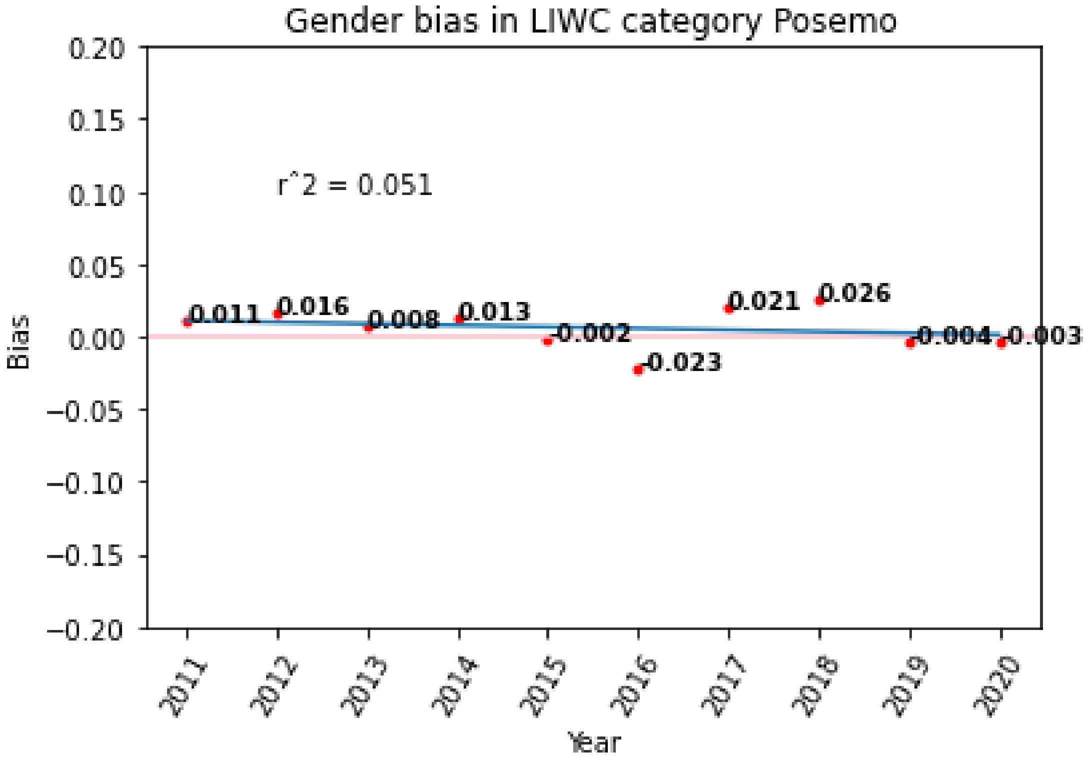

Figure 11 Visual representation of gender bias throughout the years for the LIWC category 'Posemo'

Figure 12 shows the LIWC category 'Achieve' per year. This category contains words related to achievements, such as win and lose. From 2011 to 2013 the figure shows insignificant bias values, in 2014 this category becomes slightly biased towards women and turns back into slight bias towards men in 2015. From 2016 until 2019 the bias values for the shows fluctuate between slight bias towards women and an insignificant amount of bias and in 2020 there is slight bias towards men in this category. There is no trend over the last decade for how gender bias has changed, as can be seen by the r2 value and the trendline.

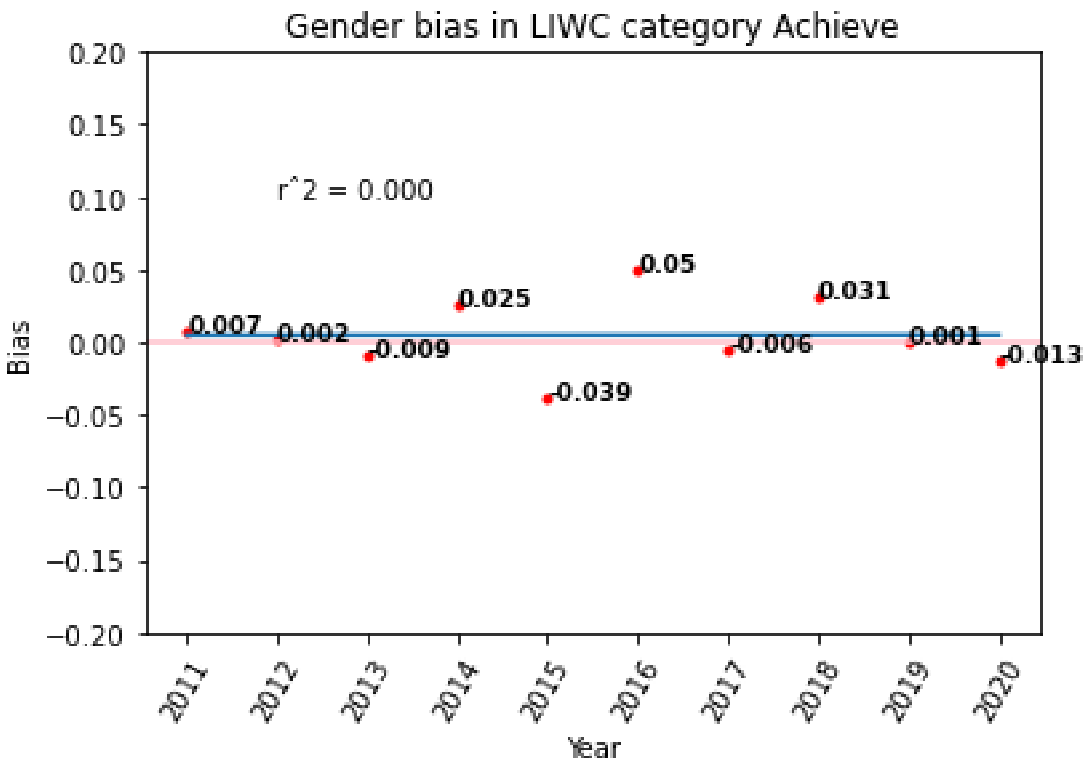

Figure 12 Visual representation of gender bias throughout the years for the LIWC category 'Achieve'

Figure 13 shows the bias values for the LIWC category 'Negemo' per year. This category contains words related to negative emotions, such as fail and hate. In 2011 for this category there is slight bias towards women, the bias values fluctuate between slight bias insignificant amount of bias towards either gender. The bias turns towards women between 2017 and 2018 and turns into almost no bias towards either gender in 2019 and 2020. There is no trend visible in the change of bias over the last decade for this category.

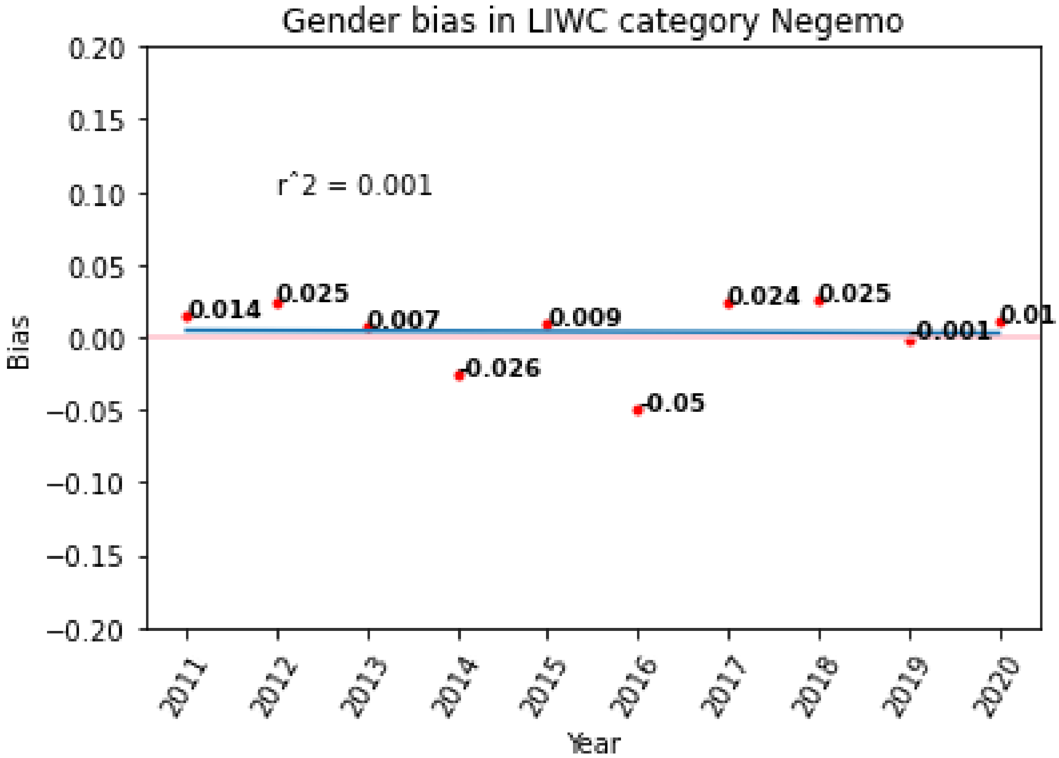

Figure 13 Visual representation of gender bias throughout the years for the LIWC category 'Negemo'

For the year 2015, a closer look is taken at the words that were calculated for bias within the categories ‘Groom’ and ‘Posemo’. Figure 14 shows the words in the category ‘Groom’ that were also found in the corpus and their bias value. Three words appear in both the ‘Groom’ category and the corpus. The words ‘haren’ and ‘schoonmaken’ are biased towards men and the word ‘schoon’ contains bias towards women.

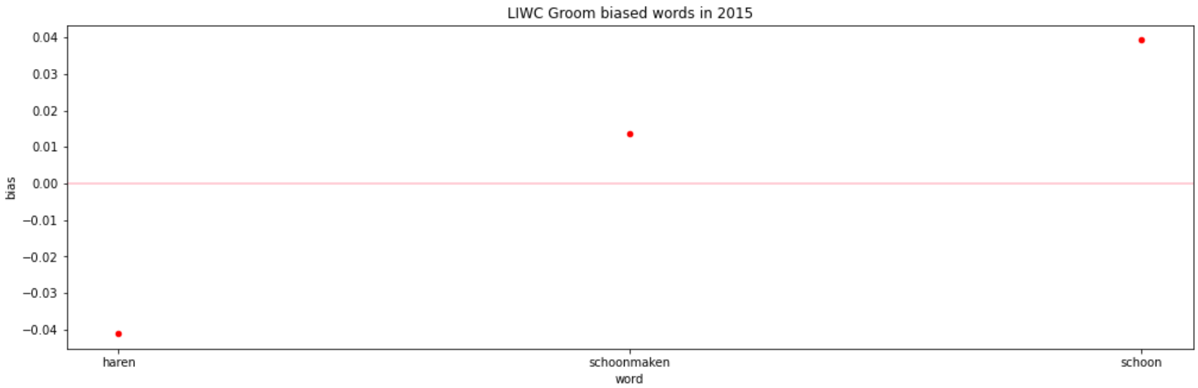

Figure 14 Visual representation of the bias value of the words in the corpus that are also present in the LIWC category 'Groom' in 2015

Figure 15 shows the words that are in the LIWC category ‘Posemo’ and are also found in the corpus in the year 2015. There are 19 words that appear in the ‘Posemo’ category as well as in the corpus. Out of these 19 words, 9 words are biased towards women and 10 are biased towards men.

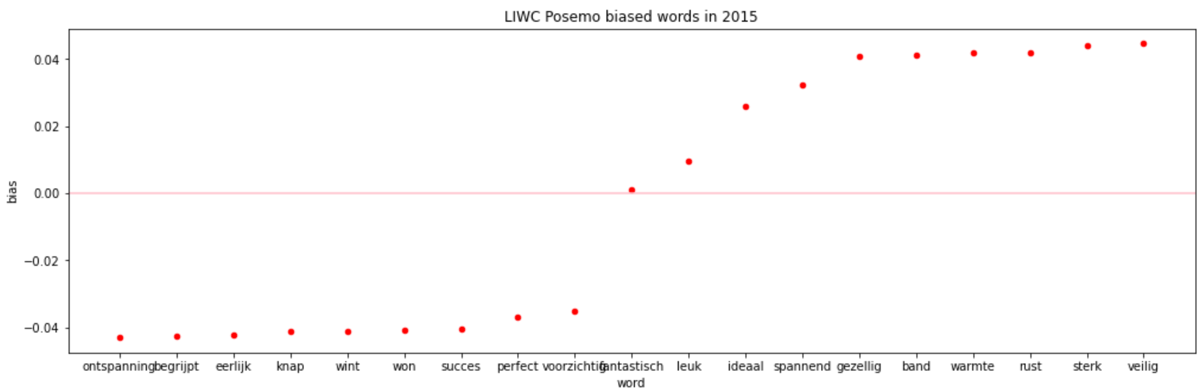

Figure 15 Visual representation of the bias value of the words in the corpus that are also present in the LIWC category 'Posemo' in 2015

For the category ‘Posemo’, a closer look is also taken at the words that appear in the ‘Posemo’ category as well as in the corpus for 2014, as can be seen in figure 16. In this corpus there are also 19 words that are found. Out of these 19 words, ten are biased towards women and 9 are biased towards men. The words that are found in 2014, are mostly different from those found in 2015.

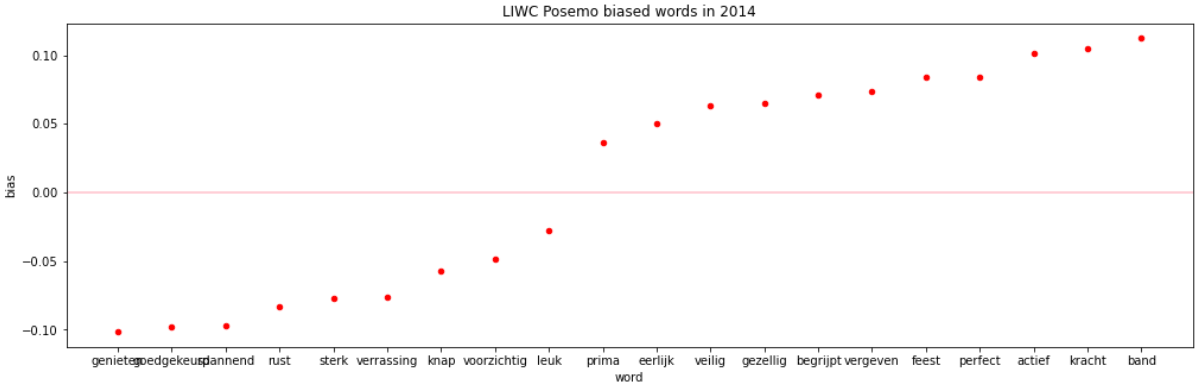

Figure 16 Visual representation of the bias value of the words in the corpus that are also present in the LIWC category 'Posemo' in 2014

Conclusion and discussion

From the results it is visible that there is slight gender bias present in certain categories in *Sesamstraat* and *Het Klokhui*s, that has changed over the last decade. These results lead to circling back to the research question proposed in the introduction: “In what aspects has gender bias in children’s television shows changed throughout the last 10 years?”.

In the LIWC categories ‘Groom’ and ‘School’, the bias has shifted throughout the decade from being slightly biased towards women to being less biased towards women. This means that words that are related to grooming and words that are related to school were closer associated with female related words, than with male related words in the beginning of the decade. At the end of the decade, words related to grooming and words related to school were still closer related to female related words than to male related words, but less than in the beginning of the decade.

The remaining categories ‘Occup’, ‘Posemo’, ‘Achieve’, and ‘Negemo’, show no trend in how the bias has changed, with ‘Occup’ en ‘Posemo’ having mostly an insignificant amount of bias over the last decade. This means that the words related to occupations and positive emotions are not associated more with words related to either gender. They are associated with male related words and female related words almost equally.

For words related to achievements and negative emotions, there have been fluctuations in slight bias towards either gender or an insignificant amount of bias, with no trend over the years. For example, words related to negative emotions were closer related to male related words than to female related words in 2016, while words related to negative emotions were closer related to female related words than to male related words in 2017 and 2018, as can be seen in figure 14.

The lack of a trend in these categories could be explained in two ways. The first explanation could be that there simply is no trend. The second explanation could be that the amount of data is insufficient to show the underlying trend. This explanation could be possible, especially when looking at the word count of the LIWC categories in the corpus combined with the number of words per year, as can be seen in figure 1 and figure 7.

Figure 14 and figure 15 show the difference in the number of words that are found in the corpus relating to the categories ‘Groom’ and ‘Posemo’. There are less words in the corpus relating to ‘Groom’ than there are words related to ‘Posemo’. This would mean that the occurrence of a word in the category ‘Groom’ has more weight for the bias value than the occurrence of a word in the category ‘Posemo’. The unequal representation of every category could have led to bias values that do not represent the actual bias.

Overall, from this research with the chosen categories and the chosen dataset, slight bias has been found in the categories ‘Groom’ and ‘School’ towards women, which has become less over the last decade, but still slightly biased towards women; mostly an insignificant amount of bias has been found for words related to occupations and words related to positive emotions, with no trend over the last decade; fluctuating slight bias towards either gender and an insignificant amount of bias has been found for words related to achievements and words related to negative emotions. No clear bias has been found in the categories that have been investigated, meaning that the investigated shows are not clearly biased towards either gender in the investigated categories, despite the number of female words being used in the corpus being far greater than the number of male words which could have led to the expectation of there being bias towards women.

For example, from the data exploration it was found that the Dutch word for mom ('ma') is found in the corpus around four times as often as the Dutch word for dad ('pa'). The associations however do not show any clear bias in any category towards women, while just looking at the data exploration would paint a different image.

Implications and Limitations

As discussed, there is slight bias present in the two analyzed children's shows in the selected categories. The bias differs per category, but it is still slightly present and could potentially be picked up on by children watching these shows. However, not every word in the selected LIWC categories was also present in the corpus the model was trained on. For example, figure 14 shows only three words, while figure 15 shows 19 words. The number of words per category is thus unequal, causing the bias value in figure 6 to potentially not be as representative for bias as the other categories. Also, when looking at a year-by-year level and ignoring the lack of a trend, fluctuations in the bias value can be seen in each category. This can be caused by several reasons, like current events which have influenced episodes of the analyzed shows. It does not necessarily mean that the bias is necessarily harmful, what the value says is that a word is more often associated to a certain gender. Another reason for this fluctuation could be the different number of total words per year, as can be seen in figure 1.

Besides the unequal number of words detected per category, how much bias is actually influential to children watching these shows? There is no concrete metric that says this amount of word relatedness causes a child to think women cannot drive, for example. Even if there was, at how much exposure does it influence the child? Is it influential to children who watch every day, or is it also influential to children who watch these shows twice a week?

This brings on a limitation to this research, children's television shows are made in a way to hold the child's attention. The audio is only one part of the show, *Sesamstraat* as well as *Het Klokhuis* make use of visuals to make their shows more interesting and hold the attention of children for the duration of the show.

Another limitation to this research can be attributed to the use of automated speech recognition. The challenges of ASR that have been mentioned in the methodology section, could possibly have affected the results in this research. The chosen method for bias analysis was chosen amongst other reasons, due to its previous use in similar research by Wevers (2019) and Garg et al. (2018). A limitation for this research compared to the research done by Wevers (2019) and Garg et al. (2018), could be that they used written text for their analysis, while in this research audio is first translated into written text and then used for analysis. ASR is not without errors, so this could have caused limitations in the research. One of these errors of ASR is related to the Netherlands being a multicultural country, and both shows used in this analysis incorporate diversity into their shows (Ministerie van Onderwijs, Cultuur en Wetenschap, 2021). This sometimes results in names that are unique for ASR on Dutch audio, which could result in the name being either not included or misspelled. This limitation is also connected to another limitation in this research, which is the list of popular names in the Netherlands that is used. Over the last decade, the average birth rate per year in the Netherlands has been 171000 births per year (Geboorte - Cijfers | NJi, 2021). The list of popular names only uses the top 100 most popular names for each year. So, there is a possibility that names that are mentioned are not recognized as either male related word or female related word.

Adding on to the limitations of the list of female related words, in the Dutch language the word 'zij' can be used as a word referring to a woman or girl as well as that it can be a word referring to groups of people. In this research this could have resulted in some words being closer to the average vector for female related words, while the words might not have actually been close to a female related word.

The final limitation to this research is that it only covers two children's television shows, while children watch a much wider variety of shows. Also, of the two series that were chosen only a limited number of episodes per year was analyzed. Using all the episodes could lead to different results.

The limitations of this research make the research valuable as a steppingstone towards further research into the field of gender bias related to children's television shows.

References

Advanced readers: skipgram versus cbow. (2020). [Illustration]. https://fasttext.cc/docs/en/unsupervised-tutorial.html

Bussey, N. (2016). *Female characters in film and TV motivate women to be more ambitious, more successful, and have even given them the courage to break out of abusive relationships*. See Jane. https://seejane.org/wp-content/uploads/female-characters-in-film-and-tv-motivate-women-to-be-more-ambitious-more-successful-and-have-even-given-them-the-courage-to-break-out-of-abusive-relationships.pdf

Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, *356*(6334), 183-186.

Cao, J., & Banaji, M. R. (2016). The base rate principle and the fairness principle in social judgment. *Proceedings of the National Academy of Sciences*, *113*(27), 7475-7480.

Centraal Bureau voor de Statistiek. (2019, January 19). *Verschil arbeidsdeelname mannen en vrouwen weer kleiner*. https://www.cbs.nl/nl-nl/nieuws/2019/03/verschil-arbeidsdeelname-mannen-en-vrouwen-weer-kleiner

CLARIAH Media Suite - Version 3. (2018). Mediasuite.Clariah.Nl. https://mediasuite.clariah.nl/documentation/release-notes/v3

Cooney, J. G. (1967). The potential uses of television in preschool education.

Crawford, R. (2015, December 18). *Cleaning Matters Magazine - Women clean; men don't: The stereotype that won't go away* [Blog post]. Cleaning Matters. https://www.cleaning-matters.co.uk/page_668111.asp

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of personality and social psychology*, *56*(1), 5.

Gaddes, R., Jacobson, Z., Montgomery, S., Moore, C., & Staff, D. E. (2021). Conscious and Unconscious Gender Bias.

Gardner, K. (2021, March 31). *How Sesame Street Uses Muppets to Teach Inclusion in the US and Abroad*. American University. https://www.american.edu/sis/news/20210401-how-sesame-street-uses-muppets-to-teach-tolerance-in-the-us-and-abroad.cfm

Geboorte - Cijfers | NJi. (2021, February 15). Nederlands Jeugdinstituut. https://www.nji.nl/nl/Databank/Cijfers-over-Jeugd-en-Opvoeding/Cijfers-per-onderwerp/Geboorte

Gelman, S. A., & Meyer, M. (2011). Child categorization. Wiley Interdisciplinary Reviews: Cognitive Science, 2(1), 95-105.

González, M. J., Cortina, C., & Rodríguez, J. (2019). The role of gender stereotypes in hiring: a field experiment. European Sociological Review, 35(2), 187-204.

Hoefnagels, C., Onrust, S., Goossens, F., van Haastrecht, H., Staal, J., Bartelds, S. & Speetjens, P. (2014). Nemen kinderen het woord. *Effectevaluatie van Klokhuis over Kindermishandeling. Utrecht: Trimbos-Instituut.*

How does LIWC analyze language? (2015). LIWC. https://liwc.wpengine.com/how- it-works/

Jaarlijkse Nederlandse voornamen top 100 van 1880–2011 « Naamkunde. (2011). Naamkunde.Net. http://www.naamkunde.net/?page_id=398

Kamins, M. L., & Dweck, C. S. (1999). Person versus process praise and criticism: implications for contingent self-worth and coping. *Developmental psychology*, *35*(3), 835.

Karani, D. (2020, September 2). *Introduction to Word Embedding and Word2Vec - Towards Data Science*. Towards Data Science. https://towardsdatascience.com/introduction-to-word-embedding-and-word2vec-652d0c2060fa

Kostyrka-Allchorne, K., Cooper, N. R., & Simpson, A. (2017). The relationship between television exposure and children's cognition and behaviour: A systematic review. *Developmental Review*, *44*, 19-58.

Lekker veel Klokhuis. (n.d.). Klokhuis. Retrieved June 6, 2021, from https://www.hetklokhuis.nl/algemeen/in%20de%20klas

Madigan, S., Browne, D., Racine, N., Mori, C., & Tough, S. (2019). Association between screen time and children's performance on a developmental screening test. JAMA pediatrics, 173(3), 244-250.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Ministerie van Onderwijs, Cultuur en Wetenschap. (2021, February 15). *Wat doet de Nederlandse Publieke Omroep (NPO)?* Rijksoverheid.nl. https://www.rijksoverheid.nl/onderwerpen/media-en-publieke-omroep/vraag-en-antwoord/nederlandse-publieke-omroep

National Endowment for the Arts. (2008, December). *WOMEN ARTISTS: 1990 TO 2005*. https://www.arts.gov/sites/default/files/96.pdf

Nederlands Jeugdinstituut (2015). Factsheet Media in het gezin.

NHS. (2017). *Improving access to psychological therapies (IAPT) data set reports*. https://digital.nhs.uk/data-and-information/data-collections-and-data-sets/data-sets/improving-access-to-psychological-therapies-data-set/improving-access-to-psychological-therapies-data-set-reports

Office for National Statistics. (2017, September). Suicides in great Britain: 2016 registrations https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/suicidesintheunitedkingdom/2016registration

Peters, I. (2011, October). Sesamstraat in de groep: Praktijkbeschrijvingen van het gebruik van Sesamstraat met jonge kinderen.

Philipson, I. J. (1993). On the shoulders of women: The feminization of psychotherapy. New York: Guilford Press.

Radionova, Z. (2017, March 10). Gender Inequality: What happened when a man and woman switched names at work for a week. Independent. https://www.independent.co.uk/news/business/news/gender-inequality-man-woman-switch-names-week-martin-schneider-nicky-knacks-pay-gap-a7622201.html

See Jane. (2019). *Historic Gender Parity in Children's Television*. https://seejane.org/wp-content/uploads/see-jane-2019-full-report.pdf

Smith, R., Kelly, B., Yeatman, H., & Boyland, E. (2019). Food marketing influences children's attitudes, preferences and consumption: A systematic critical review. *Nutrients*, *11*(4), 875.

Stoet, G., & Geary, D. C. (2020). Gender differences in the pathways to higher education. *Proceedings of the National Academy of Sciences*, *117*(25), 14073-14076.

Vogel, D. L., Heimerdinger-Edwards, S. R., Hammer, J. H., & Hubbard, A. (2011). "Boys don't cry": Examination of the links between endorsement of masculine norms, self -stigma, and help-seeking attitudes for men from diverse backgrounds. *Journal of counseling psychology*, *58*(3), 368.

Wevers, M. (2019). Using word embeddings to examine gender bias in Dutch newspapers, 1950-1990. arXiv preprint arXiv:1907.08922.

Wille, E., Gaspard, H., Trautwein, U., Oschatz, K., Scheiter, K., & Nagengast, B. (2018). Gender Stereotypes in a Children's Television Program: Effects on Girls' and Boys' Stereotype Endorsement, Math Performance, Motivational Dispositions, and Attitudes. *Frontiers in psychology*, *9*, 2435.

Young, S. (2019, October 24). *FEMALE CEOS JUDGED MORE HARSHLY THAN MEN FOR WORKPLACE FAILURES, STUDY FINDS*. Independent.

https://www.independent.co.uk/life-style/women/women-ceo-judgement-gender-corporate-work-failure-ethical-study-a9169261.html

Yu, R. (2021). Unconscious preparation: Effects of prime visibility on semantic generalization of task priming. British Journal of Psychology.