Cognitive demands of gender-neutral language: the new genderless pronoun in the Swedish language and its effect on reading speed and memory

Hellen Vergoossen

Advisor: Marie Gustafsson Sendén
MASTER’S THESIS, 30 HP, VT 2015

STOCKHOLMS UNIVERSITET
PSYKOLOGISKA INSTITUTIONEN
COGNITIVE DEMANDS OF GENDER-NEUTRAL LANGUAGE: THE NEW GENDERLESS PRONOUN IN THE SWEDISH LANGUAGE AND ITS EFFECT ON READING SPEED AND MEMORY

Hellen Vergoossen

The gender-neutral pronoun *hen* has been added to the Swedish language with the aim to reduce sexism in the language and make language more inclusive. An objection against the implementation of the word is that the word would be cumbersome to read and would take more cognitive capacity to process. The present study examined if that concern is warranted. 209 participants self-paced read five texts with three pronouns each. Participants were randomly distributed to conditions containing only *hen*, *he/she*, *she*, or *he* as pronouns. No difference was found between reading speed in the condition containing *hen* as a pronoun in comparison with the conditions using *han*, *hon* or *han/hon*. No overall difference in memory recall was found either. However, sexism was found to be a moderator for memory recall when reading a text including *hen*, leading to lower memory recall for individuals scoring high on modern sexism.

*Keywords*: gender-neutral language, *hen*, reading speed, memory recall

Consider the following riddle cited in Reynolds, Garnham, and Oakhill (2006):

A man and his son were away for a trip. They were driving along the highway when they had a terrible accident. The man was killed outright, but the son was alive, although badly injured. The son was rushed to the hospital and was to have an emergency operation. On entering the operating theatre, the surgeon looked at the boy, and said: “I can’t do this operation. This boy is my son.” How can this be? (Sanford, 1985, p. 311).

After reading this riddle, several options may come to mind. Perhaps the man in the accident is not the boy’s father, or maybe the boy had two fathers, before considering that the surgeon may be the mother. This example demonstrates the power words – in this particular example role titles – have on our mental representations. While the role title ‘surgeon’ in English can refer to individuals of all genders, it is automatically associated with a male referent because of the stereotype information associated with the word (Gygax et al., 2008). In recent years the male bias demonstrated in this example has been the target of many linguistic reforms. As a result, there are rules and regulations discouraging sexist and excluding language, such as in the Publication manual of the American Psychological Association (APA) (2012) and in Recommendation R 4 of the Council of Europe (1990). The APA recommends, for example, to eliminate the masculine generic use of ‘he’. However, rebuilding language to be non-sexist and including is still a work in progress.
Sweden is currently at the forefront of non-sexist language reforms (Milles, 2011). *Hen* has been added to the Swedish language to fulfil the need of a gender-neutral pronoun. While *hen* is neither nationally or internationally the first linguistic reform suggested for non-sexist and gender-including language, it is the first that can be used in both gender transgressing contexts, anonymising contexts, when gender is not known and in indefinite and generic references (Ledin & Lyngfeldt, 2013). At the same time it does not suffer from the male bias seen in amongst others masculine generics. This is for example demonstrated in Stahlberg, Sczesny and Braun’s (2001) study in which masculine generics caused the retrieval of fewer female than male exemplars of a certain role, for example favourite athletes or favourite musicians, from memory in comparison to neutralizing and feminine-masculine role name pairs. *Hen* also doesn’t suffer from the “order effect” in which the word first mentioned shows a linguistic norm like in the he/she or the uncommon she/he alternatives (Bruckmüller et al., 2012). The “order effect” in this context refers to the masculine form most commonly being mentioned first in the he/she-form. However, there are concerns that *hen* will increase cognitive load and may interfere with reading comprehension. There is as of yet no research done on *hen*’s cognitive demands and this study aims to provide insight.

**Gender-fair language strategies**

To understand why *hen* is a worthwhile strategy to consider for the goal of gender-fair, inclusive language, it is important to compare it to previous gender-fair language strategies to consider how it could work alongside them. These strategies have been previously adopted to attempt to make other genders more visible in language. Many of these strategies have been suffering from a male bias and show the need of a new or additional solution for gender-fairness and inclusiveness in language. The strategies can be summarized as 1) Applying masculine generics to all genders, 2) Creating/Encouraging female forms and suffixes, and 3) Gender-neutralization.

The first strategy, ‘applying masculine generics to all genders’ was commonly used for a long time, but has been criticized in recent years as it is regarded as both a symptom and a source of fundamental androcentrism (Braun, 2005). This strategy involves using ‘he’ and masculine role titles to not just refer to males, but to members of other sexes too. Applying masculine generics to all genders makes other genders less visible in language, because they are more readily associated with males (for a review, see Stahlberg et al., 2007). This is the case for both languages with a grammatical gender system (e.g., German; Braun, Sczesny & Stahlberg, 2005), which are languages with very frequent references to gender, and languages with a natural gender system (e.g., English; Gygax et al., 2008), which are languages without grammatical markings of gender.

The second gender-fair language strategy concerns the use of feminine forms of for example work titles. With this strategy, both titles can be used in reference to a person practising a profession, and the feminine form is used for women in a specific role. For example, in grammatical gender languages such as German, a teacher could be labelled as either *Lehrer* (masculine form) or *Lehrerin* (feminine form). This strategy has as a downside that women referred to with a feminine form of a job title are rated as less competent than women referred to with a masculine form (Formanowicz et al., 2013), at least when implementation of these forms into the language has happened recently. Formanowicz and colleagues also found that the negative evaluations of the feminine form was strongest among people high on sexism. The discounting of women referred to with a feminine role title instead of a masculine role
Individual characteristics can indeed influence the interpretations of gender and language. In Parks & Roberton’s study (2005) it was found that attitude toward women, measured with Tougas et al.’s Neosexism Scale (1995), partially mediated the gender effect for attitudes to sexist language. In Formanowicz et al. (2013) study mentioned earlier, a moderating effect of political attitudes was found on the evaluation of job applicants. Specifically, if female applicants were referred to with a feminine job title instead of a generically intended masculine job title, they were rated less positively by participants identifying as conservative than by those identifying as liberals. The scale used to measure political attitudes contained a scale measuring modern sexism. In the present study it could then be hypothesized that for individuals in whom there is a “manifestation of a conflict between egalitarian values and residual negative feelings toward women”, (Tougas et al., 1995, p. 843) which is the definition of neosexism on which validated modern sexism scales are based, (e.g., Ekehammar, 2000) there may be an interaction with reading speed and memory recall.

The third gender-fair language strategy is language neutralization, and this is the category hen falls into. An example of a previous strategy is the replacement of male generic pronouns with “he or she”-forms. This strategy is observed to lead to a more symmetrical choice of male and female referents, even when presented in combination with a stereotypically masculine context, (Rothmund & Scheele, 2004), but this effect is not observed in all studies (Wojahn, 2013). However, using double forms is not a wholly inclusive strategy, as it does not include a reference to intersex and trans*individuals, and it is also a rather uneconomical solution as it takes longer time to write and read. There also appears to be an order effect, as mentioned earlier, with the ‘he’ in the “he or she”-form coming first showing visuospatial androcentrism according to Hegarty & Buechel (2006). Hegaty and Buechel studied gender comparisons in scientific articles, and found that men was presented first, meaning to the left or above the women in the text, in 74% of the comparisons, as opposed to 26% in which the information on females was presented before males.

In a study by Bäck and colleagues (2015) it was found that a lexically neutral role noun (‘the applicant’) despite supposedly being neutral, still evoked a male bias, unless it was referred to with hen. In this study it is suggested that hen evokes more gender-neutral mental representations than other gender-neutral strategies, because it is subjected to the male bias to a lesser degree. Despite not having received a great deal of research yet, hen then seems like a promising solution for gender-fair and inclusive language.

The new pronoun has also evoked many negative reactions, amongst others because its association with Swedish feminist movements. Opponents of the implementation of hen have had similar arguments as opponents of the earlier mentioned linguistic reforms, such as having doubts about the value of such an “artificially” implemented word, arguing that words can be changed, but that this won’t influence what properties are associated with genders (Maass, Suitner & Merkel, 2012). Braun (2005) calls this the ‘arbitrary’ view on feminist language critique. This view suggests that there is no relation between grammatical gender and gender roles, i.e. that language doesn’t carry information about gender and doesn’t influence gender perceptions. This is the opposite of the semantic view that feminist language
critics maintain and which states that there is an interaction between grammatical gender and gender person reference, reflecting and reconfirming relationships between the genders (Braun, 2005), which is supported by the research mentioned earlier.

In the case of *hen*, even official institutes have called for caution with the use of the word. Språkrådet, the Swedish Institute for Language and Folklore, stated in 2011 that “it looks odd, and it is hard to believe that such a common word could be introduced as a pronoun without support in the spoken language” (in Milles, 2013, p. 123). This statement was later replaced by a more accepting statement in which it was stressed that *hen* doesn’t involve any language barriers, but that it can cause irritation amongst the readers, so other gender neutral writing strategies can be considered if it would decrease focus on the text’s content (Språkrådet, 2013).

The idea that *hen* may interfere with reading comprehension is an argument that would fit Blauberg’s categorizations of different arguments against gender fair amendments of language (1980). Blauberg investigated word reforms and what types of arguments that were put forward by antagonists. Eight categories were found, among which the *Change is Too Difficult, Inconvenient, Impractical, or Whatever* category. To this category belong other beliefs such as that language, in particular gendered pronouns, are too deeply ingrained in the language’s fabric to be changed (Parks & Roberton, 1998).

Caution and criticism hamper the acceptance of the word in the Swedish language. It is important to investigate the validity of these points of criticism. In the present study the focus will lie on the question if *hen* interferes with cognitive processes involved in the reading and remembering of a text. What cognitive costs arise when a text with *hen* is read? Does it lead to a slower reading speed and worse memory recall of the text’s contents?

**Reading behaviour and eye-tracking evidence**

One of the common arguments against using *hen* has been that it interferes with the understanding of a message. The first way the effect of *hen* on reading behaviour will be measured is by measuring reading speed.

When encountering an incongruity between the word that is expected and the word that is present (also called a ‘mismatch’) as may be the case when reading a man-suffix term (for example *chairman*) and a feminine reflexive pronoun (for example *she* or *her*), eye movements back to the initial man-suffix term (regressions) can be expected and were observed in for example Khan & Daneman’s study (2011). In this study much more time was spent looking back at *chairman* when encountering *herself* afterwards than was the case with *chairperson*, supporting the evidence that a gender is immediately ascribed to the subject when encountering it, and the gender is more likely to be male. This reading difficulty, which occurs when encountering a mismatch, is called the ‘mismatch-cost’ (Kreiner et al., 2008). Either this effect is caused by the influence of the lexical information (the presence of “man” in the noun), or the stereotypical association between the role of chairman/chairperson with male referents. *Hen* observed in this situation could lead to a similar mismatch if gender initially gets derived from the role noun and *hen* conflicts with the expectation of a gendered pronoun.

In studies using eye tracking it has been found that when encountering incongruities, the reader’s fixations increase and the reader is more likely to regress to earlier parts of the text and re-read (Warren, 2011; Rayner et al., 2006), which also takes longer time in total of
reading a text. If the presence of *hen* in a text increases the difficulty of reading and processing the text, this should be reflected by longer reading times.

While the present study does not contain eye-tracking measurements, previous evidence aids in predicting what effect on reading speed *hen* will have. Additionally, it strengthens the position of reading speed as a valid variable for measuring reading disruption in the present study. An example of a study connecting mismatches to slower reading speed is Kennison & Trofe’s study (2003) in which either role titles stereotypically strongly associated with males and females (for example *executive* or *secretary*) were presented in a sentence and in a later sentence referring back to with a pronoun congruent with this stereotype, or incongruent. They found that reading time was significantly slower when gender of the pronoun mismatched the gender stereotype of the preceding role title.

The expectation of the present study is that either the expectation of another pronoun than *hen* will cause this mismatch-cost, or the novelty of *hen*. No eye-tracking information has been gathered on *hen* as of yet, but it can be defined as a linguistic anomaly to those who are least familiar with it. Novel words are more likely to be fixated and have longer reading times than familiar words. Wochna and Juhasz’s (2013) demonstrate this in their study where they presented in a paragraph one of two words with the same meaning, of which one is common in the English language and one pseudoword. The novel words were more likely to be fixated and had longer reading times than the words participants were familiar with. Based on this finding, *hen* can be expected to take longer reading time than the common, gendered pronouns.

While the previous studies solely focused on reading speed and eye movement, the present study adds another variable to encircle the concept of cognitive load.

**Cognitive load and memory recall**

Cognitive load is a central concept in the present study as it is assumed to be affected when reading a text with *hen*. *Hen* is supposed to add additional strain to the processing of the text, possibly inhibiting the comprehension of the rest of the text. In the present study cognitive load is measured alongside reading speed. Whereas reading speed is a measurement of the time it takes to process a text, a measurement of cognitive load generated by *hen* shows if *hen* also interferes with the understanding and retaining of information of the text. *Hen* could increase the time it takes to read the text, but this does not necessarily mean it interfered with the comprehension of the text. In the present study this is accounted for with a memory task, on which a lower performance would be an indicator of lower reading comprehension.

Cognitive load is a construct with three measurable dimensions: mental load, mental effort and performance (Kirschner, 2002). The mental load is the aspect of the cognitive load that is exclusively increased by task- and environmental demands. This is the aspect of cognitive load the present study is concerned with. Working memory (WM) is the part of the executive function involved with encoding, activating, storing and manipulating information while performing cognitive tasks (Diamond, 2013). WM is crucial in reading as reading requires keeping words in short-term memory to make sense of longer strings of words, like sentences. Reading comprehension involves creating complex mental representations of the words read that have been referred to as *situation models* (Zwaan & Radvansky, 1998). Novel information, such as novel words like *hen*, can lead to a higher cognitive load as it must be processed in working memory in order to construct such models in long-term memory (van
Merriënboer & Sweller, 2010). While fully automated models drawn from the long-term memory don’t put strain on WM, dealing with novel information for which no schemas are available, WM has limitations.

WM theories provide a way to operationalize the construct of cognitive load for the present study. Because a common assumption of WM models is that a limited amount of information can be simultaneously processed (Baddeley, 2003) it then leads to the assumption that increases in mental load are associated with reduced performance in reading comprehension.

A way to test cognitive load in reading is by examining the impact of another task on reading performance. Interruptions, which *hen* is hypothesized to cause, are an example of a way to increase cognitive load. Literature on the effect of interruptions on reading comprehension is consistent with the view that interruptions will decrease reading comprehension. This is for example demonstrated in Lorch’s study (1993) in which being interrupted led to slower reading times for the sentences following the interruption, especially when no information was provided about where reading had been interrupted. Lorch suggests that after an interruption important text information must be reinstated in working memory as to successfully continue the development of the situation model.

Clevinger’s study (2014) investigated the extent to which loads on working memory capacity influenced how people read and comprehend written information. Specifically emotional arousal, which is likely to act as a cognitive load if arousal is too high, was investigated. It was found that arousal induced via emotionally charged words (such as ‘cancer’) negatively influenced reading performance. These findings are not based on cognitive load in general, but on the semantic value of words processed that led to comprehension difficulty. In this study the measurement of reading comprehension (in the present study memory recall) was operationalized similarly to the present study.

This is in line with earlier research investigating the effect of emotional word stimuli’s interference with performance if the emotional words were processed semantically, for example Huang et al.’s study (2008) in which emotionally charged words (for example ‘sick’) captured more attention than neutral words.

At the present there are no studies that have tested whether *hen* takes more cognitive resources than another pronoun and the present study intends to fill this gap.

**Study objectives**

The overall aim of this study is to investigate if *hen* interferes with the processing of text. If *hen* creates a mismatch between the expected pronoun and the pronoun present, or if it suffers from the costs associated with the presence of a novel word, then reading a text with *hen* as a pronoun should take more time than it would for the texts with other pronouns and memory recall should be lower. If *hen* is not processed differently, information in the text should be processed at the same speed and information about the text should be retained as well as for texts with other pronouns.

**Hypotheses:**

1. Reading speed will differ between experimental conditions, with the condition containing *hen* taking more time to read than the other conditions.
2. Memory recall will differ between experimental conditions, with less questions being answered correctly after reading the texts in the condition containing *hen* in comparison with the other conditions.

3. The level of sexism will interact with reading speed, such that individuals scoring high on sexism to take a longer time reading the texts including *hen* than individuals scoring low on sexism.

4. The level of sexism will interact with memory recall, leading to individuals scoring high on sexism answering fewer memory questions correctly when reading texts including *hen* than individuals scoring low on sexism.

**Methods**

*Participants*

209 participants (144 women, 53 men, 3 non-binary, 9 unknown) with the mean age of 33.47 years (*SD* 9.4 years) completed a survey experiment. An additional 120 individuals started on the questionnaire but didn’t finish it, of which the majority quit the questionnaire at the very beginning. Participants were recruited through social media and other online platforms. The questionnaire was hosted by Qualtrics, an online survey software ([www.qualtrics.com](http://www.qualtrics.com)), and distributed online. The participation took 20 minutes; no compensation was offered for participation.

Consent was required before starting the questionnaire. 3 participants were excluded from analyses because they were below the legal age of majority in Sweden. Another 11 participants were excluded because they spent such a short time on the pages displaying the texts that it seemed highly implausible that they had read them, also invalidating subsequent performance on memory questions. 3 participants were excluded because several of their reading speed values exceeded the third quartile with 1.7 times the IQR, which raises the concern that the individuals were distracted while reading the text or another variable was interfering with their self-paced reading speed.

*MATERIALS AND PROCEDURE*

The questionnaire was introduced as a test of reading comprehension. Participants were randomly distributed to a condition in which all the texts at the beginning of the questionnaire contained either only the pronoun *hen* (genderless pronoun), *han* (“he”), *hon* (“she”) or *han/hon* (“he/she”). A total of 51 participants were distributed to the *hen* condition, 52 to the *han* condition, 44 to the *hon* condition, and 62 to the *han/hon* condition. There were five texts on neutral subjects such as the planting of a tree and the golden rule contained roughly 550 words and three pronouns each. All of the pronouns used in the texts were singular, generic forms. That is, the pronouns were not used to refer to a real person. The order of the texts was counterbalanced.

Prior to the present study, a small pilot was conducted to get feedback on the valence, level of interestingness and familiarity of the subjects discussed in the texts, the appropriateness of the level of difficulty of the recall questions, and the definition of the texts’ content as female or male.

An example of a text is given below (own translation from Swedish, bold emphasis added to show the location of the pronouns).
It is important that a person that wants to plant a tree carefully picks a spot where the tree shall stand. Then he/she needs to dig a big pit that is at least 3 times the diameter of the root clump and twice as deep.

Next, he/she should mix the excavated soil with potting soil, about half of each. The person will then need to fill up the hole with freshly mixed soil and make sure the tree is at the correct height. It is better if he/she plants the tree a little high as the soil will settle and the tree feels better from standing a bit higher rather than sunk.

Then the person fills up with the mixed soil so the tree is stable, after that the tree should get water.

**Procedure.** After self-paced reading all the texts, three content based multiple-choice questions per text followed, as to test the participant’s ability to recall information from the texts. To provide the reader with some context, every section had the same title as the one the text was presented with. Initially the memory section also included free recall questions where the respondents were asked to freely write down what they recalled from the text, but these were removed right after the release of the questionnaire as it led to participants quitting the questionnaire and a poor response rate.

**Devices.** Before running further analyses the effect of devices on reading speed and memory recall were checked. The most common devices used to fill out the questionnaire on were laptops or the equivalent tablets with keyboard (39.2%) and smartphones (36.8%). Other devices used were a tablet without a keyboard (8.1%) and a stationary computer (11.5%). There was no significant difference between the standardized reading times when filling out the questionnaire on different devices, $F(3,196) = 2.137, p = .097$, and no significant difference for the amount of memory questions answered correctly, $F(3,196) = 1.723, p = .164$.

**Variables.** The first dependent variable was the speed with which the texts were read. Reading speed was measured as time spent on the page the individual text was displayed on by the Qualtrics software. Reading speed was standardized by dividing the reading speed with the amount of characters in the text. This method has previously been used by Irmen and Roßberg (2004).

The performance on the memory questions was the second dependent variable, hereafter named ‘memory recall’. Memory recall was measured by 3 multiple choice questions per text, so 15 in total. The percentage of correct answers on all questions was computed.

Sexism was measured by a subscale of the Classical and Modern Sexism Scale (Ekehammar et al., 2000), a scale specifically developed for a Scandinavian context, was used to assess level of sexism. The subscale ‘Modern Sexism’ was included in the present study. In total the scale consists of 8 items such as “the government puts too much emphasis on women’s issues” and “society treats men and women the same way”, accompanied by a 7-point scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). In the present study, the item “The school curriculum should be adapted to girls’ needs“ was left out because it didn’t fit in with the rest of the scale. For the current sample the Cronbach’s $\alpha$ was .84.
Results

Reading speed
Mean reading times were standardized by dividing total reading time with the number of characters in the respective text (Irmen & Rossberg, 2004). The average standardized reading speed was 61 milliseconds per symbol (SD = 20), which translates to 33.32 seconds (SD = 11.19) for an average 550-symbol text. Means per experimental condition are displayed in Table 1.

Table 1. Means of the standardized reading speed divided on experimental condition (he/she, hen, she, he) in milliseconds, standard deviation within parentheses.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Reading Speed</th>
</tr>
</thead>
<tbody>
<tr>
<td>He/She</td>
<td>65 (22)</td>
</tr>
<tr>
<td>Hen</td>
<td>62 (20)</td>
</tr>
<tr>
<td>She</td>
<td>59 (21)</td>
</tr>
<tr>
<td>He</td>
<td>55 (17)</td>
</tr>
</tbody>
</table>

First, the hypothesis that reading speed would differ between conditions was tested. Reading the five texts was done fastest in the condition in which he was used as a pronoun. In the feminine generic she condition, the texts were read slightly slower. The texts were read slightly slower still in the gender-neutral hen condition and slowest in the he/she condition. While these differences were not significant, as determined by a one-way ANOVA with texts as independent factors and reading speed as the dependent variable, $F(3,205) = 2.175$, $p = .093$, it does show a tendency to spend more time on texts containing hen and he/she than texts containing she or he.

To reveal if there was any effect of the background variables on reading speed, the independent variables text version, motivation to perform well and modern sexism were regressed on the dependent variable reading speed. Age was added as a control variable. Pearson correlations are displayed in Table 2.

Table 2. Reading speed, memory recall and background variables: Correlations

<table>
<thead>
<tr>
<th>Variables</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Reading Speed</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Memory Recall</td>
<td>.15*</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Age</td>
<td>-.07</td>
<td>-.15*</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Motivation to perform well</td>
<td>.23**</td>
<td>.17*</td>
<td>-.15*</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>5. Modern Sexism</td>
<td>.11</td>
<td>-.08</td>
<td>.01</td>
<td>.09</td>
<td>-</td>
</tr>
</tbody>
</table>

Note. * Significant at the .05 level (2-tailed)  
** Significant at the .01 level (2-tailed)

A multiple regression was performed. A significant model emerged ($F(4,197) = 4.283$, $p = .002$). Table 3 shows information for each predictor. The independent variable Motivation to perform well was a significant predictor of reading speed, where a higher motivation to perform well led to more time spent on reading the texts. There were no effect of the text version, and not any main effect of modern sexism either.
Table 3. The regression coefficients for the variables entered into the model to analyse how they predicted reading speed.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>B</th>
<th>SE B</th>
<th>β</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text version *</td>
<td>-.002</td>
<td>.003</td>
<td>-.046</td>
</tr>
<tr>
<td>Age</td>
<td>.000</td>
<td>.000</td>
<td>-.087</td>
</tr>
<tr>
<td>Motivation to perform well</td>
<td>.005</td>
<td>.001</td>
<td>.244*</td>
</tr>
<tr>
<td>Modern Sexism</td>
<td>.002</td>
<td>.001</td>
<td>.093</td>
</tr>
</tbody>
</table>

*Note. * Text version: 1 = text containing *hen*, 2 = text containing other pronoun

* Significant (*p* < .001)

In order to test the hypothesis that level of sexism might affect the reading speed for the different text versions, a moderation model was tested. Reading speed was the dependent variable, the text version (containing *hen* versus not containing *hen*) independent variable and sexism the moderator.

To investigate the moderating effect of the independent variables on reading speed, the Hayes Process Macro was used. With this macro, the indirect moderating effect of the variables was computed by calculating the product of coefficients (Hayes, 2012; 2013; Model 1) with 95% bias-corrected bootstrapped confidence interval (CI). This analysis revealed that there was no significant direct effect of sexism on reading speed for the individuals reading the texts with *hen* as a pronoun (*b* = -.002, *SE*(boot)= .003, CI= [-.012, .001]).

**Memory recall**

Memory recall was calculated as the percentage of memory questions answered correctly. On average, 74.55% of all questions were answered correctly (SD = 13.47). This translates to just over 11 out of 15 questions. Means per experimental condition are displayed in Table 4.

Table 4. Means of the memory recall divided on experimental condition (*he/she, hen, she, he*), standard deviation are given within parentheses.

<table>
<thead>
<tr>
<th></th>
<th>Memory Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>He/She</td>
<td>75.35% (12.17)</td>
</tr>
<tr>
<td>Hen</td>
<td>78.07% (11.60)</td>
</tr>
<tr>
<td>She</td>
<td>71.82% (16.28)</td>
</tr>
<tr>
<td>He</td>
<td>72.45% (12.34)</td>
</tr>
</tbody>
</table>

When a longer time was spent on reading the text, generally the performance on memory questions was slightly better (*r*(207) = .15, *p* < .05) as well as when a participant rated themselves as motivated to perform well on the test (*r*(207) = .17, *p* < .05).

The hypothesis that memory recall would differ between conditions was tested. There was no difference in memory recall between the different pronoun versions (*F*(3,208) = 2.305, *p* = .083).

To reveal if there was any effect of the background variables on memory recall, a regression analysis was performed where the memory recall was the dependent variable, and the background variables text version, age, motivation to perform well and modern sexism as independent variables. A significant model emerged (*F*(4,197) = 4.879, *p* = .001).
shows information for each predictor. The condition, which was recoded as either a version containing *hen*, or a version not containing *hen*, was a significant predictor of memory recall, where a text containing *hen* would lead to a better performance on memory questions. *Motivation to perform well* was a significant predictor of memory recall, where a higher motivation to perform well led to more memory questions answered correctly. *Age* was also a significant predictor of memory recall, where a higher age led to less memory questions correctly.

Table 5. The regression coefficients for the variables entered into the model to analyse how they predicted memory recall.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>B</th>
<th>SE B</th>
<th>β</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text version</td>
<td>-5.399</td>
<td>2.079</td>
<td>-.177*</td>
</tr>
<tr>
<td>Age</td>
<td>-.233</td>
<td>.095</td>
<td>-.168*</td>
</tr>
<tr>
<td>Motivation to perform well</td>
<td>2.085</td>
<td>.829</td>
<td>.172*</td>
</tr>
<tr>
<td>Modern Sexism</td>
<td>-1.426</td>
<td>.828</td>
<td>-.118</td>
</tr>
</tbody>
</table>

*Note.* Text version: 1 = text containing *hen*, 2 = text containing other pronoun

* p < .05

In order to test the hypothesis that level of sexism may affect the memory recall for the texts containing *hen* and the texts not containing *hen*, a moderation model was tested. Memory recall was the dependent variable, the text version (containing *hen* versus not containing *hen*) independent variables and sexism the moderator.

To investigate the moderating effect of the independent variables on reading speed, the Hayes Process Macro was used again. This analysis revealed a marginally significant model (p = .07) with a significant effect (p = .02) of sexism on memory recall for the individuals reading the texts with *hen* as a pronoun (b = -4.756, SE(boot) = 1.994, CI = [-8.689, -0.823]), with individuals scoring high on sexism having a lower memory recall.

**Discussion**

*Study findings*

The aim of this study was to test whether texts including *hen* were associated with higher cognitive load. Concerns have been expressed that the use of *hen* may lead to difficulty reading and comprehending texts, which has been used as an argument for objection to further integration of the word into the Swedish language. The present study was the first study to investigate if such concerns are warranted.

The concept of cognitive load was operationalized with two measurements: reading speed and memory recall. Regarding reading speed it was hypothesized that texts containing the pronoun *hen* would be read slower than texts containing the pronouns *he, she,* and *he/she*. No evidence that it takes more time to read texts containing *hen* was found. It was also hypothesized that sexism would be a moderator for reading speed, with higher scores on *Modern Sexism* leading to longer reading times on texts containing *hen*. Such interaction was not observed in the present study. The argument that texts containing *hen* take more time is
not supported by these results, at least when *hen* is used as a generic reference in reference to an undefined individual.

The second measurement in the study design aimed to deduct the influence of *hen* on cognitive load were the memory recall questions. It was hypothesized that reading texts containing *hen* would lead to a lower performance on memory recall questions, but such effect was not found. There was no evidence that *hen*, whether because of novelty or a mismatch, interfered with memory recall. If anything, individuals that read texts containing *hen* as the pronoun seemed to perform slightly, yet not statistically significantly, better on memory recall. This can be speculated to be due to a ‘spotlight effect’ caused by variables such as the emotions experienced when reading the word, or the novelty of the word, where the latter has been established to lead to a bigger chance of fixation and has longer reading times than familiar words (Wochna & Juhasz, 2013). In addition, it could be explained by the emotional arousal evoked by the word, which is in line with research showing that both positive and negative emotional stimuli are likely to capture attention, which in Gotoh’s study (2008) led to increased response times for same-different decisions.

In addition to this, sexism was found to be a moderator for memory recall, with individuals scoring high on sexism answering fewer questions correctly. While this was only a moderately significant effect, it’s an interesting starting point for future research. What exactly the cause of *hen*’s disruption of memory recall is, especially without an observable effect on reading speed, can at this point only be speculated about and requires further research. The lack of increase in reading time could for example be due to the shortness of the word and the relatively little attention short words tend to get with fixations (Blythe & Joseph, 2011). It could also be that some negative emotion is evoked by the word, leading to less attention on further text, but perhaps not enough emotion to effectively cause time-consuming regressions to earlier text or longer fixations. There could also be other variables connected to a high score of sexism leading to mental representations that cause a bigger disruption when coming across a genderless pronoun, such as less ‘flexible’ mental representations of stereotypically gendered role titles or simply less exposure to the word. The cost associated with *hen* may also get lower the longer the word has been present in language and the less offensive and explicitly connected to the feminist movement it becomes.

With ‘noise’ in the reading speed and memory recall variance caused by not controlling the devices participants used to complete the questionnaire, it is possible (larger) effects of *hen* on reading speed and memory recall remained veiled. Larger effects may also be found when *hen* is presented in different contexts than just being a generic reference as it was in the present study. This may have meant that no mismatch was caused because *hen* didn’t refer to a person or a stereotyped role, while in other contexts it would have (see Garnham, 2002), for example, if *hen* was used to refer to a real person. *Hen* can be used also to refer to trans*persons, to a single individual of unknown sex, to an individual whose gender is irrelevant, and even to cisgender individuals as a substitute of gendered pronouns. These ways of using *hen* are evoking the most heated debate (Milles, 2011) and could cause more evoked emotion when presented in a text. There’s the concern by some that people who do identify as cisgender aren’t allowed to identify themselves as such anymore. With such contexts lacking in the texts used in the present study, perhaps the violation provided wasn’t large enough to detect any differences in reading speed or memory recall. Placing what
would normally be a semantic violation into a context in which it does not refer back to a concrete noun or person can partially and even entirely avoid disruption to it (see Ferguson & Sanford, 2008).

Limitations and future research
Central in the present study was the assumption that a slower reading speed reflected a higher requirement on language processing resources, but this does not necessarily have to be true. A rather crude way of measuring reading time was used. This may suffice in this study in which it is not so important when the effect is occurring. An issue however, is that when measuring word processing times, as was attempted for the pronouns in the present study, it is assumed that every word is being fixated, which is not the case in reality. It is suggested by some that some words are not being fixated at all (Rayner et al., 2006), while others maintain that every word is fixated, but sometimes together with another word (Just et al, 2013). Despite this disagreement, it can be agreed that there is variation in how much attention is devoted to a word individually, whether it is fixated at all or together with another word. The standardized reading time used in the present study, which was the total reading time of a text divided by the number of symbols in the text, would accurately reflect (word) processing time if it was certain every word was fixated. Without eye tracking data we cannot be sure if the pronouns in the hen condition were even looked at, especially when lacking a control question asking what pronoun was observed in the texts prior. Perhaps the presence of hen was not registered.

In future research it is then important to use eye tracking material to observe what is going on “behind the curtain” and take measurements of first-fixation duration (the duration of the first fixation in a region), single-fixation duration (cases when only one fixation is made on a word), gaze duration (the sum of all fixations on a word prior to moving to another word), and total fixation time (the sum of all fixations, including regressions, on a word) to establish if a mismatch occurs and exactly where it occurs. In addition, the probability of fixating on hen and the frequency of regressions away from hen should also be investigated. Such measures are also highly relevant to investigate the different usages of hen, such as references to trans*persons or anonymous individuals. The different contexts in which hen can be used as a reference may differ in eye behavior. It can also provide new insights in what happens when no gender can be derived from lexical information as we so habitually attempt to do (e.g. Kreiner et al., 2008).

While devices did not have a significant effect on the reading speed and memory recall, it would be desirable to standardize the apparatus used to read texts, especially when devices differ in amount of scrolling needed, responsiveness and presence or absence of a mouse.

Individual differences among readers
Another question that received relatively little attention in the present study is how text-external factors such as individual differences besides sexism may have affected the reading of the texts. For example, in Daneman et al. (1983) it was found that individuals with less skill in reading were less likely to detect anomalies than individuals with more skill in reading. Because reading is a learned skill there is variability between individuals. There may be variability in individual’s ability to temporarily store and process verbal information. Individuals with smaller reading comprehension have also been observed to have a reduced ability to integrate text with stored knowledge (e.g., Seigneuric et al., 2000). Because of these individual differences in working memory, it may be the case that if texts containing hen
increase cognitive load, people with less working memory capacity should experience more difficulty reading *hen* in texts than those with more capacity. In future studies it should therefore be considered to include a measurement of quality of (verbal) working memory.

**Conclusion**

This is the first study testing whether *hen* is associated with an increased cognitive load. Having a good understanding of how gender-neutral language influences people is valuable when navigating the debate around the question if *hen* should be integrated further into the language. In the present study no evidence was found that the presence of the word interfered with reading speed or memory recall, at least when it is used as a general reference to a mixed and undefined group of people. The argument that *hen* should be used cautiously as to not distract the reader is not supported by the present study.

The moderating effect of sexism on memory recall for texts containing *hen* makes sense in the light of previous findings of individuals scoring high on sexism also being unaccepting of feminist language planning. However, considering that *hen* lead to heated debate only 3 years ago, the effects found are not large. Integrating *hen* into the language may be less problematic than has initially been thought, and with the present study showing there is no need for concerns about *hen* complicating the reading of a text, we may be one step closer to the successful integration of *hen* into the Swedish language.
References


Språkrådet. (2013). *Pronomenet hen*. [The pronoun *hen*]. Retrieved from [http://www.sprakochfolkminnen.se/download/18.4a09e1b214565240e191148/1398839699856/pronomenet-hen.pdf](http://www.sprakochfolkminnen.se/download/18.4a09e1b214565240e191148/1398839699856/pronomenet-hen.pdf)


