Language and Gender:

A matched-guise study exploring linguistic stereotyping using voice morphing

Mattias Östling
Abstract

The matched-guise test, which was originally developed by Lambert et al. (1960), has been useful in determining covert attitudes towards accents, dialects, or languages. In this paper, the author presents a fresh take on matched-guise tests in sociolinguistics utilizing modern technology to digitally manipulate a female voice into a male voice. This makes it possible to create recordings in which the tone, intonation, stress, and personality traits are the same, yet the apparent guise is different, thus allowing us to ascertain attitudes towards the sexes as well. Using this method on 169 respondents of mixed sex, the author has shown that there are statistical differences between how the male and female guises are interpreted. When all collaborative and competitive aspects are combined, the female guise is considered more collaborative, while the male is considered more competitive. Contradictory to old stereotypes it would also seem as though men are considered less knowledgeable than women. Less sexist people also rate the male and the female guise more unevenly than more sexist people do. A less sexist person favours the female guise on collaborative variables, while a more sexist person favours neither guise.

Keywords:
sexism, stereotyping, matched-guise, voice morphing, sociolinguistics, collaborative speech, competitive speech
Acknowledgements

I express my sincere gratitude to my supervisor Mats Deutschmann for the multitude of useful comments, tips, ideas, and engagement through the writing process of this master thesis. Furthermore, my deepest thanks go out to Anders Steinvall who introduced me to the topic of sociolinguistics as well as all the support along the way. Also, I like to thank the respondents in my survey, who have willingly shared their time when filling out the survey. I would also like to thank Vetenskapsrådet in Sweden for ensuring that funding was available for the RAVE project which allowed me to have access to all the tools I needed for this thesis. Lastly, my dear fiancé Sari who has stood by my side, supported me, and inspired me throughout this process; thank you.
# Table of contents

1 **Introduction**........................................................................................................................................3  
   1.1 Introduction..................................................................................................................................3  

2 **Aim and Research Questions**...........................................................................................................4  

3 **Theoretical framework**.......................................................................................................................5  
   3.1 Performing Gender ..........................................................................................................................5  
   3.1.1 Hegemonic masculinity & femininity..........................................................................................5  
   3.2 Competitive and Collaborative Language .......................................................................................8  
   3.2.1 Hedges......................................................................................................................................8  
   3.2.2 Minimal responses......................................................................................................................9  
   3.2.3 Monologues and playing the expert..........................................................................................10  
   3.2.4 Verbal sparring.........................................................................................................................10  
   3.2.5 Topic choice..............................................................................................................................11  
   3.2.6 Turn-taking patterns & Interruptions.......................................................................................11  
   3.2.7 Questions..................................................................................................................................13  
   3.2.8 The Gender Similarities hypothesis vs Gender Differences model........................................14  
   3.2.9 Speech style and gender construction......................................................................................15  
   3.3 Linguistic Stereotyping..................................................................................................................15  
   3.4 Matched-guise Technique..............................................................................................................16  

4 **Material and Method**........................................................................................................................18  
   4.1 The Respondents............................................................................................................................18  
   4.2 The Script.......................................................................................................................................19  
   4.3 The Implementation.......................................................................................................................23  
   4.4 Rating Dimensions.........................................................................................................................25  
   4.5 Technical Issues.............................................................................................................................25  
   4.6 Ethical Dilemmas...........................................................................................................................26  

5 **Results and Analysis** ........................................................................................................................27  
   5.1 Data...............................................................................................................................................27  
   5.2 Data Summary ...............................................................................................................................38  

6 **Discussion** .........................................................................................................................................39  

7 **Conclusion and Suggestions for Further Research** ..........................................................................42  

8 **References**..........................................................................................................................................43  

9 **Appendix**...........................................................................................................................................48  
   9.1 Disney Text.....................................................................................................................................48  
   9.2 Gender discussion text....................................................................................................................51  
   9.3 Explicit Sexism Test.......................................................................................................................55
9.4 Extra Tables and Figures ........................................................................................................56
1 Introduction

1.1 Introduction

Decades of research in social psychology has shown that we draw on pre-existing attitudes and beliefs about stereotypes when forming initial impressions of others, and that these stereotypes have a deep impact on how we perceive the people we meet (Higgins & Bargh, 1987) (Macrae & Bodenhausen, 2001). More importantly for this essay, is the research done by Hay, Warren, and Drager, and others, that show that these stereotype-based categorizations also affect how we interpret and process speech (Hay, Warren, & Drager, 2006) (Niedzielski, 1999) (Johnson, Strand, & D'Imperio, 1999) and that language affects how individuals are judged in relation to e.g. intellect and empathy (Deutschmann et al., 2011; Cavalloaro & Ng, 2009; Fuertes et al., 2012).

The matched-guise test, which was originally developed by Lambert et al. (1960), has been useful in determining covert attitudes towards accents, dialects, or languages. Traditionally you would have a speaker or several speakers record controlled samples (for example one Cockney accent and one RP accent) who the research participants would then listen to and rate the ‘guise’ they hear on dimensions such as ‘intelligence’, ‘likability’, ‘honesty’, etc. Today we have the tools to use matched-guise tests to ascertain implicit attitudes towards males and females as well. Modern technology allows us to ‘morph’ or distort sound recordings so that a male speaker sounds like a female speaker or vice-versa, so while before it was impossible to design a matched-guise test investigating gender that adequately fooled the participants, it is now possible to record a controlled sample. This sample can then be digitally manipulated so that the voice of one or more participants is altered and thus we can produce a manipulated text where the apparent genders of the participant/s are altered but in which the tone, intonation, stress, and personality traits are the same. This method is still new in sociolinguistics and only a handful of studies have been conducted in which voice morphing is used. One such example is a series of studies where Deutschmann et al. (Forthcoming) used the matched-guise method with voice morphing to develop and explore experimental pedagogic approaches for raising gender awareness. Other than that, however, very little research with regards to how gender stereotypes affect people’s interpretation of language input using matched-guise tests has been conducted.

The aim of this study is to further explore the method of voice-morphing to investigate how linguistic stereotyping works with regards to gender. Do people interpret the same text differently depending on whether they think they are listening to a man or a women, and what aspects differ? The presentation of this thesis is organized into six main sections below, starting at aim and research questions and ending with a discussion of the data and a conclusion.
2 Aim and Research Questions

The aim was to explore linguistic stereotyping in regards to gender. The research questions are stated below.

- Are there statistically significant differences between how people interpret the male voice sample compared to the female voice sample, and what specific differences as regards language features and personality traits can be elucidated?
- How does the sexism of the respondent affect how they perceive the guise?
- In what way does the gender of the respondent impact the way they perceive the guise?
- What potential differences can we find when comparing how the guises are perceived in a recording with a competitive speech style to a recording with a collaborative one?
3 Theoretical framework

3.1 Performing Gender

When Candace West and Don Zimmerman published “Doing Gender” (1987), it was a conceptual breakthrough that has since been confirmed and supported in over 20 years of sociological scrutiny and research (Butler, 1990; Messerschmidt J. W., 2009; Risman, 2009; Wallenberg & Thanem, 2016; Pecis, 2016). The principles are that when discussing gender and sex we must realize that while they are often used as almost perfect synonyms they actually work in distinctly different ways. While a person’s sex is determined by their genetic make-up, it is not necessarily the case when referring to gender. A person does not ‘possess’ a gender the same way they do a sex. Instead, gender is something that is constructed in interactions with others; it is culturally defined (Butler, 1990) (Messerschmidt J. W., 2009). Socially and culturally defined membership of a group (male or female) is used to accept or reject behaviors as either masculine or feminine, that is, gender category is something into which all people are placed in social situations. Their actions and behaviors are then judged based on their biological sex (Messerschmidt J. W., 2009) (West & Zimmerman, 1987). Dozier (2005) and Messerschmidt (2009) found that people can present an easily recognized biological sex, but may diverge from this group when performing their gender, that is, perform a gender that does not conform to their sex. Further, the meaning assigned to a person’s gender is dependent on other people’s perceptions of that person’s biological sex. In other words, in social interactions we see gender and sex as two inseparable parts of a whole, this means that people who perform a gender that does not conform to their sex are often punished socially (e.g. feminine men or masculine women) (Butler, 1990) (Messerschmidt J. W., 2000). Importantly, gender is often an unconscious action. Martin (2003) and Messerschmidt (2009) found that while accountability of one’s actions and behaviors can enforce a conscious choice of gender performance, most often people’s ultimate choice of behavior is an unconscious choice.

3.1.1 Hegemonic masculinity & femininity

Hegemonic masculinity, a concept popularized by R.W. Connell in the late 80s and early 90s deals with the ‘ideal’ form of masculinity; the image of masculinity which is culturally normative and something everyone must relate themselves to (Connell, 1987) (Connell, Masculinities, 2005). The characteristics and behaviors of hegemonic masculinity are traits that all males are encouraged to internalize into their own character and personal code. These characteristics include: competitiveness, stoicism, courage, toughness, risk-taking, adventure and thrill-seeking, violence and aggression, and achievement and success (Donaldson, 1993). As the name might imply,
hegemonic masculinity is only dominant in its relation to other masculinities, that is, subordinate forms of masculinity. Homosexual men may perhaps be the most salient form of subordinate masculinity in Western society today. Negative attitudes towards this subordinate masculinity are evidenced in street violence, cultural exclusion, legal violence, and economic discrimination against this group (Connell, 2005). This also becomes evident in the treatment of effeminate men who are often culturally scorned as well (Butler, 1990) (Messerschmidt J. W., 2000). Importantly, while hegemonic masculinity is the ideal, it is not the numerically most common form of masculinity. Often it takes the form of a fantasy character such as Rambo or Batman (Connell, 1987) who possesses traits and behaviors that may not be a common pattern in real life. Indeed, characters can act as symbols that have cultural authority to influence people’s perception of ideal masculinity, ideals which no one can fully live up to. In other words, they act as an unattainable goal that men should strive for (Connell & Messerschmidt, 2005).

When Connell first talked about hegemonic masculinity, she noted that there was no hegemonic femininity since all forms of femininity were constructed in subordination of women to men. She argued therefore that there was no femininity that held the same role as hegemonic masculinity does for men (Connell, 1987). Instead, she called it emphasized femininity and explained it as below.

One form [of femininity] is defined around compliance with this subordination and is oriented to accommodating the interests and desires of men. I will call this ‘emphasized femininity’. Others are defined centrally by strategies of resistance or forms of non-compliance. Others again are defined by complex strategic combinations of compliance, resistance and co-operation. (pp. 184–185)

According to Connell, there are multiple femininities (compliance/non-compliance to subordination) but does not mention or further analyze femininity in isolation since most of her aim here was to analyze masculinities. Schippers notes that, “although emphasized femininity is central to men’s dominance over women, it is not the only mechanism for ensuring men’s domination over women” (2007, p. 87). She goes on to explain that by only considering emphasized femininity without a hegemonic femininity there is no adequate apparatus for distinguishing subordinate masculinities from femininities (p. 88), a notion echoed by others (Halberstam, 1998; Lorber, 1998; Martin P. W., 1998). Indeed, it was not until 2003 that Pyke and Johnson tried to use Connell’s framework of hegemony to develop a definition of hegemonic femininity. However, while they called it hegemonic femininity, they were quick to explain that they were “not arguing that hegemonic femininity and masculinity are equivalent structures. [...] Whereas hegemonic masculinity is a superstructure of domination, hegemonic femininity is confined to power relations among women” (Pyke & Johnson, 2003, pp. 50-51). In other words, while hegemonic femininity is the same as hegemonic masculinity in the sense that both are the dominant gender
constructions, they are also dissimilar since hegemonic femininity does not hold dominance over masculinity. Schippers defined it as:

hegemonic femininity consists of the characteristics defined as womanly that establish and legitimate a hierarchical and complementary relationship to hegemonic masculinity and that, by doing so, guarantee the dominant position of men and the subordination of women. (2007, p. 94)

In other words, she argues that hegemonic femininity is a collection of womanly traits that maintain men in the dominant position and women in subordination through the idolized relationship between masculinity and femininity. Women who go against idolized norms of femininity are seen as refusing to complement hegemonic masculinity in relation to subordination/dominance. For example, if a woman is not sexually available, then a man cannot have the sex he is entitled to, and if a woman is aggressive, she is not submissive to a man’s violence.

According to Schippers, when relating hegemonic femininity to other subordinate femininities, the other femininities are pariah femininities (p. 95). She argues that they should be referred to pariah femininities because they are not necessarily subordinate to other femininities, but rather contaminate the hegemonic relationship between femininity and masculinity. For example, a woman who rejects the advances of a man is “a lesbian, a “slut” [...] or “cock-teaser”, “a bitch.” (Schippers, 2007, p. 95). Examples of how these femininities and behaviors are undesired or stigmatized can be found in Messerschmidt’s empirical research on adolescent gender and violence (Messerschmidt J. W., 2003). What this means is that pariah femininities are women who embody characteristics of hegemonic masculinity (aggression, competitiveness, woman-desire of a feminine object (lesbian), authority (bitch), etc.) (Schippers, 2007).

In summary, gender is a continuum rather than a dichotomy. On the one hand, you have hegemonic masculinity, and on the other hegemonic femininity, and between these two poles there are a multitude of subordinate masculinities/femininities. Importantly, what we see from Messerschmidt’s (and others’) research is that the social implications of gender production behaves differently depending on the biological sex of the performer. For example, a woman embodying masculine traits is socially stigmatized just as a man embodying feminine traits is socially stigmatized. They are stigmatized because their gender does not conform to their sex. In section 3.2 below a thorough account of how hegemonic gender and gender production pertains to language production will be presented.
3.2 Competitive and Collaborative Language

According to sociolinguists such as Trudgill, regardless of social class, or age, when men talk they have a statistical tendency to ignore other men’s utterances or disagree with them, while women tend to acknowledge and build on them. When men converse they tend to seek power while women seek solidarity (Coates, 2004, p. 126). On the topic of the different speech styles Cheshire and Trudgill writes:

It seems clear that, other things being equal, women and men do have a preference for different conversational styles. Women - in most western societies at least - prefer a collaborative speech style, supporting other speakers and using language in a way that emphasizes their solidarity with the other person. Men, on the other hand, use a number of conversational strategies that can be described as a competitive style, stressing their own individuality and emphasizing the hierarchical relationships that they enter into with other people. (Cheshire & Trudgill, 1998: 3)

These two speech styles described by Cheshire and Trudgill are characterized by some key linguistic features. For the cooperative speech style these are hedges and minimal responses. And for the competitive speech style these are monologues, interruptions, playing the expert, and verbal sparring. Some other differences include topic choice, floor space, turn-taking patterns, and what types of questions are asked (Coates, 2004).

3.2.1 Hedges

Hedges are filler words used to save face, for either the speaker or other participants in the interaction. They are normally used when talking about intimate topics or topics that could be too face-threatening (Coates, 2004, p. 129). Since they mitigate utterances that would otherwise be face-threatening they can be seen as promoting or encouraging discussions and dialogues (Coates, 2004). For example, utterances such as *umm, sort of, I mean, probably, well, could*, etc. could be considered hedges depending on the circumstances around the utterance. Coates brings up the following example from her studies in her book Women, Men, and Language (hedges underlined):

Meg: but I did see what amounted to *sort of* chest hair, black, she’s a very dark sort of dark skinned and sallow complexion and a l- I mean l - I mean I hope I’m just reporting this without any edge to it. *you know*, so *I mean I probably*-.

(Coates, 2004, p. 129)
3.2.2 Minimal responses

Minimal responses are short utterances, such as *mhm, yeah, or okay*, well-placed in-between chunks of coherent syntactical and semantic speech. They are used to show interest, participation, support, and recognition of a topic’s development and the different stages of a conversation (Minchin, 2007). Noteworthy is that, for instance, a delayed *mhm* could be seen as sarcastic and challenging, thus all *mhm’s, etc.* are not minimal responses.

Minimal responses have traditionally been hard to define (see e.g. (Zimmerman & West, 1975) (Woods, 1988), but in 1995 Reid characterised minimal responses by six criteria.

1. They must be brief since they are only intended to be indicators of participation in the conversation.
2. They must be made in response to another speaker. This ensures they really are a ‘response’.
3. They contain little semantic content since they serve only to indicate participation or, at most, agreement.
4. They do not generally interrupt the flow of speech from the first speaker.
5. The second speaker, that is, the one who produces the minimal response, is not attempting to take over the floor.
6. Each verbal minimal response constitutes either a completed or continuing intonation unit.

(Reid, 1995)

She also added that non-verbal indicators of a participant showing interest, support, participation, etc. such as head nods, could also be considered minimal responses, although whether you recognize this idea or not is of little importance as this paper only deals with verbal cues. The following example comes from the end of a topic:

(five women are discussing parents' funerals)

Gina: *there's two things aren't there/ there's the the other people like your mother or father who's left and or or siblings/ and there's also how you feel at that time about (.) the easiness of going (to the funeral)/

Gina: *mhm/

Mary: *mhm/

Bea: *mhm/

Meg: *mhm/

Sally: *yeah/

(Coates, 2004, p. 129)
3.2.3 Monologues and playing the expert

Monologues and playing the expert often go hand in hand in conversations as the former opens up the latter, and the latter usually entails the former. A monologue is when a participant holds the conversational floor space for an extended period. And playing the expert is when a person talks about a subject in which he is an expert. This does not necessarily mean that when conversing a particular person always takes the role of the expert, on the contrary it would seem that participants take turns at being the expert in the conversation and holding a monologue (Coates, 2004, p. 134). The following exert from Coates shows how playing the expert looks in conversations (minimal responses in italics):

cos you know we've got BT internet at home (mhm) and I've set it up so that . um through the BT internet WAP portal so that Kate can read . her email that she gets . um on her phone (oh right) which is qui- which is quite useful if you're kinda not behind a computer but I was musing the other day on . on how funny it is that the sort of graphics you get on WAP phones now . is like you use to get on the ZX81 (yeah) and every- everything's having to adapt to that kind of LCD based stuff (that's right) um computers have got to the point they've got to . and now we've gone all the way back with WAP technology...
(Coates, 2004, p. 134)

In this example, we clearly see how one speaker talks for an extended stretch of time and is holding the role of an expert in the conversation. Importantly, we can see the minimal responses by the other participant signalling interest and that he wants the speaker to go on.

3.2.4 Verbal sparring

Verbal sparring can be seen as the opposite of monologues and playing the expert and occurs in competitive speech. Verbal sparring is rapid-fire turn taking, often in which the participants are disagreeing in some way. The following example from Pilkington shows this in conversations:

Ray: crate!
Sam: case!
Ray: what?
Sam: they come in cases Ray not crates
Ray: oh same thing if you must be picky over every one thing
Sam: just shut your fucking head Ray!
Ray: don't tell me to fuck off fuck ( . . . )
Sam: I'll come over and shut yo-
Jim: yeah I'll have a crate of apples thanks [laughingly using a thick sounding voice)
Ray: no fuck off Jim
Jim: a dozen...
Dan: shitpicker! [amused]
(Pilkington, 1998, p. 265)

In this excerpt from Pilkington we see Sam disagreeing with Ray, Ray with Sam, Jim disagreeing with Ray, and Dan criticising Jim, all with short one sentence utterances. While the language is quite explicative and vulgar, the dialogue includes laughter, amusement and joking indicating that it is a friendly sparring session and not an actual quarrel.

### 3.2.5 Topic choice
A great deal of previous research show that collaborative speech and competitive speech usually discuss different topics in same-sex groups (e.g. Aries & Johnson, 1983; Seidler, 1989; Pilkington, 1998). The evidence shows that people with a collaborative speech style prefer to talk about emotions or people, while those with a competitive speech style like to talk about impersonal things, such as current affairs, travel, technology, and sports. Further, when the conversations become more personal they tend to be about achievements or drinking habits rather than feelings or people (Coates, 2004).

### 3.2.6 Turn-taking patterns & Interruptions
According to Coats, in female same-sex group discussions, the ‘one-speaker talks at the time’ rule seldom exists. Coates termed this way of talking a *jam session* (2004, p. 131), because just like jazz musicians jamming, they get together "for the spontaneous and improvisatory performance of talk, usually for their own enjoyment" (Coates, 2004, p. 131). What this means is that the floor space is often open for all participants in a discussion, rather than taking turns.

Tina: she didn't actually TEACH them but
Lyn: ................................................................................................................

Tina: She just provided a model
Lyn: Provided a model
(Coates, 2004, p. 131)

In the example from Coates above, we see how Lyn helps construct the sentence together with Tina, and how Tina then uses the same construction (‘provided a model’) to signal that she accepted Lyn’s contribution. Below is a slightly more intricate example of this:

[talking about child abuse]
Bea: I mean in order to accept that idea you're having to completely change your view of your husband.
Mary: mhm. completely review your [view of your husband/]

(Coates, 2004, p. 132)

In this dialogue, both co-construction and overlapping words can be seen. What is important to recognize here is that these overlapping words do not act as interruptions, but rather serve as a way to signal interest and help the co-conversationalist to form utterances.

According to Coates (2004) and Talbot (1992), turn-taking pattern in male same-sex talk is vastly different, however. Where women prefer a jam session, men prefer monologues and extended periods of time where one speaker holds the conversational floor. Thus, because of this, in mixed-sex talk where women use collaborative speech methods (such as co-constructing utterances) men might interpret it as interruptions. Talbot (1992) has an excellent example of this in which two heterosexual couples are having a dinner. One of the men starts telling a story and his partner co-constructs parts of the story; it finally ends in the man telling his partner that ‘I wish you'd stop interrupting me!' because he interpreted her collaborative methods as interruptions. He felt that she infringed upon his solo floor space.

Interruptions are speech functions used when someone infringes on a co-participant’s floor space. That is, someone breaks in on a person speaking in order to take the floor (Minchin, 2007). West and Zimmerman define an interruption as a ‘deep incursion into the turn-space of a current speaker’ (West & Zimmerman, 2011). This definition is slightly vague because it opens up the possibility of minimal responses being interpreted as interruptions since not all interruptions are intended or recognized as attempts to claim the floor. Jack Bilmes defines interruptions a bit more concretely by noting that only when the person being interrupted reacts to the interruption or when the person interrupting apologizes for interruption can it be considered a real interruption (Bilmes, 1997). However, this says nothing about attempted interruptions that may go unrecognized by either participant(s) if it is a heated discussion. West and Zimmerman argued that there are three notable reasons why people use interruptions. First they argued that interruptions can be used as a form of dominance or control by the speaker interrupting; second they can be used as an actual control device by which the person interrupting tries to disorganize the speaker’s construction of a topic; and third, interruptions can be used to steer the topic away from e.g. sensitive topics (such as, talking about a divorce, passed away child, etc.) (West & Zimmerman, 2011).

An example of interruptions in dialogues can be found in the gender discussion transcript used for the competitive recording in this paper (see appendix 2). In it,
John is talking about how research conducted on language patterns of genders simply seek evidence to confirm stereotypes, when Jane interrupts him.

**John:** Yea researchers simply seem to seek evidence to confirm stereotypic behaviour and the question remains if we should even include this research in our teachings [...]  

**Jane:** [...]Oh come on! That’s ridiculous I mean I would say the total opposite. The role of sociolinguistics research is to expose structural frameworks and how they are reflected in language.

### 3.2.7 Questions

Because playing the expert is a rare occurrence in same-sex female conversations according to Coates, questions that act as information-seeking devices are uncommon. Instead, questions in same-sex female groups are often hedges, devices to introduce new topics, to check the views of other conversational partners, or to instigate stories according to research done by Coates (2004, p. 130).

[Talking about doctors getting younger]  
Karen: I suppose if you're ill you don't care do you?  
Pat: I suppose not, but there are um- there are limits aren't there?  
(Coates, 2004, p. 130)

In this example from Coates, we see how Karen and Pat use tag questions to involve the other participant in the conversation. Sometimes questions in collaborative talk are used to check a consensus or affirm a view while not necessarily requiring or expecting an answer. In other words, the questions are sometimes rhetorical. In the two examples below from Coates we see this:

[end of topic Relationships]  
Liz: it's strange isn't it? the life some people lead/  
(Coates, 2004, p. 131)

[gossip about school friends]  
Claire: but they're so stupid right? cos then- cos Nina said...  
(Coates, 2004, p. 131)

When questions are used like this it does not matter if they are in the middle or at the end of an utterance, the purpose is merely to check the consensus or world view and no answer is expected.

Questions in same-sex male groups on the other hand are the exact opposite in the sense that they are often used to promote a person taking the role as an expert according to Coates. That is, questions are used to seek information. However, this does not mean that just because an information seeking question is
posed the other participant immediately becomes the expert in the conversational topic. Below is an example in which an information-seeking question is posed but the respondent does not have the knowledge to play the role as the expert.

Peter: what else do they use it for apart from the deaf? or do they have other applications- I don't mean the deaf, I mean the dumb, do they have other applications?
Rob: well they didn't develop it for the dumb, I can't remember why they did develop it, um - I don't know
(Coates, 2004, p. 135)

Sometimes questions are also asked to introduce topics in which the speaker rather than the addressee is an expert. Below is an example of this:

Rob: do you know about the Pennsylvania experiment?
Peter: no tell me about it
[Rob proceeds to tell Peter about it, playing the role of the expert)
(Coates, 2004, p. 135)

In this example, Rob relies on illocutionary force (the intention to perform a certain illocutionary act) (see: (Alston, 2000)) and pragmatics to effectively ask Peter if he wants Rob to tell him about the Pennsylvania experiment. Peter’s “no tell me about it”, demonstrates that he interprets Rob’s statement as such.

3.2.8 The Gender Similarities hypothesis vs Gender Differences model

It is important to note that the research conducted does not represent individual reality, but rather statistical tendencies. Individual production of gender and language can vary depending on different factors, such as personality and context. In other words, a person’s production of language depends on their constructed gender. So when we talk about how ‘men tend to ignore other men’s utterances or disagree with them’, it is essential that we understand that these are not representations of all individuals behaviour, and not all the time, it is a statistical tendency.

We must consider both the ‘gender differences’ model and the ‘gender similarities’ hypothesis. The gender similarities hypothesis holds that men and women (as well as boys and girls) are similar on most but not all fundamental psychological variables. That is, they are more alike than they are different (Hyde, 2005). Meanwhile, the gender differences model suggests that there are noticeable differences between how women and men are in many different variables. Gilligan (1982), suggested in her book in a different voice: Psychological theory and women’s development that women use a moral “voice” of caring and men use a voice of justice. However, her argument for large gender differences is disproven by meta-studies
(Jaffee & Hyde, 2000; Thoma, 1986; Walker, 1984). Gilligan’s ideas have since spread through American culture, and the idea of large differences between the sexes has helped maintain and create stereotypes, such as women are nurturing and men are not (Hyde, 2005). In other words, we can note that some researchers have found that there are more similarities between the genders than differences. This means that while some data shows that men’s language is competitive, and women’s collaborative, we must recognize that other data shows the opposite, that is, that both men and women can be competitive and collaborative, indeed, that individuals can use more than one speech style. For the purpose of this paper, however, what is of most importance is how men and women are perceived to talk and what stereotypes influence that perception.

### 3.2.9 Speech style and gender construction

If we consider some of the traits of hegemonic masculinity detailed in section 3.2. above, we can quite easily see how some of them correlate with competitive speech. Similarly, if we consider hegemonic femininity we can see how collaborative speech correlate with those traits as well. The behaviours and characteristics of hegemonic masculinity includes: competitiveness, violence and aggression, and achievement and success (Donaldson, 1993), and all these traits are familiar with the competitive speech style. Interruptions are both competitive and aggressive. Playing the expert could be seen as achievement/success and competitiveness, verbal sparring is both aggressive and often competitive, and we see how taking a great deal of the floor space relates to both aggressiveness and success (everyone has to listen to what they have to say). Further, if we consider the type of questions that are used in competitive speech we see that those are questions that lead to someone playing the expert. The cooperative speech style, on the other hand, includes traits relating to hegemonic femininity. For example, hedges are used to mitigate statements and express doubt, which relates to submission, as does taking little floor space. Minimal responses and the function of questions in cooperative speech relates to attentiveness.

### 3.3 Linguistic Stereotyping

The linguistic stereotyping hypothesis began roughly half a century ago with Lambert and his colleagues. They proposed that even brief samples of speech varieties (e.g. accent, intonation, minority language, etc.) associated with social groups can affect how individual people are seen (e.g. intelligent, criminal, submissive, dominant, etc.) (Lambert, Hodgson, & Fillenbaum, 1960) (Bradac, Cargile, & Hallett, 2001). This means, for example, that although Jamaicans’ opinion towards Jamaican Creole is fairly positive they still statistically regard an English speaker to be more intelligent than a
Creole speaker (Jamaican Language Unit. (2005)). That is, people attribute characteristics and traits to individuals that they stereotypically ascribe to the social group to which they belong, and speech is a major trigger (Lippi-Green, 1997) (Johnson F. L., 2000). People’s knowledge about stereotyping and how they unconsciously rate and judge people differently based on their sex, ethnicity, age, etc. is often lacking; it is difficult to maintain knowledge about all stereotypes in all contexts. Further, even when knowledge about stereotypes is present, it can be difficult for people consciously to try to act and judge in ways where they do not treat people differently based on certain features. These stereotype-based categorizations deeply affect how people interpret and process speech and language (Hay, Warren, & Drager, 2006; Niedzielski, 1999; Johnson, Strand, & D’Imperio, 1999). This encourages a shift from what dissimilarities there are between different social groupings, to what beliefs there are about the language behavior of different social groupings (Edlund et al. 2007) (Deutschmann et al. Forthcoming).

With specific focus on gender stereotypes, it has been shown that students evaluate a text more favorably if they think a male teacher has written it (Abel & Meltzer, 2007) (Centra & Gaubatz, 2000). Furthermore, both female and male teachers have a higher chance of positive reviews and evaluations if they conform to gender stereotypes rather than if they deviate from them (Basow, 1995) (Deutschmann et al. 2011). This suggests that the results in this paper should show that the male guise is rated as more knowledgeable and that there will be a difference between how the guises are perceived in the different recordings.

### 3.4 Matched-guis Technique

The matched-guis test was originally developed by Lambert et al. (1960) to indirectly ascertain covert attitudes towards English and French in Montreal. Traditionally, you would have a speaker or several speakers record controlled samples that would then be rated on different aspects. The recorded speaker is often someone with native-like speaking ability in the different languages or variations if it is one speaker that is recorded for the different guises (Agheyisi & Fishman, 1970; Anisfeld & Lambert, 1964; Lambert W., 1967). This is naturally concealed from the respondents so they are unaware that the different guises are the same person. The respondents’ evaluation of the speaker’s intelligence, likability, bluntneess, attributes, etc. are then entirely based on the accent or language. The main idea behind this method is that there is a control for every other variable in the study, such as intonation, script, reading speed, voice and quality. This way, we know that any differences in how the respondents evaluate the guises are only because of the accent or language (Agheyisi & Fishman, 1970). This means that if there is any uniformity in how the respondents rate the guises, we can accurately determine what status a language or accent has and if any stereotypes associated with it is reflected in their responses. Using respondents as their own
controls is an inexpensive and effective way that often reveals that people rate the same person differently based on their language or accent. Because of this, this method is still widely used today in social psychology, sociolinguistics, business research, and medicine (Buchstaller, 2006; Cargile & Giles, 1998; Dixon, Mahoney, & Cocks, 2002; Lawson & Sachdev, 2000; Cargile A., 1997; Carson, Drummond, & Newton, 2004).

With modern technology, it is possible to manipulate voices in recordings. This, for example, allows a man to sound like a woman or vice-versa. Previously in matched-guise studies there was always the issue with two speakers likely having different ways of speaking, speed, intonation, etc. or the same speaker being recognized as one and the same or reading the different accents or languages differently (Tsalikis, DeShields, & LaTour, 1991). But now it is possible to have a single person record a controlled sample and then digitally manipulate it to a female or male voice, creating two identically read passages with two distinctly different voices. This opens the possibility to study gender and language in ways that have not been possible before, with no risk of participants recognizing the speaker as the same person or being affected by different tempo or intonation.

Using modern technology it is also possible to add another dimension to the recording. With software such as The Sims or Second Life one can create avatars (a character in the virtual world) that represents how a person with a particular accent or language stereotypically looks, e.g. it is possible to use a black avatar to represent a speaker talking African American Vernacular English, or a male/female avatar to represent a man/woman. Deutschmann et al. argued that this combination of digitally manipulated voice along with an appropriate avatar creates a convincing illusion. However, they go on to write that they did not get the effect they hoped for from their respondents, partly because there were too many distractions (Deutschmann, Steinvall, & Lagerström, Forthcoming). So, if one were to use virtual worlds such as Second Life it is best to leave the respondents out of the world and simply let them view and listen to the character since any data relying on respondents’ performance in a virtual world is simply too inconsistent (Deutschmann, Steinvall, & Lagerström, Forthcoming).
4 Material and Method

4.1 The Respondents

The data collected for this essay were gathered from four controlled groups and one uncontrolled group. With the ‘controlled group’, aspects such as age, gender, and nationality were known before they responded to the guises, and with the ‘uncontrolled group’, these aspects were not known before. The first controlled group comprised teacher students studying at a University in a medium sized city in Sweden who had also completed a course on sociolinguistics dealing with speech styles. They are referred to as ‘Group 1’ in this essay, and comprised 24 individuals in all, 15 male and 9 female. How a group versed in sociolinguistics might affect how they perceived the guises is discussed further in the data section. The second controlled group was upper secondary school students in a different medium sized city in Sweden, referred to as ‘Group 2’ in this essay, and comprised 15 individuals in all, 7 male and 8 female. The third controlled group was university students studying English in a medium sized city in Sweden, referred to as ‘Group 3’ in this essay, and comprised 13 individuals in all, 7 male and 6 female. The first uncontrolled group, referred to as ‘Group 4’, comprised 16 individuals, 9 male and 6 female. It consisted of respondents from all over the world found through the social media Facebook and Reddit.com/r/samplesize. For the Facebook respondents, friends on Facebook were asked to post the link to the survey informing their readers that it was to help them help a friend, that is, no one knew for whom they were really doing the survey. However, this posed a problem of snowball/skewed sampling since it is likely that friends of friends share similar views with each other, otherwise they would probably not be friends and/or assist with the study. This problem was not as direct for the respondents from Reddit since the respondents there did not necessarily have a connection to each other. Further, since Group 1, 2, 3, and 5 provided data initially, it was possible to compare and see if any strong deviations occurred from average response values. Moreover, since there were very few respondents (16 total) from the Facebook (6 respondents) and Reddit group (10 respondents), snowball/skewed sampling was not an issue. The reason so few respondents were found through social media may be that the survey was too long. Roughly 75-80% of the respondents stopped the survey before finishing it. Lastly, the fifth group was purchased from SurveyMonkey1. SurveyMonkey found random people within the UK with an age limit of 45 aiming for a 50/50 split in males/females, who then respondent to the survey. This group comprised 101 individuals in all, 48 males and 53 females. Below you find a table detailing the relevant information of the groups, including ‘recording sample’ which is elaborated on in section 4.2.

---

1 https://sv.surveymonkey.com/
Table 1, Respondent group information

<table>
<thead>
<tr>
<th>Group</th>
<th>Details</th>
<th>Males</th>
<th>Females</th>
<th>Total</th>
<th>Recording sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>Teacher students</td>
<td>15</td>
<td>9</td>
<td>24</td>
<td>Disney</td>
</tr>
<tr>
<td>Group 2</td>
<td>U-secondary students</td>
<td>7</td>
<td>8</td>
<td>15</td>
<td>Gender</td>
</tr>
<tr>
<td>Group 3</td>
<td>English students</td>
<td>7</td>
<td>6</td>
<td>13</td>
<td>Gender</td>
</tr>
<tr>
<td>Group 4</td>
<td>Facebook &amp; Reddit</td>
<td>9</td>
<td>6</td>
<td>15</td>
<td>Both</td>
</tr>
<tr>
<td>Group 5</td>
<td>SurveyMonkey</td>
<td>48</td>
<td>53</td>
<td>101</td>
<td>Disney</td>
</tr>
</tbody>
</table>

‘Recording sample’ found in column six describes which of the two recording samples the group listened to, more on this will be presented in section 4.2 and 4.3 below.

The reason such disparate groups were sampled was because of the difficulty of gaining access to contexts where a group of people have 15-30 minutes to spare. Especially if you take into account the possibility of respondents looking at each other’s monitors, which might taint the data. Therefore, Group 2 was chosen because they were upper-secondary students so it was possible to incorporate the survey into their regular education. Group 1 and 3 were chosen because the data accumulation for this paper is also a part of ‘Raising Awareness through Virtual Experiencing’ (RAVE) a larger research context founded by Vetenskapsrådet (The Swedish Research Council) in Sweden. The RAVE project gave access to the respondents found in Group 1 and 3. Group 4 was chosen to increase the number of respondents; the idea was that by trying to get respondents from all over the world it would be possible to bolster the results with more data. Finally, when the results for Group 4 were showing signs of inadequacy in the number of respondents, Group 5 was purchased through SurveyMonkey that were able to provide the adequate number of respondents.

4.2 The Script

Audio samples of two discussions between two linguists were (see appendix 1 and 2) recorded and digitally manipulated. They were first recorded using Avid Pro Tools HD 12.0.0 and then edited in the same software. Pitch shifting was processed with X-Form (Rendered Only) using Elastic Audio properties in Pro Tools. See picture 1, 2, and 3 for the relevant settings.
In one of the samples a more collaborative speech style was used (Appendix 1, henceforth “Disney discussion”), and in the other a more competitive speech style (appendix 2, henceforth “Gender discussion”). This was primarily to find out how participants rate (see 4.3 and 4.4 below) a speaker who go against the stereotypical norm (woman using a competitive style/man using a collaborative), and how they
rate (see 4.3 and 4.4 below) a speaker who adheres to the stereotypical norms (woman using a collaborative style/man using a competitive). Further reasons as to why two samples were used were to circumvent limitations of matched-guise tests. One of the limitations Lambert brings up (1972) is that there is a possibility the recording encourages the use of stereotypes, that is, rather than creating a balanced accent/dialogue/sample/etc., a much more stereotype based accent/dialogue/sample/etc. is designed in order to evoke stronger reactions from the respondents. Therefore, by recording a sample of each style we navigate around this limitation, since for one recording the female guise will have their stereotypical speech style exposed and for the other recording the male guise will have their stereotypical speech style exposed.

The topic of the recordings was related to gender; the collaborative one was about princesses in Disney movies, and the competitive recording was about how gender research affects stereotypes. The topics were related to gender so that Group 1 and Group 2 could properly be deceived into thinking the conversation in the survey was relevant to their standard education, that is, the survey was made to be seemingly part of their normal course package. In other words, if the topics were e.g. politics, there might be a risk that the respondents would think that there were ulterior motives behind having the survey during regular scheduled teaching, since that would not be a part of their normal course package. This was essential, since our aim here was to elicit their impressions of the recordings, not for them to count the number of e.g. interruptions. The main difference between the scripts was turn-taking patterns and contradictions. In the competitive text, both participants are trying to play the expert by taking long turns and asking information seeking tag questions to encourage the participant holding the floor instead of questions used to hedge and check the view of the participant (6 information seeking tag questions from the rated guise). They also contradict each other a great deal. In the collaborative text, both participants are holding the floor for shorter amounts of time and co-constructing sentence. Interruptions, floor space, minimal responses, and hedges were kept moderately even between the two discussions in order to make them seem more natural, with the main difference being what kind of questions are asked (information seeking for the competitive, and mitigating/view checking for the collaborative). In both recordings, Jane holds the floor 51% of the time. Below you find some excerpts from both texts.
(Disney discussion)

John: But from what I gather, it’s improved in recent years uh. The more recent movies seemed to be more balanced, at least some of them.

Jane: hmm yea, but there’s other aspects than floor space to look at [...] 

John: [...] yeah yea, that’s true. But, uhm, you know, they did actually look at other aspects as well, such as the number of male and female characters I think uh,

(Gender discussion)

John 11: Well well, yea, but, but the question remains, right? Whether the research is representative or not. I mean I I I think loads of studies that fail to find, you know, what they’re looking for they, they are never published you know? They they are not interesting and and

Jane 11: oh come on now, I mean that is really a cynical view. I mean that is a cynical view of research. I mean we have to assume that researchers have some integrity to publish their results

In the table below you can see the differences between the texts in table form.

<table>
<thead>
<tr>
<th></th>
<th>Disney Discussion</th>
<th>Gender Discussion</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Length</strong></td>
<td>4:21</td>
<td>6:06</td>
</tr>
<tr>
<td><strong>Floor space</strong></td>
<td>51%</td>
<td>51%</td>
</tr>
<tr>
<td><strong>Style</strong></td>
<td>Jam session</td>
<td>Playing the expert</td>
</tr>
<tr>
<td></td>
<td>Agreeing</td>
<td>Contradicting</td>
</tr>
<tr>
<td><strong>Interruptions</strong></td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td><strong>Hedges</strong></td>
<td>19</td>
<td>24</td>
</tr>
<tr>
<td><strong>Minimal responses</strong></td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td><strong>Information seeking questions</strong></td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td><strong>Mitigating questions</strong></td>
<td>3</td>
<td>0</td>
</tr>
</tbody>
</table>

This meant that while the texts were clearly different in style, they were not so different as to make it obviously constructed. If the guise was overly competitive or overly collaborative, it is likely that it would sound unnatural or weird.
4.3 The Implementation

In order to unearth the demographics of the respondents, a series of simple direct questions were added in the beginning of the survey, covering age, gender, and nationality. In addition, to find out whether sexism is a variable that affects respondents' impressions of the text, two tests with focus on gender were shortened and combined and included in the survey. The first eight questions came from Modern Sexism Scale by Ekehammar et al. (2000) and the remaining six were taken from the Ambivalent Sexism Inventory (Glick & Fiske, 1996). These models were chosen because they are established tools in social psychology for measuring sexism. In order to circumvent the issue of priming the respondents before they listen and respond to the recording, the sexism test was placed at the end of the survey. In other words, if the respondents filled in the explicit sexism test before the video and response survey they would be implicitly prepared for gender issues. Their exposure to one stimulus (explicit sexism test) would influence their response to another stimulus (the recording). Avoiding priming could also have been accomplished by writing dummy questions, however, since priming is not an issue once they have responded to the recording, it could be placed after the recording. This made it possible to remove 28 or more questions that would have cluttered the survey and increased the odds of the respondents becoming bored or disengaged, resulting in incomplete responses.

The recording was located after the demographic elicitation but before the sexism test. Importantly, the response survey was kept separate from the recording so that the respondents could not go back and listen to the recording after they had read the questions, or be distracted by the questions while listening. When each respondent in the uncontrolled group started the survey, they were first randomized a style (competitive recording or collaborative recording), and then randomized a 'morph' that is a digitally manipulated recording (the male or the female). This ensured equal chances for each recording which would yield the best chances of getting an even split between the four recordings. For the controlled groups each group was assigned a specific recording (competitive or collaborative) and then randomized a morph (male guise or female guise). Importantly, the respondents could never ‘go back’ in the survey after they had progressed to another page, so they could never go back and listen to the recording again, for example. Below you find a list detailing how the survey was designed.
1. Consent
2. Personal information
3. Randomized Sound files
4. Sound file response survey relevant to the sound file the respondent had listened to.
5. Modern Sexism survey
6. Explanations of what the experiment was about
7. Another consent form and an opportunity to leave the survey

For Group 1 the test was administered during two ordinary lessons of the ordinary course in Sociolinguistics that the students were attending. Only half the class was present on each occasion (sub-groups A and B). The students were told that they would listen to a recording of a discussion on gender issues between two researchers and that they would then be asked to answer some questions. Each student had access to a computer/tablet with headsets, and listened through the recording, answered the questions and they were then free to leave the class. Group A listened to the ‘male-morphed’ version of the Disney text while Group B listened to the version of the un-manipulated female voice. After the event, both groups were brought together and the design of the experiment was revealed.

For Group 2 the test was administered during ordinary lessons in the subject of English that the students were attending. They were not split into two groups, but instead what guise they listened to was randomized. The students had access to headphones and a laptop or smartphone. After everyone had completed the survey, the design of the experiment was revealed and discussed in relation to their ordinary teachings.

For Group 3 the test was administered during an ordinary lesson in their coursework. Whether they listened to the male or female guise was randomized as they started the survey. Everyone had headphones and most of the students had a laptop, and those who did not instead used a Smartpad. After everyone was finished, it was discussed in class.

For Group 4 the test was first posted on Reddit.com/r/samplesize (a subforum on Reddit where people help answer each other’s casual and academic surveys). No information other than it being for academic purposes was provided and they filled in the survey on their own volition. Roughly five respondents filled in the complete survey during the first two days. The link was reposted at a later time to find the last five that responded on Reddit. Further people were found through Facebook where the same method was applied, except instead of posting it myself, friends were chosen to post the link to the survey instead. Because the survey was provided online the study’s real aim was revealed at the end of the survey rather than in person.

For Group 5 the test was administered by SurveyMonkey whose instructions to the respondents were none other than what was provided in the survey. That is, that their responses would be used in research studies involving areas such as education,
discourse analysis and IT, gender/class/ethnicity and communication. Just like in Group 4, the study's real aim was revealed at the end of the survey rather than in person.

4.4 Rating Dimensions
The respondents rated the guises on a seven point Likert scale, ranging from 1 (disagree completely) to 7 (agree completely) on a number of dimensions, five related to speech style, and two related to characteristics/personality, namely:

- The American speaker interrupts a lot.
- The American speaker signals interest in what the other person is saying.
- The American speaker argues in a forceful way.
- The American speaker takes a lot of space in the conversation.
- The American speaker contradicts the other person a lot.
- The American speaker comes across as sympathetic.
- The American speaker comes across as knowledgeable.

Similarly, when the respondents filled in the explicit sexism survey they rated 14 questions (see appendix 3) on a five point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly disagree) (reversed for reverse questions). For the analysis, a respondent who, on average, rated high on the explicit sexism test was considered ‘more sexist’, similarly, a respondent who rated low was considered ‘less sexist’.

The seven point Likert scale was chosen for the responses to the guise because a five point Likert scale showed inadequacy in fine-tuning. The differences found in test runs simply showed that it might be too hard to isolate differences between the guises if the respondents only have five options. For the explicit sexism test, however, a five point Likert scale was used in order to be compatible with Ekehammar and Glick & Fiske’s sexism tests.

4.5 Technical Issues
The digitally manipulated voice in the recording sounded ‘robotic’ to some people in a test run so it was decided to add some background noise and distort the sound quality slightly. This made both the digitally manipulated voice and the other professor sound less clear, creating the illusion that any poor quality and/or robotic sound was only caused by a poor microphone or recording device, instead of being caused by voice manipulation. In the second test run, none of the respondents commented on the sound of the voice.
It was decided that the best way to present the respondents to the recording was to use an embedded video from YouTube. The audio track was added to a still-page of two people representing the co-participants in the recording. This gave the respondents a face to the voice, allowing them to deeper immerse themselves in the audio. However, this proved to add an additional issue, namely YouTube’s video recommendation page that comes up after a video is played. This meant that after listening to the recording it gave the other recording (the manipulated/the un-manipulated) as a suggestion for further watching, showing them that there are two different recordings, thus allowing them to figure out that they were deceived before they responded to the survey. To overcome this issue other channels were promoted so the respondents were redirected to other recommendations instead of the videos on the channel containing the different guises. This eliminated the problem as YouTube no longer suggested the manipulated/the un-manipulated video. In the final version, another step was taken as well, instead of using YouTube to embed videos, Vimeo was chosen instead, as it had no recommendation feature in the same way YouTube did.

Unfortunately, roughly 15 responses from Group 2 had to be discarded as there was an issue with the sound in one of the videos when they responded to the survey. This meant that all the answers pertaining to the collaborative recording for Group 2 were lost. After that data had been perused and the issue noted the recording was fixed and no further issues were revealed.

4.6 Ethical Dilemmas

For research ethics, the guidelines of Good Research Practice (Good Research Practice, 2011) were adhered to. Since the respondents are not informed what the study is about before listening to the recording and responding to the survey, it was considered ethically mandatory to include a page after the survey where they were informed about the purpose of the survey. Because respondents’ minds may be affected by the study as a byproduct of participating, the survey included two separate consent agreements, one before starting the survey, and one at the end after the aims of the study were revealed. This gave the respondents an opportunity to decline having their answers recorded after they knew what it was about as well as before. Ideally, there would only be one consent form, but since it is impossible to give proper consent without knowing what the study is about, a second one was mandatory since a proper explanation of the study before they respond to it would defeat the purpose of a matched-guise study.
5  Results and Analysis

5.1  Data

Before analysing all the data, the initial stance was that Group 1 (the one familiar with sociolinguistics) would potentially differ from the other groups in the way they perceived the guises. However, statistical tests showed that ‘group’ as a variable had no impact on the results, that is, the group to which the respondents belonged to did not affect their responses. Therefore, for all further analysis, ‘group’ was considered a non-variable. In table 3 below, you find an overview of all the respondents joined into one group for the relevant recording, that is all the respondents for ‘Disney discussion’ are combined into one group, and all the respondents for ‘Gender discussion’ are combined into another. However, if a more detailed overview is desired you find Appendix table 1 in the appendix section which presents an overview of all the respondents split up into their respective groups (group 1, 2, 3, etc.).

For the analysis, four different statistical procedures were used: MANOVA, MANCOVA, ANOVA, and ANCOVA. MANOVA (multivariate analysis of variance) is a model for comparing multivariate sample means used when comparing two or more dependent variables (we are considering seven here as dependent (interruptions through knowledgeable)) and later two (competitive and collaborative). MANCOVA is used in the same way, except we have an explaining variable (sexism) to ‘factor out’ noise or error. ANOVA (analysis of variance) is a statistical test to see whether the means of two or more groups are different when considering one variable at the time. ANCOVA is used in nearly the same way, except we also consider an explaining covariate (sexism) to ‘factor out’ noise or error. If further information is desired on the different procedures, see the footnotes\(^2\) for links to websites where information and mathematical formulae can be found.

When using the statistical procedure MANOVA there were no statistical differences between the two guises if we consider the seven response variables (interruptions through knowledge). However, when using the statistical procedure ANOVA we find that there are statistical differences between certain variables when considered in isolation from each other (e.g. just looking at interruptions, just interest, etc.). Data of this is first presented below in Figure 4, which is a graphical representation of the differences between the responses. The right columns refer to the male guise, and the left to the female guise. A higher column means a higher average rating by the respondents. This shows that the competitive speech variables are rated higher for the male guise, and the collaborative speech variables are rated higher for the female.
However, note that only floor space and contradictions are statistically significant, and interruptions are approaching significance (p value 0.069).

Below you find table 4 that details the differences more clearly. On the Y-axis you find category and interaction patterns, (e.g. male/female guise, Disney/gender discussion, male/female respondent), on the X-axis you find the variable tested (e.g. interruptions, interest, etc.), and the numbers represent the p value.
In table 4 above, we see that the male and the female guise versions of the conversations were interpreted in significantly different ways, in particular in relation to the perceived floor space that the speaker occupied (p value = 0.013), and the degree to which the speaker contradicted the conversational partner (p value = 0.04). Interruptions were approaching significance (p value = 0.07), showing that the respondents tended to interpret the male guise as interrupting more than the female guise. Also, note that there seems to be a strong tendency for male respondents rating the perceived floor space of the guise differently than female respondents. We see this in Table 4 where 'guise x gender' is approaching significance (.058). This means that female respondents perceived the female guise as taking less floor space than the male guise. Further, there were significant differences in the responses between the two texts (Disney and Gender discussion). Note that treatment responses for interest (.044), contradictions (.000), forcefulness (.000), sympathy (.002), and knowledge (.002) all are statistically significant showing that treatment (Disney/Gender discussion) affects how they rate the different aspects. Interestingly, while the different texts are rated differently, there does not seem to be an interaction between treatment and guise. That is, the difference between the guises was not affected by treatment.

Below, Figure 5, 6, and 7 show how the male guise is rated significantly higher on contradictions, floor space, and approaching significance on interruptions. Similar to Figure 4, the average rating is presented on the Y-axis, and on the X-axis the male guise on the right, and female on the left. A higher placement on the Y-axis means that the guise was perceived as e.g. contradicting more.
Further, below we have Figure 8 that shows the strong tendency of how female respondents rate the female guise as taking much less floor space than the male guise, meanwhile, the males rate the two guises fairly evenly. For a more detailed view, see Appendix Table 6.
Below you find Table 5, which details how the responses differed when considering sexism as a covariate with the mean value of 2.4137. Similar to Table 4, the Y-axis shows the category, the X-axis the response variable, and the numbers represent p value.

Table 5. p values with sexism as a mean value of 2.4137

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Interruptions</th>
<th>Interest</th>
<th>Forceful</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guise</td>
<td>.251</td>
<td>.034</td>
<td>.000</td>
</tr>
<tr>
<td>Gender</td>
<td>.072</td>
<td>.423</td>
<td>.417</td>
</tr>
<tr>
<td>sexism</td>
<td>.811</td>
<td>.291</td>
<td>.116</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Guise</th>
<th>Gender</th>
<th>.045</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Floorspace</th>
<th>contradicts</th>
<th>sympathetic</th>
<th>knowledgeable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>.124</td>
<td>.001</td>
<td>.000</td>
</tr>
<tr>
<td>Guise</td>
<td>.009</td>
<td>.063</td>
<td>.518</td>
</tr>
<tr>
<td>Gender</td>
<td>.166</td>
<td>.073</td>
<td>.663</td>
</tr>
<tr>
<td>sexism</td>
<td>.086</td>
<td>.494</td>
<td>.127</td>
</tr>
<tr>
<td>Guise</td>
<td>Gender</td>
<td>.039</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>--------</td>
<td>------</td>
<td></td>
</tr>
</tbody>
</table>

Figure 6. Guise and gender interactions - floor space
In this table, we see that when using sexism as an explaining variable we can tell that the gender of the respondent seems to affect their responses to the guise. The fact that the 'gender' p value increased as sexism was introduced as an explaining variable suggests that these two explain the same thing. That is, that rather than the gender, it may be the sexism of the respondent that affects their rating. However, we also see that the gender of the respondents affected how they rated the interest of the guise in a significant way, which suggests that gender of the respondent has some effect. We see that females perceived the female guise as showing more interest on average than the male guise. Males on the other hand, perceived the male guise as showing more interest than the female. This is graphically shown in Figure 9 below and the numbers on this can be found in Appendix Table 5.

Figure 7. Guise and gender interactions – interest. Sexism as mean value

When testing the interaction variables ‘sexism x guise’ (i.e. if sexism affects how they rate the guise) we found no statistically significant differences, although, we saw strong tendencies. The sexism of the respondent seems to affect how they rated three aspects of the guise: contradictions (.060), sympathy (.052), and knowledge (.072). This can be seen in Table 6 below which details how the responses differed when considering sexism and guise as interaction variables. Similar to Table 5, the Y-axis shows the category, the X-axis the response variable, and the numbers represent p value.
When first conducted, the MANOVA and MANCOVA analysis did not show any statistically significant differences between the guises when accounting for all seven variables (interruptions through knowledgeable), however, the data showed some responses were strongly related to other responses. This is presented below in Table 7 that illustrates how certain statements were likely to be rated in the same way. In the Y-axis you find the variable, on the X-axis you find the different groups (1 and 2), and the numbers in the table show how strongly the correlated with each other. A higher number means a higher chance of the respondent answering similarly to both variables.

<table>
<thead>
<tr>
<th></th>
<th>Interruptions</th>
<th>Interest</th>
<th>Forceful</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>.235</td>
<td>.048</td>
<td>.000</td>
</tr>
<tr>
<td>Guise</td>
<td>.093</td>
<td>.154</td>
<td>.998</td>
</tr>
<tr>
<td>Gender</td>
<td>.831</td>
<td>.412</td>
<td>.119</td>
</tr>
<tr>
<td>sexism</td>
<td>.651</td>
<td>.003</td>
<td>.942</td>
</tr>
<tr>
<td>Sexism x guise</td>
<td>.211</td>
<td>.208</td>
<td>.822</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Floorspace</th>
<th>contradict</th>
<th>sympathetic</th>
<th>knowledgeable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>.164</td>
<td>.000</td>
<td>.000</td>
<td>.002</td>
</tr>
<tr>
<td>Guise</td>
<td>.049</td>
<td>.021</td>
<td>.041</td>
<td>.034</td>
</tr>
<tr>
<td>Gender</td>
<td>.105</td>
<td>.077</td>
<td>.631</td>
<td>.701</td>
</tr>
<tr>
<td>sexism</td>
<td>.144</td>
<td>.324</td>
<td>.228</td>
<td>.045</td>
</tr>
<tr>
<td>Sexism x guise</td>
<td>.182</td>
<td>.060</td>
<td>.052</td>
<td>.072</td>
</tr>
</tbody>
</table>
As we can see here in Table 7, the collaborative responses correlated with each other, and the competitive responses correlated with each other. This shows that people who rate interruptions high also rate forcefulness, floor space and contradictions high. In other words, they rate all competitive aspects high. It also shows how those who rate interest high also rate sympathetic and knowledgeable high. Knowledgeable and sympathetic are not related to speech style, but it is easy to see how people who think that the guise is interested in the other person’s opinion is more knowledgeable and sympathetic than those who think they argue and disagree a great deal. For our purposes here, knowledge and sympathy will be considered collaborative variables when discussing them. Instead of using a MANOVA procedure using seven variables it was possible to compare two variables (competitive and collaborative speech variables) with the data derived from table 5. This showed that there is a statistically significant (.013) difference between how the competitive and collaborative speech questions were rated between the guises. This is presented in Figure 10 below, and the detailed numbers on this can be found in Appendix Table 7.
Here in Figure 10, we see that for the competitive speech questions the female guise is perceived as being significantly less competitive, as it is rated lower on all the competitive variables.

Interestingly, the male respondents were significantly more sexist than the female respondents. Below you find Figure 12, which details this difference split into the mean values and spread. The Y-axis shows how high they rated on sexism (a higher score means more sexist), the X-axis shows the gender of the respondent, the boxes detail the spread of the sexism values, and the black line in the boxes signify the mean value. Here we see how male respondents were generally more sexist than females. The female respondents had a mean value of 2.05 while the male respondents had a mean value of 2.506 on a 5-point scale (p value .000).
Figure 9. Sexism differences between male and female respondents

When we considered the variable groupings we saw in Table 5, the results showed that when we grouped interest, sympathy, and knowledge together and compared that to floor space, interruptions, contradictions, and forcefulness with ‘sexism x guise’ as interaction variables. We noticed that there were statistically significant ways in which the guise was rated depending on the sexism of the respondent. This is presented in Table 8 below.

Table 8. Collaborative and Competitive speech variables with sexism x guise interaction variable

<table>
<thead>
<tr>
<th></th>
<th>Collaborative</th>
<th>Competitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>.004</td>
<td>.001</td>
</tr>
<tr>
<td>Guise</td>
<td>.030</td>
<td>.089</td>
</tr>
<tr>
<td>Gender</td>
<td>.499</td>
<td>.416</td>
</tr>
<tr>
<td>sexism</td>
<td>.016</td>
<td>.913</td>
</tr>
<tr>
<td>Sexism x guise</td>
<td>.046</td>
<td>.262</td>
</tr>
</tbody>
</table>

Here we find that the sexism of the respondent strongly affected how they rated the collaborative speech variables of the guise. For every 1 point higher in sexism of the respondent, the female guise is rated .475 less in the collaborative variables (see Appendix table 8 for more details). However, we also see that there were no significant differences for the competitive speech variables. That is, the respondent
being more or less sexist did not affect how they rated the guise on floor space, contradictions, interruptions or forcefulness.

5.2 Data Summary

In this section, we have found statistical support for many different aspects. These differences are summarized below.

- There are significant differences between how the guise is rated when it comes to floor space and contradictions. Interruptions are approaching significance as well.
- The gender of the respondent seems to affect how they rate the floor space and the interest of the guise.
- The difference found in how they rate differently based on gender can be partially explained by the respondent’s sexism.
- The sexism of the respondent seems to affect how they rate the guise on three variables: contradictions, sympathy, and knowledge. None of these are statistically significant, but all are approaching significance.
- The competitive speech variables (interruptions, floor space, contradictions, and forcefulness) correlated with each other, as did the collaborative (interest, sympathy, and knowledge).
- By comparing the competitive speech variables to the collaborative speech variables, we found statistical differences between how the guises were rated. The female guise was rated higher on the collaborative variables and the male was rated higher on the competitive.
- Using a similar method and considering sexism’s interaction with the guise, we also found statistical evidence that the collaborative speech variables are rated lower on the female guise the higher the sexism of the respondent. Although, the female guise is always still rated higher on variables.
- The male respondents are more sexist than the female respondents.
- The gender discussion recording was rated significantly higher on all competitive aspects (floor space, interruptions, forcefulness, contradictions) and the Disney discussion recording was rated significantly higher on all collaborative aspects (interest, knowledge, and sympathy).
6 Discussion

The results presented in this paper suggest that there are statistically significant differences between how a male and female are rated differently based on the listener’s stereotyping. Utilizing the matched-guise method, we should have eliminated all other variables but the listener’s subjective ideas of what is female and what is male (Agheyisi & Fishman, 1970), however, we introduced the issue of sound quality. Many different methods were considered before finally settling on the one presented in the method section, but the quality was still a potential issue. Although, any robotic sound should primarily affect sympathy and arguably interest, while not have a particularly significant effect on e.g. floor space.

It was hypothesized that the male guise would be rated higher on competitive (and knowledge) speech variables, and lower on the collaborative (and sympathy) variables. This was the initial stance as it harkens back to the statistical stereotypes of how men and women talk (see e.g. (Coates, 2004; Cheshire & Trudgill 1988; Abel & Meltzer, 2007; Centra & Gaubatz, 2000). The results show that the differences found are regarding precisely those variables, only not all of them as was expected. Instead of ‘interruptions’, ‘contradictions’, ‘floor space’, ‘forcefulness’, and ‘knowledge’ being rated higher for the male guise, we instead found statistically significant differences between two of them and one approaching significance. These differences were found with floor space, contradictions, and interruptions. In other words, there were no discernible differences between the guises in how they were rated on forcefulness or knowledge, nor interest or sympathy. This means that when looking at the different guises without considering sexism, there were differences between two (or three) of seven variables. Both (or all three) of these variables were competitive variables ((interruptions), floor space, contradictions).

We found that certain responses correlated with other responses, this made it possible to divide those variables into two separate groups. The responses that correlated with each other were those relating to speech style. Namely, collaborative speech variables (interest, sympathy, knowledge) constituted one group, and the competitive speech variables (floor space, interruptions, contradictions, forcefulness) the other group. We found that people who rated interruptions high also rated forcefulness, floor space and contradictions high. That is, they rate all competitive variables high. This means that we were able to compare two speech styles, which gave us statistically significant differences in how the guises were perceived. From these results we can say that the male guise is rated higher on competitive aspects, while the female guise is rated higher on the collaborative aspects (and knowledge/sympathy).

Contradictory to old stereotypes (see e.g. (Abel & Meltzer, 2007; Centra & Gaubatz, 2000)) it would seem as though men are considered less knowledgeable than women. This can be understood through three sets of data, first if we consider a comparison between the guises with knowledge responses in isolation. Here we see
that while not statistically significant, the male guise was considered less knowledgeable. Secondly, we can say with certainty that both guises in the gender discussion were considered less knowledgeable. Through this, we can understand that a more competitive speech style is seen as less knowledgeable since the gender discussion was perceived as much more competitive. Thus, because the male guise was considered more competitive, we can conclude that the male guise was regarded as less knowledgeable. Thirdly, we can see that the collaborative speech variables (including knowledge and sympathy) were rated higher for the female guise, which gives further credence to the notion that the male guise was perceived as less knowledgeable. Although, it is important to emphasize that because the differences were not statistically significant when comparing guises in regards to the knowledge responses alone, we cannot say this with certainty. Further research with more respondents would be required to determine if the conclusion holds true. Why we find relation of knowledge between a competitive speech style and the male guise might be explained by recent trends. If we consider the most recent PISA results, we find that girls are performing better each year, while boys are performing worse. Moreover, if we consider the Global Education Digest (2010) we see that more women than men are focusing on tertiary education, which suggests that it is more likely to encounter highly educated women than men. These tendencies of women focusing more on school than men might give us some idea of why the male guise is rated lower for knowledge. The results indicate that what is considered an intelligent speech style has shifted somewhat as the collaborative speech variables correlated with knowledge and the competitive did not. It may be, for example, that the traditional view of how to appear knowledgeable has changed and that the modern generation view someone who is prepared to listen as more serious (Barack Obama, for example) than someone who is overly competitive (Donald Trump, for example).

There also appears to be a relationship between the gender of the respondent and how certain variables are rated. For example, we see that female respondents perceived the female guise as taking much less floor space than the male guise, meanwhile, the males perceived the female guise and male guise moderately evenly. This relationship between guise and gender can also be found in how males and female respondents rated ‘interest’ of the guise differently. The results show how male and female respondents rated opposite from each other. Where males thought that the female guise showed less interest and the male guise more, the females strongly thought that the male guise showed less interest and the female more. These differences can possibly be accounted for if we look at sexism instead of gender. If we consider the split between sexism of the respondents we find that they seem to relate to the differences between how the guises are rated. Surprisingly, this shows that males (or more sexist people) seem to rate the guises more evenly, while females (or less sexist people) rate the guises unevenly favouring the female guise for

---

collaborative speech variables. It is hard to determine whether it is gender or sexism that causes these differences, all we can tell with certainty is that there are both differences between how men and women rated the guises and that these differences are in line with the differences in sexism between the male and female respondents. It is difficult to determine why less sexist/females rate more unevenly and favour the female guise. It may be because of overcompensation where people who are less sexist according to the sexism test are more aware of gender inequalities and have had greater exposure to gender difference studies, which in turn has led to stereotypic views. For example, unpublished research by Deutschmann and Steinvall suggests that when students were made aware of gender differences in language they started favouring the female guise for collaborative aspects and the male for competitive. Alternatively, perhaps because they are less sexist they find a female linguist much more interesting and knowledgeable than a male, that is, they are interested in what the guise has to say because it is a woman arguing/discussing with a man. On the other hand, it could also be because of the digitally manipulated voice. While it showed no issues in the testing before the respondents listened to it, it might still affect how people interpret the voice. Although, this should only have affected sympathy and interest, it is hard to argue that a slightly more robotic voice may, for example, take more floor space or interrupt more.

The data suggests that sexism of the respondent affects how they rated the guise. A more sexist person rated the female guise lower on all collaborative speech variables. Importantly, this does not mean that people who scored higher on the sexism test disfavoured the female guise for the collaborative aspects, rather, the less sexist people favoured the female guise. That is, a more sexist person rated the guises more evenly. Because the competitive speech variables were not favoured for males when accounting for sexism, we can draw the conclusion that the difference we find here is not because of the likelihood of less sexist people being better versed in gender studies, since they would then have rated the male guise higher on the competitive variables as well. Rather, it may be because, as argued above, less sexist people overcompensate. They particularly notice that it is a female linguist and therefore find her more interesting, more knowledgeable, and more sympathetic.

There does not seem to be a statistical difference between how the two different texts were rated when it comes to the guises. That is, the respondents did not seem to react more when a speaker goes against or adheres to stereotypical norms (see e.g. (Cheshire & Trudgill, 1998; Coates, 2004). In other words, the difference between the guises does not change if we compare the guises in the Disney text to the guises in the Gender text. A woman is rated higher on the collaborative variables in both recordings, just as the man is rated higher on the competitive variables in both. This suggests that, for example, a woman being overly aggressive/competitive in speech does not negatively affect her more than it would a man, which suggests that speech style might not relate to hegemonic gender as much as it once has. Although, we
cannot say this for sure since we do not have any data on variables such as ‘being a bitch’, and so on.

7 Conclusion and Suggestions for Further Research

With the data collected with this matched-guise study, we have been able to answer all research questions presented in the Aims section. The results are promising and indeed show that there are statistically significant differences between how the male guise and female guise are rated. Firstly, the competitive speech aspects are rated higher for the male guise, and the collaborative speech aspects are rated higher for the female guise. Secondly, we find that the female guise appears to be considered more knowledgeable and sympathetic as well. Thirdly, gender or sexism of the respondent seems to affect how they rated the guise as well. Generally, males were more sexist, and their responses were in line with the responses of the ‘more sexist’ group. However, the respondents who were more sexist never favored the male guise, rather, the respondents who were less sexist favored the female guise on all collaborative variables. Fourthly, there did not seem to be a difference between how the female guise and male guise were rated in relation to recording. That is, the differences found between the guises were similar in the Gender discussion and the Disney discussion.

For further research, it is important to consider several aspects. First, we have to consider the demographics of the respondents. This study included almost only respondents from the West (most from the UK and Sweden), which means that while the respondents are moderately varied here, further data is required to make stronger assertions. Noteworthy, however, is that since the groupings did not affect the differences, it is unlikely we would find other results than what we have found here, at least in the West. Second, there is a great deal of value in trying to optimize the survey so that fewer respondents opt out when they see the length of it. As it stood here, more than 75% of the respondents from the Internet group cancelled the survey after only answering half the questions. The best places to save time are the explicit sexism test, which can be made shorter, and the recordings themselves that can be condensed down to 3-4 minutes instead of the 4-6 they are now. Third, by getting more respondents we could potentially see the differences presented in table 5 become clearer. Here we already see the general pattern of how the female guise is rated higher on collaborative speech aspects while the male is rated higher on the competitive, but by getting more respondents we could potentially find that all of these are statistically significant in isolation as well as when combined as one variable (competitive vs collaborative variables). Finally, doing this research diachronically could shed some light on the differences found. Perhaps a cultural shift may be the reason why we find that the female guise is rated higher on knowledge than male.
8 References


44

Good Research Practice (2011), at https://publikationer.vr.se/produkt/good-research-practice/ (last accessed 29 May 2016)


9 Appendix

9.1 Disney Text

The following conversation is an informal panel discussion between two linguists (John and Jane) where the general theme was “How does popular children’s media affect linguistic gender stereotypes?.

Floor space: Jane: 51% John 49%
Dialogue is more of a Jam session than one person playing the expert.
Interruptions by Jane: 4
Heges by Jane: 19
Minimal responses / back channelling by Jane: 7
Information seeking question by Janes: 1
Collaborative questions (hedges, view checking, etc.) by Jane: 3

John: I mean I’d say popular media is definitely responsible for perpetuating gender stereotypes in conversation. You know, I-I read this really interesting study recently on the Disney Princess movies in Wha. […]

Jane: […] oh yea yeah I saw that too! That’s the one where they looked at the percentage of words spoken by female and male characters in several different movies right?

John: Yeah, yea, that’s the one! Yea, and it was really interesting to see how females occupied as little as 20-30% of the floor space, you know, in spite of these movies having main characters that were girls.

Jane: Yeah, yea that is really disturbing. Bu-but what I gathered that was particularly the case for the movies produced in the nineties.

John: Yeah, yea, I think you’re right there.

Jane: I mean, the statistics were actually a lot more positive for the earlier films,

John: mm, mm

Jane: Like, like (??) Snow White and Cinderella, and I think in the movie Sleeping Beauty from like 59, they were, they were females were producing like 70% of the dialogue.
John: Yeah yea, I think you’re right there. I, uh, didn’t really pay that much attention to the details.

Jane: hmm, yeah

John: But from what I gather, it’s improved in recent years uh. The more recent movies seemed to be more balanced, at least some of them.

Jane: hmm yea, but there’s other aspects than floor space to look at […]

John: […] yeah yea, that’s true. But, uhm, you know, they did actually look at other aspects as well, such as the number of male and female characters I think uh,..

Jane: mhm, mhm

John: uh, uh, and they’ found that male characters sort of greatly outnumbered female characters in all the movies […]

Jane: right right […] Right right right, like all the side-kicks were apparently male. And all the other active roles, other than the female main characters of course, were apparently populated by men […]

John: […] yeah, yea yea, an-an-and all the funny side characters li-like the Genie in Aladdin, the lobster in the Little Mermaid and, and. They all seem to be male…. […]

Jane: mm, yea, […] and not only those! Did you ever see a female police officer in an animated Disney Movie for example?

John: hmmm, nope. Can’t think of any […]

Jane: no, no nope […] or a shopkeeper or a soldier or really any other profession?

John: No no I Can’t think of any but bu-but yea hang on a moment maybe Cruella de Vil in you know Pongo and the Dalmatians or whatever it was called I mean she might qualify as as professional ‘bitch’ maybe?
Jane: ha ha. Yea, yea, but yea well didn’t they look also at the content in conversations?

John: Yeah, an-and that’s where it was really really interesting actually. Cause they looked at compliments […]

Jane: hmm [...] yea, that’s right that’s right! It was compliments. And […]

John: […] yea yea, and I think that the majority of compliments that were given to female characters in the movies, had to do with looks rather than things that they did […]

Jane: […] yea of course! Yea But if I also, also, this was not just the case that it was also in more recent movies that was all reversed.

John: yeah, I think you’re right there. Uh, so there might be a bit of hope after all then

Jane: hmm, hmm. You know, I don’t know, maybe maybe the most important question is I mean do these movies actually affect children’s behaviour or their expectations when it comes to like gender roles.

John: Mmm. I-I think they definitely do and uh, they probably have more influence than we like to think, you know?.

Jane: Yeah.

John: uh, and I-I mean I often see kids at the prep school, you know, I mean, where my son is, an-and they sorta sit in corners and they play all these characters you know princesses and its sorta part of their everyday playing activities […]

Jane: […] yea I think you’re right, and the age groups these films they appeal to like 4-10 year-olds, I mean those are really formative years when it comes to gender roles.

John: mmm. Yep, yep

Jane: And in fact studies like this one they show the importance of linguistic research. I mean after reading this study, I’m thinking that Disney will probably pay more attention to the gender aspects of their dialogues.
John: yea I think you’re right.

Jane: You know, they may not even be aware of these facts … […]

John: […] yeah yea, I mean the movies are probably just a reflection of societal norms, and they can’t really be blamed …

Jane: […] hmmm, yea but they do, they do have extra responsibilities given the impact that these movies have on kids of all ages.

John: ye-yea that’s true! Well at least we can conclude that linguists seem to have an important role then yea?

Jane: Yeah yea! Well I’m not sure everyone agree but that’s a different story.

9.2 Gender discussion text

The following is a panel debate between two linguists with the title “What does science say about gender differences in language”.

Floor space: John: 49%; Jane: 51%
Both participants try to play the expert.

Interruptions by Jane: 4
Heges by Jane: 24
Minimal responses / back channelling by Jane: 12
Information seeking question by Janes: 6
Collaborative questions (hedges, view checking, etc.) by Jane: 0
John 1: I’d probably say that much of the research that has been carried out in relation to gender and conversational styles is actually damaging because it adds strength to existing stereotypes.
Jane 1: [what?]
John 2: Yea researchers simply seem to seek evidence to confirm stereotypic behaviour and the question remains if we should even include this research in our teachings...
Jane 2: [...]Oh come on! That’s ridiculous I mean I would say the total opposite. The role of sociolinguistics research is to expose structural frameworks and how they are reflected in language
John 3: Yea, yea
Jane 3: and then we can actually start combating structures and stereotyping. I mean of course that would be the kind of knowledge that that’s what belongs on the curriculum. I mean, are you suggesting we start censoring courses?
John 4: No, no no, of course not. I mean I agree with you to a certain point
Jane 4: Okay?
John 5: but then again, you know, researchers tend to find what they are looking for, and and that what they choose to publish.
Jane 5: Hmmm
John 6: uh And and I don’t know how many studies I’ve read that seek to test hypotheses that are direct extensions of popular stereotypes.
Jane 6: What do you mean, you’re going to have to explain that.
John 7: Well well, okay. Take, take the research on aspects such as collaborative vs competitive conversational styles. And
Jane 7: yea, yea
John 8: I mean, based on previous research [uh] it’s just assumed that collaborative styles are typical for women and a positive result in a study seem to be a confirmation of this, right? This stereotype.
Jane 8: Yea yea, but you’re confusing like systemic structural patterns with stereotypes. I mean if society equates aggressive conversational behaviour power. You know, which is in turn equated with being male.
John 9: Yea
Jane 9: Then it’s not so surprising that study after study confirm this relationship. I mean, it’s just a reflection of society
John 10: Yea, yea, yea I guess
Jane 10: I mean I don't even see how this is a problem. Shouldn't we teach our students about such structures?

John 11: Well well, yea, but, but the question remains, right? Whether the research is representative or not. I mean I I I think loads of studies that fail to find, you know, what they're looking for they, they are never published you know? They they are not interesting and and

Jane 11: oh come on now, I mean that is really a cynical view. I mean that is a cynical view of research. I mean we have to assume that researchers have some integrity to publish their results

John 12: Yea yea, but, but

Jane 12: (sounds?) even those that don't fit the original assumptions. If we don't I mean the whole foundation of science can be questioned. We might as well give up.

John 13: Yea, okay. You got a good point there right?, but but, the problem is that although researchers may cover all sorts of different aspects, only a fraction of this research reaches the public, yea?

Jane 13: ye, okay, true.

John 14: yea And some studies get greater publicity and have greater impact than others. And I would argue that those studies surprise surprise that fit the popular world view are the ones that get published.

Jane 14: Yea, yea I see your point, and okay that's valid but, the way I see it you're indirectly suggesting that the findings from current research is incorrect, and that that they there aren't any gender differences in conversation

John 15: no no no that’s not what I'm saying at all, you know? There may well be gender differences, but but, gender as such is (???) I I I would argue.

Jane 15: hmmm okay

John 16: because

It it's power that's the interesting variable.

Jane 16: hmmm yea, okay

John 17: uhh, an-and you know by constantly making conversational styles a gendered issue you actually strengthen stereotypes. And in addition, you know popular science with books such as, you know, men are from mars and women are from venus, and what have you. I mean they just further cement these views.

Jane 17: Yea, okay you have a point there, but you know where there's a commercial possibility the market's always going to capitalize

John 18: Yea exactly
Jane 18: *I mean, I think this type of research has made some important contributions anyway, I mean* it highlights beneficial uses of different conversational styles

John 19: *yea*

Jane 19: *I mean* there’s a collaborative or repertoire type conversation, for example. *I mean* that’s being sold as successful leadership skills for both men and women

John 20: *Yea yea, but but but* those styles are still being associated with being female or male. *You know? And sometimes the-they even label it as such, speaking as a woman speaking as a man you know? Men’s style and and*

Jane 20: *hmm, hmmm yea that’s true. That’s true.*

John 21: *And and. Yea yea, and that’s a problem I think*

Jane 21: *Well why would that be a problem?*

John 22: *We-well stereotypes that are portrayed as positive such as women are nurturing, they’re just as problematic and perhaps even more so than stereotypes that are more negative such as women tend to gossip, you know? You know, positive stereotypes are easier to relate to, we-we all want to be flattered.*

Jane 22: *hmm, yea.*

John 23: *yea, bu-but they still maintain the status quo. It’s a bit like saying I’m poor but I’m honest, you know that kind of stereotype. It’s flattering yea, but it does nothing towards changing structures*

Jane 23: *yea, well that that could be but it doesn’t change the facts, I mean, looking at the world statistically, we can conclude that men are more likely to be aggressive in conversations than women, or that men take up more space in conversation, or that women are more active listeners and tend to build conversation from context etc. I mean, what are you proposing? I mean that we hide*

John 24: *hang on hang on. You’re doing, you know, just what I’ve been criticising here, you know, equating behaviour with group traits such as gender, you know? If I were to show you statistics that afro-american are overrepresented in jail, right? An-and call afro-americans criminals you’d call me a racist.*

Jane 24: *hm yea* [stifled laughter]

John 25: *But-but for some reasons same logic doesn’t seem to apply to gender.*

Jane 25: *yea, but perhaps gender is a better predictor of behaviour than ethnicity*

John 26: *yea, but isn’t that reinforcing stereotypes?*

Jane 26: *no that’s not what I’m saying. I mean, I think you’re underestimating the intelligence of our students*
John 27: [mumbling] I guess
Jane 27: I-I mean I know they are perfectly capable of distinguishing between statistical tendencies and individual behaviour. And I’m sure that they can walk (???) at the same time.
John 28: yea, mhm
Jane 28: I mean, I don’t know. Pointing out systematic structures and highlighting these does not exempt the individual from being an individual, and responsible for his or her actions.
John 29: yea-yea I guess you’re right
Jane 29: an-and how can we even start to address inequality and investigate the reasons behind them if we’re scared to present the evidence?

9.3 Explicit Sexism Test

1. Discrimination of women is no longer a problem in Sweden.
2. Humiliating treatment of women in adverts is unusual.
3. Society treats men and women the same.
4. The government puts too much emphasis on women’s issues.
5. The women’s movement serves no purpose and should be abolished.
6. No matter how accomplished he is, a man is not truly complete as a person unless he has the love of a woman
7. Every man ought to have a woman whom he adores.
8. Women are too easily offended
9. Feminists are making entirely reasonable demands on men.
10. There are actually very few women who get a kick out of teasing men by seeming sexually available and then refusing male advances.
11. Better measures should be taken to achieve equality between the sexes in workplaces.
12. It’s easy to understand the viewpoints of women’s groups.
13. The school curriculum should be adapted to girls’ needs.
14. In a disaster, women ought not necessarily to be rescued before men.

Blue = Reversed coding
## 9.4 Extra Tables and Figures

### Appendix Table 1. Overview of all groups

<table>
<thead>
<tr>
<th></th>
<th>Treatment</th>
<th>Gender Discussion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Disney Discussion</td>
<td>Gender Discussion</td>
</tr>
<tr>
<td></td>
<td>Guise</td>
<td>Male</td>
</tr>
<tr>
<td></td>
<td>Group</td>
<td>Group</td>
</tr>
<tr>
<td></td>
<td>Internet Respondents</td>
<td>Survey Monkey Respondents</td>
</tr>
<tr>
<td>Interruptions Mean</td>
<td>4.00</td>
<td>2.50</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>2.30</td>
<td>1.96</td>
</tr>
<tr>
<td>Count</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Forceful Mean</td>
<td>3.00</td>
<td>3.09</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>1.30</td>
<td>1.38</td>
</tr>
<tr>
<td>Count</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Floor space Mean</td>
<td>4.00</td>
<td>3.50</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>1.41</td>
<td>1.17</td>
</tr>
<tr>
<td>Count</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Contrasts Mean</td>
<td>3.00</td>
<td>4.03</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>1.82</td>
<td>1.51</td>
</tr>
<tr>
<td>Count</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Sympathetic Mean</td>
<td>3.60</td>
<td>4.02</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>1.52</td>
<td>1.03</td>
</tr>
<tr>
<td>Count</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Knowledgeable Mean</td>
<td>4.80</td>
<td>5.22</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>.84</td>
<td>.85</td>
</tr>
<tr>
<td>Count</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Sexism Mean</td>
<td>2.32</td>
<td>2.02</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>.81</td>
<td>.63</td>
</tr>
<tr>
<td>Count</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

### Appendix Table 2. Variations between interruptions, no sexism

#### Parameter Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>B</th>
<th>Std. Error</th>
<th>t</th>
<th>Sig.</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>4.551</td>
<td>.299</td>
<td>15.212</td>
<td>.000</td>
<td>3.950 – 5.141</td>
</tr>
<tr>
<td>[Treatment=Disney Discussion]</td>
<td>-.365</td>
<td>.284</td>
<td>-1.288</td>
<td>.200</td>
<td>-.926 – .195</td>
</tr>
<tr>
<td>[Treatment=Gender Discussion]</td>
<td>0</td>
<td>.2</td>
<td>0.000</td>
<td>.999</td>
<td>-9.02 – 9.02</td>
</tr>
<tr>
<td>[Guise=Female]</td>
<td>-.433</td>
<td>.237</td>
<td>-1.827</td>
<td>.069</td>
<td>-.982 – .025</td>
</tr>
<tr>
<td>[Guise=Male]</td>
<td>0</td>
<td>.2</td>
<td>0.000</td>
<td>.999</td>
<td>-9.02 – 9.02</td>
</tr>
<tr>
<td>[Gender=Female]</td>
<td>.029</td>
<td>.234</td>
<td>0.125</td>
<td>.901</td>
<td>-.433 – 9.462</td>
</tr>
<tr>
<td>[Gender=Male]</td>
<td>0</td>
<td>.2</td>
<td>0.000</td>
<td>.999</td>
<td>-9.02 – 9.02</td>
</tr>
</tbody>
</table>

*This parameter is set to zero because it is redundant.*

56
Appendix Table 3. Variations between floor space, no sexism

<table>
<thead>
<tr>
<th>Parameter</th>
<th>B</th>
<th>Std. Error</th>
<th>t</th>
<th>Sig.</th>
<th>95% Confidence Interval</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>4.938</td>
<td>.261</td>
<td>18.926</td>
<td>.000</td>
<td>4.423</td>
<td>5.453</td>
<td></td>
</tr>
<tr>
<td>[Treatment=Disney Discussion]</td>
<td>-.344</td>
<td>.247</td>
<td>-1.391</td>
<td>.166</td>
<td>-.812</td>
<td>.144</td>
<td></td>
</tr>
<tr>
<td>[Treatment=Gender Discussion]</td>
<td>0°</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>[Gague=Female]</td>
<td>-.525</td>
<td>.207</td>
<td>-2.538</td>
<td>.012</td>
<td>-.933</td>
<td>-.117</td>
<td></td>
</tr>
<tr>
<td>[Gague=Male]</td>
<td>0°</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>[Gender=Female]</td>
<td>-.438</td>
<td>.204</td>
<td>-2.143</td>
<td>.034</td>
<td>-.841</td>
<td>-.034</td>
<td></td>
</tr>
<tr>
<td>[Gender=Male]</td>
<td>0°</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
</tbody>
</table>

a. This parameter is set to zero because it is redundant.

Appendix Table 4. Variations between contradictions, no sexism

<table>
<thead>
<tr>
<th>Parameter</th>
<th>B</th>
<th>Std. Error</th>
<th>t</th>
<th>Sig.</th>
<th>95% Confidence Interval</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>5.112</td>
<td>.262</td>
<td>19.588</td>
<td>.000</td>
<td>4.595</td>
<td>5.630</td>
<td></td>
</tr>
<tr>
<td>[Treatment=Disney Discussion]</td>
<td>-.887</td>
<td>.248</td>
<td>-3.671</td>
<td>.000</td>
<td>-1.378</td>
<td>-.397</td>
<td></td>
</tr>
<tr>
<td>[Treatment=Gender Discussion]</td>
<td>0°</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>[Gague=Female]</td>
<td>-.429</td>
<td>.208</td>
<td>-2.067</td>
<td>.040</td>
<td>-.840</td>
<td>-.019</td>
<td></td>
</tr>
<tr>
<td>[Gague=Male]</td>
<td>0°</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>[Gender=Female]</td>
<td>.459</td>
<td>.205</td>
<td>2.236</td>
<td>.027</td>
<td>.054</td>
<td>.884</td>
<td></td>
</tr>
<tr>
<td>[Gender=Male]</td>
<td>0°</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
</tbody>
</table>

a. This parameter is set to zero because it is redundant.
Appendix Table 5. Variations between interest, gender x guise. Sexism included.

Parameter Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>B</th>
<th>Std. Error</th>
<th>t</th>
<th>Sig.</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>5.029</td>
<td>.540</td>
<td>10.863</td>
<td>.000</td>
<td>[4.764, 6.295]</td>
</tr>
<tr>
<td>[Treatment=Disney Discussion]</td>
<td>.673</td>
<td>.267</td>
<td>2.542</td>
<td>.012</td>
<td>[.445, 1.110]</td>
</tr>
<tr>
<td>[Treatment=Gender Discussion]</td>
<td>0 *</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>[Guise=Female]</td>
<td>-.271</td>
<td>.319</td>
<td>-.850</td>
<td>.400</td>
<td>[-.990, .459]</td>
</tr>
<tr>
<td>[Guise=Male]</td>
<td>0 *</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>[Gender=Female]</td>
<td>-.700</td>
<td>.340</td>
<td>-2.060</td>
<td>.041</td>
<td>[-1.370, -.029]</td>
</tr>
<tr>
<td>[Gender=Male]</td>
<td>0 *</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>[Guise=Female] × [Gender=Female]</td>
<td>.905</td>
<td>.447</td>
<td>2.033</td>
<td>.045</td>
<td>[.021, 1.788]</td>
</tr>
<tr>
<td>[Guise=Female] × [Gender=Male]</td>
<td>0 *</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>[Guise=Male] × [Gender=Female]</td>
<td>0 *</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>[Guise=Male] × [Gender=Male]</td>
<td>0 *</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>Sexism</td>
<td>-.597</td>
<td>.178</td>
<td>-3.364</td>
<td>.001</td>
<td>[-.948, -.245]</td>
</tr>
</tbody>
</table>

* This parameter is set to zero because it is redundant.

Appendix Table 6. Variations between floor space, gender and guise interactions, no sexism

Parameter Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>B</th>
<th>Std. Error</th>
<th>t</th>
<th>Sig.</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>4.736</td>
<td>.280</td>
<td>16.933</td>
<td>.000</td>
<td>[4.194, 5.288]</td>
</tr>
<tr>
<td>[Treatment=Disney Discussion]</td>
<td>-.392</td>
<td>.247</td>
<td>-1.591</td>
<td>.114</td>
<td>[-.880, .095]</td>
</tr>
<tr>
<td>[Treatment=Gender Discussion]</td>
<td>0 *</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>[Guise=Female]</td>
<td>-.127</td>
<td>.292</td>
<td>-0.435</td>
<td>.664</td>
<td>[-.705, .450]</td>
</tr>
<tr>
<td>[Guise=Male]</td>
<td>0 *</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>[Gender=Female]</td>
<td>-.008</td>
<td>.303</td>
<td>-0.028</td>
<td>.978</td>
<td>[-.606, .589]</td>
</tr>
<tr>
<td>[Gender=Male]</td>
<td>0 *</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>[Guise=Female] × [Gender=Female]</td>
<td>-.782</td>
<td>.406</td>
<td>-1.910</td>
<td>.058</td>
<td>[-1.590, .027]</td>
</tr>
<tr>
<td>[Guise=Female] × [Gender=Male]</td>
<td>0 *</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>[Guise=Male] × [Gender=Female]</td>
<td>0 *</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>[Guise=Male] × [Gender=Male]</td>
<td>0 *</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

* This parameter is set to zero because it is redundant.
Appendix Table 7. Variations between collaborative speech responses and sexism

<table>
<thead>
<tr>
<th>Parameter</th>
<th>B</th>
<th>Std. Error</th>
<th>t</th>
<th>Sig.</th>
<th>95% Confidence Interval</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.286</td>
<td>0.482</td>
<td>-0.593</td>
<td>0.554</td>
<td>-1.238</td>
<td>0.666</td>
<td></td>
</tr>
<tr>
<td>[Treatment=Disney Discussion]</td>
<td>0.536</td>
<td>0.181</td>
<td>2.982</td>
<td>0.004</td>
<td>1.179</td>
<td>0.894</td>
<td></td>
</tr>
<tr>
<td>[Treatment=Gender Discussion]</td>
<td>0*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Guise=Female]</td>
<td>1.271</td>
<td>0.579</td>
<td>2.194</td>
<td>0.030</td>
<td>1.127</td>
<td>2.415</td>
<td></td>
</tr>
<tr>
<td>[Guise=Male]</td>
<td>0*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Gender=Female]</td>
<td>-1.107</td>
<td>0.158</td>
<td>-0.778</td>
<td>0.449</td>
<td>-0.419</td>
<td>0.205</td>
<td></td>
</tr>
<tr>
<td>[Gender=Male]</td>
<td>0*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sexism</td>
<td>-0.064</td>
<td>0.193</td>
<td>-0.353</td>
<td>0.725</td>
<td>-0.425</td>
<td>0.296</td>
<td></td>
</tr>
<tr>
<td>[Guise=Female] * Sexism</td>
<td>-0.470</td>
<td>0.234</td>
<td>-2.012</td>
<td>0.046</td>
<td>-0.912</td>
<td>-0.059</td>
<td></td>
</tr>
</tbody>
</table>

a. This parameter is set to zero because it is redundant.

Appendix Table 8. Variations between competitive and collaborative speech components, guise. Sexism mean value included

<table>
<thead>
<tr>
<th>Parameter</th>
<th>B</th>
<th>Std. Error</th>
<th>t</th>
<th>Sig.</th>
<th>95% Confidence Interval</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.720</td>
<td>0.199</td>
<td>3.813</td>
<td>0.000</td>
<td>0.347</td>
<td>1.092</td>
<td></td>
</tr>
<tr>
<td>[Treatment=Disney Discussion]</td>
<td>-0.585</td>
<td>0.179</td>
<td>-3.270</td>
<td>0.001</td>
<td>-0.938</td>
<td>-0.232</td>
<td></td>
</tr>
<tr>
<td>[Treatment=Gender Discussion]</td>
<td>0*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Guise=Male]</td>
<td>-0.377</td>
<td>0.150</td>
<td>-2.520</td>
<td>0.013</td>
<td>-0.672</td>
<td>-0.062</td>
<td></td>
</tr>
<tr>
<td>[Guise=Female]</td>
<td>0*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Gender=Female]</td>
<td>-0.112</td>
<td>0.148</td>
<td>-0.780</td>
<td>0.448</td>
<td>-0.404</td>
<td>-0.179</td>
<td></td>
</tr>
<tr>
<td>[Gender=Male]</td>
<td>0*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. This parameter is set to zero because it is redundant.
Appendix figure 1. Treatment differences between competitive speech responses. Sexism included as mean value.