RaY

Research at the University of York St John

For more information please contact RaY at
ray@yorksj.ac.uk

Word Norms and Measures of Linguistic Reclamation for LGBTQ+ Slurs

1.   Introduction

Word norms--averaged, self-report ratings of linguistic stimuli for particular properties--are highly valuable to experimental studies of language; they provide standardised measurements for use as independent variables, enable comparisons between different aspects of emotionality and allow for greater replicability of experimental research within and across research labs (Bradley & Lang 1999: 1).

I conducted a survey in order to produce a set of normed word ratings for British English LGBTQ+ slurs, for use in experimental research investigating the cognitive processing of such language, and the potential effects of linguistic reclamation on this process. Respondents rated slurs for relevant emotional and non-emotional word properties, and measures of reclamation behaviour. The resulting database is novel, both for its specific focus on slurs and for generating quantitative data on slur reclamation. I used this data to reveal a consistent pattern of relationships between word properties and reclamation behaviours, as well as noteworthy differences in ratings based on respondent age, gender identity and sexual identity.

2.   Research Background

I begin by defining the terms most central to my research. I take *taboo language* to refer to words or propositions with meanings that are socially prohibited "for a specific community…at a specifiable time in specifiable contexts" (Allan & Burridge 2006: 11), typically because they are "perceived to be a potential cause of discomfort, harm, or injury" (Allan 2019: 3).

Slurs are a particular class of taboo language, used by more hegemonically powerful (Gramsci 1971; 1985) social groups to derogate (Chen 1998; Croom 2010: 343) members of less hegemonically powerful social groups (Herbert 2015). The inclusion of a reference to the Gramscian concept of *hegemony*--which Baker (2008: 123) neatly summarises as "the exercise of power, whereby everyone acquiesces in one way or another to a dominant person or social group"--is my own. Research has typically defined slurs as just labels which derogate members of particular social groups, but Herbert (2015: 133) argues that slurs specifically reflect "long standing [practices] of systematic dehumanisation, marginalisation, and exclusion from social spaces", stressing that they should be distinguished from other negative labels for groups which do not reinforce these broader power structures.

Linguistic reclamation describes a sociolinguistic phenomenon particular to slurs, which Bianchi (2014: 35) defines as "uses by targeted groups of their own slurs for non-derogatory purposes, in order to demarcate the group, and show a sense of intimacy and solidarity". This definition is useful: it reflects an understanding of slurs as typically derogatory and group-targeting, and it outlines several key purposes reclamation can serve. However, Bianchi's use of "their own slurs" is potentially misleading. Herbert (2015: 131) neatly summarises the problem with framing reclamation as a "taking back" of language, in that it presumes that control of the term--insofar as this is possible to claim--was possessed by the targeted group prior to its reclamation. Therefore, my understanding of linguistic reclamation adjusts Bianchi's definition to provide additional clarity: it refers to non-derogatory uses, by members of marginalised groups, of the slurs used by members of hegemonically powerful groups to derogate them. This understanding positions reclamation as a politically-motivated process, and for this reason it is important to acknowledge that even within targeted groups, not all members engage in reclamation projects equally, and the individual decision to engage or not can be made for a number of personal and political reasons (Brontsema, 2004: 5-12). I therefore consider the reclamation of slurs to be a collective consequence of the individual linguistic choices made by marginalised people, with varying scale and success.

I focused this research on slurs targeting the LGBTQ+ (Lesbian, Gay, Bisexual, Transgender and Queer +) community, as it has a rich history of navigating oppression and marginalisation through its use of language, which includes subjecting a number of LGBTQ+ slurs to processes of reclamation (Baker 2008; Coleman-Fountain 2014; Russell, Clarke, & Clary 2009; White 1980).

A number of databases of word norms for taboo language already exist. One of the most comprehensive of these was produced by Janschewitz (2008), whose study produced normed ratings for American English taboo and non-taboo words, for the following properties: word familiarity, frequency of personal use, emotional valence, emotional arousal, offensiveness, and tabooness. Familiarity simply represents how recognisable a respondent considers a word to be based on their own experience, while personal use

represents how frequently a respondent uses a word themselves. Valence and arousal are two distinct dimensions of emotionality that are common in the literature: the former concerns whether a stimulus is positive or negative, and the latter how stimulating and attention-grabbing it is (Janschewitz 2008: 1065). Offensiveness and tabooness respectively represent how offensive a word is to the respondent themselves, and how socially unacceptable they judge it to be to a more general population (Janschewitz 2008: 1066; Madan, Shafer, Chan, & Singhal 2017: 796).

The relevance of each of these word properties to performance in taboo language experiments has since been proven out in the literature. For example, in a Canadian English study of taboo language performance in lexical decision and free recall tasks, Madan, et al. (2017: 802-804) found that a combination of *non-emotional properties* (e.g. familiarity, personal use) and tabooness explained performance in the lexical decision tasks: faster lexical decision was best predicted by higher word frequency, familiarity, personal use and contextual diversity, but also lower tabooness. Conversely, *emotional properties* like arousal, offensiveness and tabooness best explained performance in the free recall tasks: increases in all of these led to enhanced recallability. Their study also included normed age of acquisition ratings (henceforth AoA; the age at which a word is first learned) as one of the non-emotional properties relevant to lexical decision performance, based on research conducted after the production of Janschewitz's norming study (Brysbaert & Cortese 2011; Juhasz, Yap, Dicke, Taylor & Gullick 2011).

I have found no existing database of word norms specifically for slurs, much less for LGBTQ+ slurs; Janschewitz's data, for example, contains only 7 examples. Similarly, I have also found no word norms for linguistic reclamation. As such, I set out to create the database described in this paper.

3.   **Methodology**

3.1. Respondents

The survey had 155 respondents. Of these, 127 respondents were recruited through social media, particularly through the social media pages of LGBTQ+ groups based in the North-West and East Midlands of England. The remaining 28 were a combination of students at the University of Nottingham--recruited either as a means of gaining module credit or via posters advertising the study--and members of the public recruited via posters placed in local community centres, with permission. To aid recruitment, respondents were offered a chance to enter a prize draw for a small online shopping voucher upon completion of the survey, excepting the students participating for module credit instead. No contact information was retained once the winner had been selected.

All respondents were over the age of 16 and native or near-native speakers of British English. I collected additional respondent demographic data in order to make group comparisons based on age group, gender identity, cisgender status (i.e. whether their gender identity matched the one they were assigned at birth) and sexual identity (see Table 1). For the gender identity and sexual identity questions, respondents were able to give free and open responses, from which I created categories for analysis.

Descriptive analysis indicated that a majority of respondents were aged between 16 and 29, a slight majority identified as women, a large majority of respondents were cisgender and a slight majority of respondents were not heterosexual.

Table 1: Respondent Demographic Data

Age group	*N*	*%*	Gender Identity	*N*	*%*
16-19	36	23.23	Woman	82	52.90
20-29	60	38.71	Man	66	42.58
30-39	17	10.97	Non-Binary	4	2.58
40-49	15	9.68	Agender	1	.65
50-59	17	10.97	Unclear Response	2	1.29
60-69	9	5.81	**Total**	**155**	**100.00**
70+	1	.65			
Total	**155**	**100.00**			

Cisgender Status	N	%	Sexual Identity	N	%
Cisgender	140	90.32	Straight, Heterosexual	69	44.52
Not Cisgender	15	9.68	Gay, Lesbian, Homosexual	43	27.74
Total	**155**	**100.00**	Bisexual,Pansexual,Polysexual	31	20.00
			Queer	5	3.23
			Asexual	4	2.58
			Unclear Response	3	1.94
			Total	**155**	**100.00**

3.2. Design

I conducted a preliminary survey to identify LGBTQ+ slurs for use in this study. This survey asked respondents to freely list as many expressions as they could think of which they considered slurs targeting members of the LGBTQ+ community, and generated 160 unique examples. However, a majority of these (51.85%) were only identified by single respondents. I decided to only include those which had been identified by a minimum of 5 respondents in the preliminary survey; this threshold allowed me to eliminate the most low-frequency slurs, while still leaving enough to use as stimuli in future lab experiments. As I discuss later in this section, respondents to the norming survey were asked not to provide ratings for slurs they did not know; none had an excess of non-responses, demonstrating the threshold was not too low. Accordingly, 41 slurs were included in the norming survey (see Table 2).

Table 2: LGBTQ+ Slurs Included in Norming Survey

Slur	N	Slur (cont.)	N	Slur (cont.)	N
Dyke	86	Fairy	19	Trap	8
Faggot	83	Shemale	19	Cock-Sucker	7
Tranny	78	He-She	18	Gaylord	7
Fag	74	Rug-Muncher	16	Muff-Diver	7
Queer	74	Fudge-Packer	15	Bull-Dyke	6
Poof	69	Pansy	14	Bum-Boy	6
Homo	38	Arse-Bandit	12	Pervert	6
Lesbo	38	Shirt-Lifter	12	Camp	5
Lezzer	37	Bent	10	Hermaphrodite	5
Gay (Pejorative)	25	Batty	8	Queen	5
Bender	23	Bummer	8	Shit-Stabber	5
Gayboy	23	Butch	8	Sissy	5
Batty-Boy	22	Gender-Bender	8	Sodomite	5
Poofter	20	Ladyboy	8		

The first set of survey questions asked respondents to rate LGBTQ+ slurs for 7 word properties identified by both Janschewitz (2008) and Madan et al. (2017): familiarity, personal use, arousal, negativity, offensiveness, tabooness and AoA. I asked respondents for negative valence ratings specifically, for three key reasons. Firstly, slurs being highly derogatory words by definition (Chen 1998; Croom 2010: 343) means their valence ratings are very likely to skew negative, and if participants felt a word was not at all negative, they could provide the lowest rating to indicate this. Secondly, more precise ratings could be given if all points on the scale related to just negative valence, rather than to both positive and negative valence. Thirdly, this decision kept the formatting of the rating scales more consistent across the survey.

Regarding question wording, I deliberately avoided describing any of the words as slurs, in order to avoid priming more adverse judgements. I took care to describe word properties in a way that would be accessible, without losing focus on what they represented. For example, the arousal question asked respondents how shocking they considered a word to be. I acknowledge that this is not a perfect translation of what arousal-- as an unconscious emotional reaction--represents, but considered it the best word choice for the purpose of making a conscious, self-reported judgement of the same. I was also careful to clearly distinguish properties that might be easily confused, such as offensiveness and tabooness. Here, respondents were simply asked how

offensive they personally considered a word to be, and how frequently they would expect a word to be met with disapproval, respectively.

The second set of survey questions asked respondents to rate the same slurs for measures of *reclamation behaviour*--a term I coin to describe various ways in which linguistic reclamation can be engaged with and experienced. Because no existing taboo language studies have included these, I needed to create my own measures. In approaching this task, I was strongly guided by the literature on linguistic reclamation. Particularly, I was mindful of Brontsema's (2004: 5-12) observations that within communities targeted by a particular slur, reclamation of that slur may be practised by some members and not by others, but all members will have an awareness of the process taking place. I was also informed by an understanding that some slurs can be reclaimed in reference to both oneself and others, as in the case of the word *queer* (Baker 2008; Stephens 2011). As such, I decided to gather ratings for four different contexts of reclaimed use: positive uses of slurs by individuals in reference to themselves (self-self reclamation), by individuals in reference to others (self-other reclamation), by other individuals in reference to themselves (other-self reclamation) and by other individuals in reference to others (other-other reclamation). To help ensure that respondents were aware of the different contexts of use being specified for the reclamation questions, bold text was used to highlight the relevant speaker and referent, as well as to make clear that respondents were only being asked about positive uses of the words. I argue that approaching the measurement of reclamation in this way also enables me to ask questions about whether certain contexts of reclaimed use are more commonly experienced than others, and which may be better predictors of performance in lab experiments.

Some word property questions required specific adjustments to be made and communicated to respondents. In general, respondents were asked to leave a word unrated if they had never heard of it before, so that they would not give misrepresentative data. However, this was not the case for the familiarity question in particular, as instances of total unfamiliarity were relevant to the word property. For just the self-self reclamation question, respondents were provided with an additional response option to declare that they did not consider that a word applied to them at all. I reasoned that if respondents were reporting never using a slur to describe themselves positively--as the question asked--because it simply could not refer to them, this was different to never reclaiming a slur because a respondent chose never to do so. I also wanted to enable respondents to indicate whether they felt a word applied to them or not, rather than assuming this.

All but one of the properties/behaviours were rated on a 1-5 Likert scale. As an example, a tabooness rating of 1 meant a slur was not considered taboo at all, but a rating of 5 meant it was considered extremely taboo. The only exception to this was AoA, where respondents gave their answer as one of ten age brackets. These were later coded, such that a value of 1 represented the lowest 0-5 age bracket, and a value of 10 represented the highest 70+ bracket.

3.3. Procedure

The survey was conducted online. Respondents first read an information sheet, detailing the aims of the research and the completion process. They then gave informed consent to participate. Next, they provided demographic information, before responding to each slur rating question. Following completion, eligible respondents could enter the prize draw, before saving their results and closing the survey.

3.4. Data Analysis

3.4.1. Producing the Database

Each of the 41 LGBTQ+ slurs in the study received a group of ratings for each of the word properties and reclamation behaviours. Overall, the survey generated a combined total of 34,949 ratings.

Using this data, I calculated the mean rating for every property and behaviour, for every slur. I then computed three additional reclamation behaviour ratings: self-reclamation (mean self-self reclamation + mean self-other reclamation / 2), other-reclamation (mean other-self reclamation + mean other-other reclamation / 2) and overall reclamation (mean self-self reclamation + mean self-other reclamation + mean other-self reclamation + mean other-other reclamation / 4). I produced these combined ratings so that I could make broader comparisons between self-referential and other-referential reclamation behaviours, as well as to an 'overall' measure of reclamation behaviours.

In total, 574 mean ratings were generated (41 LGBTQ+ slurs x 14 word properties). These were then compiled into a complete database, which can be downloaded here:
https://www.editorialmanager.com/pragcog/download.aspx?id=4243&guid=374d804d-57db-45f5-9957-06c7e75c8833&scheme=1.

3.4.2. Correlation Analyses

Because all of the ratings had been collected on ordinal scales, I chose Spearman's rank-order correlation analysis (Field 2013: 276; Spearman 1904) as the best statistical procedure for identifying relationships between word properties and reclamation behaviours across the database. Because each property/behaviour had a potential correlation with 13 others, this meant that 182 Spearman's correlation analyses were run in total. Because half of these were not unique (e.g. familiarity x personal use was the same as personal use x familiarity), 91 potential correlations were explored overall.

For all statistically significant correlations, I carried out two further validity checks. First, I produced a scatterplot of the relevant data in SPSS, in order to confirm by visual inspection that the relationship was monotonic (i.e. if there was a positive or negative correlation). Second, I calculated upper and lower confidence intervals for the value of the r_s coefficient, to check that the range between the two did not pass through zero, thereby ensuring there was no uncertainty as to the value of the coefficient being positive or negative. In all cases, these procedures showed that all the relationships were monotonic, and that none of the ranges between upper and lower confidence intervals passed through zero.

Though not an indication of causality, I also determined how close the relationship between two particular properties/behaviours was, by calculating as a percentage how much of the variance was directly shared between those which correlated significantly. This is calculated by calculating the r_s^2 value--the square of the value of the r_s coefficient (Field 2013: 276).

3.4.3. Demographic Group Analyses

I selected the Mann-Whitney U test--a nonparametric equivalent to the Analysis of Variance (ANOVA) test--to identify whether there were group differences between word property and reclamation behaviour ratings according to respondent demographic data. As with the correlation analyses, I chose this test because all of the ratings were measured on an ordinal scale (Lehmann 2006; Mann & Whitney 1947).

Each property/behaviour rating for each LGBTQ+ slur was the dependent variable of its own Mann-Whitney U test. Three sets of these were conducted, in which the independent variable was either respondent age group, gender identity or sexual identity. Although I collected demographic to this end, I could not make a statistically valid comparison between cisgender and non-cisgender respondents (see Table 1). Overall, this meant that 1,722 Mann-Whitney U tests were conducted (41 slurs x 14 property/behaviour ratings per slur x 3 independent variables).

Within each independent variable, respondents were further divided into two groups. Groups were created from the demographic data presented in Table 1, adjusted so that there would be enough respondents within each to make valid comparisons. Table 3 lists the groups created, and how many respondents belonged to each. I only compared women and men for the gender identity comparison, as there were not enough non-binary respondents in the sample to constitute their own group (see Table 1). Similarly, I was not able to make balanced comparisons between cisgender and non-cisgender participants, nor between specific non-heterosexual sexual identities. I acknowledge that this left me unable to explore the specific judgements of non-binary and transgender participants in my analyses, and unable to distinguish a bisexual person's judgements and reclamations of LGBTQ+ slurs from those of a gay person, for example. I argue that the comparisons made are still valuable, but recommend future norming studies control their sample for even closer comparisons.

An important preliminary analysis when running a Mann-Whitney U test is to check whether distributions of the two groups of the independent variable have a similar shape. Similarly-shaped distributions enable precise comparisons of the median values in both groups, while dissimilarly-shaped distributions only allow for comparison of mean ranks, which can solely indicate whether values were generally higher or lower in one group compared to the other (Hart 2001). I visually inspected distributions for all of the Mann-Whitney U tests I conducted, and found that none were similarly shaped. This was unsurprising, because the range of ratings that could be given for each word property was small. As such, all statistically significant findings were based on comparisons of mean ranks between groups, rather than medians.

To clarify what mean rank values represent, a higher value indicates that ordinal ratings in that group were generally higher than in the other. For example, in Tables 4 to 7, I make some age group comparisons for particular LGBTQ+ slurs. The values in the mean rank columns do not denote a specific age nor rating, but represent whether ratings skewed comparatively higher or lower in that age group. For ordinal data such as mine, this information can be more useful than the descriptive medians, especially for shorter rank scales; the medians for both groups may be identical if participants only have five options to choose from, but both groups

may still have significantly differed in providing generally higher or lower ratings, which only mean rank values are able to indicate.

Table 3: Adjusted Demographic Groups and Distributions

Age group	*N*	*%*	Gender Identity	*N*	*%*	Sexual Identity	*N*	*%*
16-29	96	61.94	Woman	82	55.41	Non-Heterosexual	79	53.38
30+	59	38.06	Man	66	44.59	Heterosexual	69	46.62
Total	155	100.00	Total	148	100.00	Total	148	100.00

4. Results

4.1. Correlations between Word Properties and Reclamation Behaviours

The Spearman's correlation analyses revealed that the ratings for every word property and reclamation behaviour had a statistically significant correlation with those of every other property/behaviour. Furthermore, two distinct groups emerged, in which properties/behaviours in the same group shared positive correlations with each other, but negative correlations with those in the other group. Figure 1 illustrates these findings. Because of the number of correlations observed, here I have only represented whether the correlation was positive or negative and the degree of statistical significance.

Eighty-four (92.31%) of the 91 unique correlations found were statistically significant at the $p < .01$ level. Seven (7.69%) reached statistical significance at just the $p < .05$ level: the 3 negative correlations of familiarity with arousal, negativity and offensiveness, and the 4 positive correlations of AoA with arousal, negativity, offensiveness and tabooness.

Regarding the groups these results indicate: familiarity, personal use and all measures of reclamation behaviour correlated positively with each other, with a high level of significance. Similarly, arousal, negativity, offensiveness, tabooness and AoA all correlated positively with each other, generally also with a high level of significance. Where a property/behaviour from one group correlated with a property/behaviour from another (e.g. familiarity x tabooness), this correlation was always negative, typically reaching a high level of significance as well.

The results of the r_s^2 analysis, showing how much variance was shared between properties/behaviours which correlated significantly, are presented in Figure 2.

Considering these results alongside those presented in Figure 1, I concluded that word properties tended to share more variance when they also shared a positive correlation. The only exception to this seems to have been AoA, but this may be explained by the fact that correlations between AoA and other properties were generally weaker in terms of their significance (see Figure 2). The relationships between familiarity and arousal, negativity, offensiveness and tabooness were also quite weak, all sharing < 20% of their variance with each other. Although it is unsurprising that self-reclamation, other-reclamation and overall reclamation shared >90% of their variance with the word properties they were each calculated from, it is worth noting that all reclamation behaviours shared >80% of their variance with each other, regardless of who was using the slur and in reference to whom. Overall, these results emphasise the strength of the relationships found between familiarity, personal use and reclamation behaviours, and between arousal, negativity, offensiveness and tabooness. Interestingly, the r_s^2 values in latter group were particularly high.

Figure 1: Correlations between Word Properties and Reclamation Behaviours

	Familiarity	Personal Use	Arousal	Negativity	Offensiveness	Tabooness	Age of Acquisition	Self-Self Reclamation	Self-Other Reclamation	Other-Self Reclamation	Other-Other Reclamation	Self-Reclamation	Other-Reclamation	Overall Reclamation
Familiarity		++	-	-	-	- -	- -	++	++	++	++	++	++	++
Personal Use	++		- -	- -	- -	- -	- -	++	++	++	++	++	++	++
Arousal	-	- -		++	++	++	+	- -	- -	- -	- -	- -	- -	- -
Negativity	-	- -	++		++	++	+	- -	- -	- -	- -	- -	- -	- -
Offensiveness	-	- -	++	++		++	+	- -	- -	- -	- -	- -	- -	- -
Tabooness	- -	- -	++	++	++		+	- -	- -	- -	- -	- -	- -	- -
Age of Acquisition	- -	- -	+	+	+	+		- -	- -	- -	- -	- -	- -	- -
Self-Self Reclamation	++	++	- -	- -	- -	- -	- -		++	++	++	++	++	++
Self-Other Reclamation	++	++	- -	- -	- -	- -	- -	++		++	++	++	++	++
Other-Self Reclamation	++	++	- -	- -	- -	- -	- -	++	++		++	++	++	++
Other-Other Reclamation	++	++	- -	- -	- -	- -	- -	++	++	++		++	++	++
Self-Reclamation	++	++	- -	- -	- -	- -	- -	++	++	++	++		++	++
Other-Reclamation	++	++	- -	- -	- -	- -	- -	++	++	++	++	++		++
Overall Reclamation	++	++	- -	- -	- -	- -	- -	++	++	++	++	++	++	

+ = positive correlation, *p* < .05; + + = positive correlation, *p* < .01; - = negative correlation, *p* < .05; - - = negative correlation, *p* < .01.

Figure 2: r_s^2 Values for Statistically Significant Correlations

	Familiarity	Personal Use	Arousal	Negativity	Offensiveness	Tabooness	Age of Acquisition	Self-Self Reclamation	Self-Other Reclamation	Other-Self Reclamation	Other-Other Reclamation	Self-Reclamation	Other-Reclamation	Overall Reclamation
Familiarity		59%	14%	14%	14%	17%	58%	49%	67%	64%	66%	69%	66%	67%
Personal Use	59%		31%	36%	34%	30%	28%	58%	83%	72%	76%	81%	74%	77%
Arousal	14%	31%		85%	88%	83%	12%	34%	41%	42%	44%	38%	41%	42%
Negativity	14%	36%	85%		92%	85%	12%	28%	44%	45%	50%	41%	48%	46%
Offensiveness	14%	34%	88%	92%		83%	15%	31%	42%	40%	44%	41%	46%	42%
Tabooness	17%	30%	83%	85%	83%		12%	20%	34%	37%	44%	31%	38%	38%
Age of Acquisition	58%	28%	12%	12%	15%	12%		40%	42%	42%	41%	46%	44%	44%
Self-Self Reclamation	49%	58%	34%	28%	31%	20%	40%		71%	62%	58%	74%	69%	66%
Self-Other Reclamation	67%	83%	41%	44%	42%	34%	42%	71%		88%	86%	98%	92%	90%
Other-Self Reclamation	64%	72%	42%	45%	40%	37%	42%	62%	88%		94%	88%	90%	96%
Other-Other Reclamation	66%	76%	44%	50%	44%	44%	41%	58%	86%	94%		86%	88%	98%
Self-Reclamation	69%	81%	38%	41%	41%	31%	46%	74%	98%	88%	86%		94%	92%
Other-Reclamation	66%	74%	41%	48%	46%	38%	44%	69%	92%	90%	88%	94%		94%
Overall Reclamation	67%	77%	42%	46%	42%	38%	44%	66%	90%	96%	98%	92%	94%	

4.2. Results of Demographic Group Comparisons

In this section of the results, I detail the findings of the Mann-Whitney U tests that compared every word property rating for each LGBTQ+ slurs by age group, gender identity and sexual identity. In total, 378 of the 1,722 (21.95%) demographic comparisons conducted were statistically significant. Because of the volume of statistically significant findings, I have selected the most notable for consideration.

4.2.1. Age Group (16-29 vs. 30+) Comparisons

In total, I found 73 significant differences in ratings as a function of respondent age group. 24 (32.88%) of these were significant differences in AoA ratings. In all but one of these cases, the mean ranks were lower in the 16-29 group than in the 30+ group, indicating that AoA ratings were also generally lower in this group. As a reminder of the discussion in section 3.4.3., the values in the mean rank columns do not indicate the average age at which participants in that group reported acquiring a slur, but a higher value in one group indicates that group generally reported later acquisition of a slur than the other did.

Some of the LGBTQ+ slurs with the starkest age-related AoA differences fitting the pattern described above were also those with some of the highest familiarity ratings, as indicated in Table 4. These results suggest that slurs which are more widely recognised are also being acquired earlier by younger speakers.

Table 4: Age of Acquisition Differences by Age Group for High-Familiarity LGBTQ+ Slurs

Slur	Normed Familiarity Rating	Mean Rank (16-29)	N	Mean Rank (30+)	N	U	z	p
Dyke	3.4	27.69	40	46.73	31	952.50	3.33	< .001
Fag	3.81	31.37	45	48.85	31	1018.50	3.73	< .001
Faggot	3.66	32.10	45	47.79	31	985.50	3.33	= .001
Tranny	3.63	31.40	49	52.30	28	1058.50	4.11	< .001

The presence of *tranny* in this set of results is noteworthy, because none of the other properties/behaviours which differed by age group for this slur correlated with the AoA difference in the manner expected from Figure 1, although they correlated with each other in accordance with this trend in every other respect. Table 5 outlines these results.

Table 5: Other Age Group Differences for Tranny

Word Property	Mean Rank (16-29)	N	Mean Rank (30+)	N	U	z	p
Arousal	44.55	49	29.29	28	414.00	-2.97	< .01
Negativity	44.21	49	28.13	27	381.50	-3.23	= .001
Offensiveness	44.53	48	28.16	28	382.50	-3.37	= .001
Personal Use	34.28	45	41.38	28	752.50	1.97	= .05
Self-Other Reclamation	36.84	49	42.79	28	792.00	2.01	.04

While Figure 1 indicates that arousal, negativity and offensiveness typically correlated positively with AoA, so would therefore be expected to have higher ratings in the group that reported a later AoA for *tranny* (30+), these results show that ratings for these properties were actually higher in the group that reported an earlier AoA (16-29). Similarly, Figure 1 shows that personal use and self-other reclamation typically correlated negatively with AoA, so would therefore be expected to have lower ratings in the group that reported a later AoA for *tranny*, but the results in Table 5 show that ratings for these properties were actually lower in the group that reported an earlier AoA.

AoA was also the only property of *queer* that differed significantly by age group, but not in the same way as AoA differences by age group for any other slur. *Queer* was the only slur where AoA ratings were statistically significantly lower in the 30+ group (N = 28, mean rank = 32.27) than in the 16-29 group (N = 50, mean rank = 43.55), U = 497.50, z = -2.30, p = .02.

Another property showing many significant age group differences was familiarity, accounting for 16 (21.92%) of all age group differences. Twelve (75.00%) of these indicated lower familiarity ratings in the 16-29 group, while 4 (25.00%) indicated higher familiarity ratings in the 16-29 group. Tables 6 and 7 indicate which

slurs were less and more familiar in the 16-29 group than in the 30+ group. Where certain LGBTQ+ slurs had statistically significant age group differences for multiple word property ratings, their relationships to each other tended to be consistent with the general word property rating correlations illustrated in Figure 1. However, there were some cases in which this was not true.

There were some cases where slurs that were reported as being acquired at an earlier age by 16-29 year olds were also reported as being less familiar to the same group, which is inconsistent with the negative correlation between these two properties illustrated in Figure 1. This was the case for *arse-bandit*, *bull-dyke*, *fudge-packer* and *rug-muncher,* which I discuss further in section 5.2.

Table 6: LGBTQ+ Slurs with Lower Familiarity Ratings in 16-29 Group

Slur	Mean Rank (16-29)	N	Mean Rank (30+)	N	U	z	p
Arse-Bandit	30.18	46	52.08	31	1118.50	4.44	< .001
Bull-Dyke	32.10	46	49.24	31	1030.50	3.73	< .001
Bum-Boy	33.38	48	47.29	28	918.00	2.72	< .01
Cock-Sucker	35.86	50	46.00	28	882.00	1.96	= .05
Fairy	33.04	46	46.87	30	941.00	2.79	= .01
Fudge-Packer	32.69	50	56.11	28	1040.50	3.77	< .001
Lezzer	35.56	50	46.54	28	897.00	2.10	.04
Muff-Diver	33.14	50	50.86	28	1018.00	3.54	< .001
Pansy	34.58	46	45.16	31	904.00	2.10	.04
Poof	33.87	46	46.61	31	949.00	2.53	.04
Rug-Muncher	35.33	50	46.95	28	908.50	2.27	= .01
Shirt-Lifter	32.16	49	50.96	28	1021.00	3.95	< .001

Table 7: LGBTQ+ Slurs with Higher Familiarity Ratings in 16-29 Group

Slur	Mean Rank (16-29)	N	Mean Rank (30+)	N	U	z	p
Fag	45.74	46	29.00	31	403.00	-3.37	= .001
Faggot	44.90	46	30.24	31	441.50	-2.93	< .01
Gay (pejorative)	43.24	46	32.71	31	518.00	-2.21	.03
He-She	43.32	50	32.68	28	509.00	-2.08	.04

4.2.2. Gender Identity (Women vs. Men) Comparisons

In total, I found 118 significant differences in ratings as a function of gender identity. In 117 (99.15%) of these results, women provided lower familiarity, personal use and reclamation behaviour ratings, but higher arousal, negativity, offensiveness, tabooness and AoA ratings. Because this pattern was consistent across almost all results, Table 8 provides the results for *bent* as an example.

Table 8: Gender Differences for Bent

Word Property	Mean Rank (Women)	N	Mean Rank (Men)	N	U	z	p
Familiarity	32.43	42	45.09	33	927.00	2.57	= .01
Personal Use	29.74	35	38.66	32	709.00	2.28	.02
Negativity	36.88	32	28.12	32	372.00	-1.96	= .05
Offensiveness	37.73	33	28.12	32	372.00	-2.10	.04
Tabooness	37.65	33	28.20	32	374.50	-2.06	.04
Self-Reclamation	28.74	34	38.56	32	706.00	2.60	= .01
Overall Reclamation	28.82	34	38.47	32	703.00	2.27	.02

I note that this pattern of gender identity differences is consistent with the correlations illustrated in Figure 1. Indeed, in almost all cases where individual slurs exhibited significant gender identity differences in multiple property/behaviour ratings, the relationship between these differences matched the overall correlation pattern (see Table 9 for example). For some slurs explicitly targeting LGBTQ+ women (e.g. *bull-dyke, lesbo, lezzer, rug-muncher*), there were no gender identity differences at all, which I discuss further in section 5.3.

Table 9: Gender Differences for Muff-Diver

Word Property	Mean Rank (Women)	N	Mean Rank (Men)	N	U	z	p
Personal Use	24.07	29	31.48	25	462.00	2.66	= .01
Offensiveness	27.36	22	18.83	23	157.00	-2.26	.02
Tabooness	28.41	22	20.12	25	178.00	-2.19	.03

A final gender identity comparison was the finding that of the 10 results (8.47%) that reached statistical significance at the $p \leq .001$ level, 7 (70%) concerned one of the reclamation behaviours. All were for one of three LGBTQ+ slurs in the dataset: *fag*, *faggot* and *gayboy*. In the cases of all three, the reclamation behaviours were consistently lower among women than among men. This is perhaps unsurprising, given that these slurs are each gendered so as to target men, but the degree of statistical significance is nevertheless striking. Table 10 provides the results for *faggot* as an example.

Table 10: Gender Differences for Faggot, $p \leq .001$.

Word Property	Mean Rank (Women)	N	Mean Rank (Men)	N	U	Z	P
Other-Self Reclamation	30.96	41	43.82	31	862.50	3.34	= .001
Self-Reclamation	30.86	39	42.27	32	824.50	3.34	= .001
Other-Reclamation	31.06	41	43.69	31	858.50	3.21	= .001
Overall Reclamation	31.13	41	44.52	32	896.50	3.25	= .001

4.2.3. Sexual Identity (Non-Heterosexual vs. Heterosexual) Comparisons

In total, I found 187 significant differences in ratings as a function of sexual identity. This meant that the sexual identity group comparisons had the highest number of statistically significant differences in property/behaviour ratings. This was perhaps to be expected, given that all the words being rated were LGBTQ+ slurs.

One hundred and thirty-two (70.59%) of these differences were differences in reclamation behaviour ratings. In every case, the reclamation ratings were higher in the non-heterosexual group, regardless of the type of reclamation behaviour. Of these, 52 of these (39.39%) were significant at the $p \leq .05$ level, 34 (25.76%) were significant at the $p \leq .01$ level and 46 were (34.85%) significant at the $p \leq .001$ level. These results reflect existing thought on linguistic reclamation, in that they suggest the non-heterosexual group is able to engage in the reclamation process to an extent that the heterosexual group is not. This might also explain why the majority of these differences reached high levels of statistical significance, with 80 (61.07%) reaching either $p \leq .01$ or $p \leq .001$ level. Table 11 shows how many sexual identity group differences were found for each type of reclamation behaviour.

Table 11: Numbers of Significant Sexual Identity Group Differences for Reclamation Behaviours

Reclamation Behaviour	Number of Significant Sexual identity Differences
Self-Self Reclamation	2 (1.51%)
Self-Other Reclamation	21 (15.91%)
Other-Self Reclamation	27 (20.45%)
Other-Other Reclamation	14 (10.61%)
Self-Reclamation	23 (17.42%)
Other-Reclamation	22 (16.67%)
Overall Reclamation	23 (17.42%)

The two sexual identity differences in self-self reclamation behaviour were for pejorative *gay* and *homo*. However, I suggest that these results should be discounted from analysis. While the data indicates a small number of people outside of the LGBTQ+ community reported using these words positively in self-reference, by my definition (see section 2) I do not consider this an instance of reclamation. Additionally, in neither case were the number of responses for both sexual identity groups statistically comparable (34 non-heterosexual, 8 heterosexual for pejorative *gay*; 34 non-heterosexual, 5 heterosexual for *homo*). All other reclamation behaviours were relatively similar in this regard.

More generally, sexual identity differences reflected the overall correlation pattern in the dataset. With only one exception, where such differences occurred, heterosexual respondents provided lower familiarity,

personal use and reclamation behaviour ratings, but higher arousal, negativity, offensiveness, tabooness and AoA ratings. In this regard, sexual identity differences were almost completely consistent with the correlation pattern illustrated in Figure 1. To illustrate this, Table 12 provides the results for *gayboy*.

Table 12: Sexual Identity Differences for Gayboy

Word Property	Mean Rank (Non-Heterosexual)	N	Mean Rank (Heterosexual)	N	U	z	p
Familiarity	44.87	42	27.83	32	981.50	3.51	< .001
Personal Use	38.36	40	30.36	29	714.50	2.08	.04
Arousal	28.95	37	39.31	29	368.00	-2.29	.02
Tabooness	29.90	41	40.46	26	365.00	-2.26	.02
Age of Acquisition	30.49	41	39.54	26	389.00	-2.02	.04
Self-Other Reclamation	37.50	42	29.65	26	672.00	2.05	.04
Self-Reclamation	39.20	42	28.46	27	743.50	2.63	= .01

I argue this overall pattern represents a very clear difference between those who are targeted slurs on the basis of their sexual identity and those who are not. I note this was not true of the gender identity differences, which indicated no pattern of results that depended on whether the slur specifically targeted men or women.

The sole exception to this pattern was *tranny*. The sexual identity group differences for this slur are presented in Table 13 and indicate that the heterosexual group reported lower familiarity ratings, but also lower arousal, negativity and offensiveness ratings compared to the non-heterosexual group.

Table 13: Sexual Identity Differences for Tranny

Word Property	Mean Rank (Non-Heterosexual)	N	Mean Rank (Heterosexual)	N	U	z	p
Familiarity	44.15	42	28.77	32	951.50	3.17	< .01
Arousal	41.76	42	31.91	32	851.00	2.01	.04
Negativity	42.02	41	30.56	32	862.00	2.43	.02
Offensiveness	41.07	42	31.48	31	822.00	2.07	.04

5. Discussion

5.1. Patterns of Correlation between Word Properties and Reclamation Behaviours

As discussed in section 2., Madan et al. (2017: 802-804) found that certain groupings of non-emotional (e.g. familiarity and personal use) and emotional (e.g. arousal, valence, offensiveness, tabooness) word properties were best at explaining performance in taboo language processing tasks, as long as tabooness was included in both models. Referring to Figure 1, I suggest my results replicated these findings: familiarity and personal use correlated positively, while arousal, negativity, offensiveness and tabooness all correlated positively with each other. If two properties/behaviours did not share a positive correlation, then they always shared a negative one, suggesting two clearly delineated groups.

My correlation patterns developed existing findings by identifying where both AoA and measures of reclamation behaviour fit into these groups. First, I identified that AoA correlated positively with arousal, negativity, offensiveness and tabooness, but negatively with familiarity, personal use and all reclamation behaviours. Second, I identified that all reclamation behaviours correlated positively with familiarity and personal use, but negatively with arousal, negativity, offensiveness, tabooness and AoA. Overall, this suggested two key groups of properties/behaviours, sharing positive correlations within the group but negative correlations between groups. The first group consisted of familiarity, personal use and all reclamation behaviours. The second group consisted of arousal, negativity, offensiveness, tabooness and AoA.

The results of my r_s^2 analysis (Figure 2) also developed existing findings by identifying which properties/behaviours shared the most variance in their ratings, and therefore may be more closely related. Generally, they suggested that word properties with positive correlations tended to share the most variance, but particularly that arousal, negativity, offensiveness and tabooness all shared > 80% of their variance with each other. The same was true of the variance shared between all of the reclamation behaviours. Of all the properties

in the data, AoA seems to have shared the least amount of variance with the others, regardless of the type of correlation. However, familiarity also shared similarly little variance with arousal, negativity, offensiveness and tabooness.

Some of these relationships seem intuitive. For example, it seems natural that greater personal use of an LGBTQ+ slur would be associated with greater familiarity. Similarly, the relationship of reclamation behaviour to familiarity and personal use makes sense; a reclaimed use of a slur is, after all, an instance of personal use. However, I note that it was not only the reclamation behaviours reflecting respondents' own uses (self-self reclamation, self-other reclamation, self-reclamation) that demonstrated this association; all reclamation behaviours did. For these specific correlations, it did not matter whether respondents were reclaiming the slur themselves or were experiencing others doing the same.

Similarly expected was that slurs rated more offensive and more taboo also tended to be reported more arousing and more negative. I consider this to replicate findings from the literature which suggest both arousal (e.g. Anderson & Phelps 2002; Bayer, Sommer & Schacht 2012; Brown & Kulik 1977; Kensinger & Corkin 2003; Madan et al. 2017) and valence (Bayer et al. 2012; Kensinger & Corkin 2003; Madan et al. 2017) are closely linked to the processing of emotional language, with highly taboo language tending to demonstrate the most exaggerated versions of these effects (Jay, Caldwell-Harris & King 2008; Kensinger & Corkin 2003; LaBar & Phelps 1998; Madan et al. 2017).

One finding which I did not predict was that those who reported first encountering a slur later in their lives generally also reported it being more arousing, negative, offensive and taboo. Considering this alongside the finding that later AoA ratings tended to be given alongside lower familiarity and personal use ratings, it is possible that these results suggest a desensitisation effect of acquiring a word earlier in life.

Because of the demonstrated importance of arousal, valence, offensiveness and tabooness to taboo language processing, the most important finding concerning the impact of reclamation was that all of these properties correlated negatively with every measure of reclamation behaviour in the data. Therefore, respondents' ratings for arousal, negativity, offensiveness and tabooness tended to be lower when their ratings for the various reclamation behaviours were higher. I consider this evidence for the possibility that, by its nature, linguistic reclamation might lessen perceptions of arousal, negativity and tabooness for slurs, and may affect the way they are processed compared to other kinds of taboo language. However, I also acknowledge that the reverse could be true; an increase in reclamation behaviours might be a consequence of (rather than a cause for) a slur starting to be viewed as less arousing, negative, offensive, or taboo in general. This requires experimental study to test.

5.2. Age Group Differences: Age of Acquisition and Familiarity Ratings

The comparison between respondents aged between 16-29 and 30+ revealed that almost a third of age group differences were differences in AoA, and just over a fifth were differences in familiarity. In the case of the AoA differences, the results indicated that in all but one case of age difference, there were earlier ages of acquisition reported in the 16-29 group compared to the 30+ group.

Of all my findings, these are the ones to treat the most critically. It is not surprising to have found a number of AoA differences when age itself is the variable being compared. It also makes sense that AoA ratings might be lower and less variable in the younger group; those in the 16-29 group weren't old enough to be able to select outside of the first three possible answers to the AoA question.

However, I argue that differences in age do not presuppose differences in AoA; if the use of a slur had remained constant intergenerationally, the age at which the slur is first encountered should be the same, regardless of the age of the respondent. This was not the case for many of the LGBTQ+ slurs, suggesting that the use of these words is not constant, and that generally speaking these words are being acquired at an earlier age by younger speakers. I suggest that this is because LGBTQ+ identities have become much more visible in recent decades, so it follows that slurs targeting LGBTQ+ people might have become more widely used (and therefore earlier encountered) as a result.

As indicated in Figures 1 and 2, AoA had one of its strongest correlations with familiarity, which was negative. Table 4 also indicates that the 4 LGBTQ+ slurs with the most statistically significant age group differences for AoA also had some of the highest familiarity ratings in the data. Together, these results suggest that slurs which exhibited AoA differences might also have increased in use over time.

The sole exception to the pattern of AoA differences was with *queer*, for which AoA was the only property that differed significantly by age. In this case, the 30+ group instead reported an earlier age of acquisition than the 16-29 group. One possibility is that *queer* used to have a number of meanings largely

unrelated to its use as an LGBTQ+ slur, such as *strange* or *unusual*, which have since declined (Brontsema 2004: 2; Gaucher, Hunt, & Sinclair 2015: 128). If the word used to have entirely non-taboo meanings--which are less restricted in use than taboo ones and therefore likelier to be acquired earlier--that have become obsolete, this might also explain why the word has started to be encountered for the first time at an older age.

Regarding familiarity, the word property with the second highest number of age differences, these results were useful for determining which slurs might be increasing or declining in use over time. I note from Table 6 that many of the slurs that were more familiar to older respondents and less familiar to younger ones belonged to semantic fields of sexual acts and non-normative gender expressions, e.g. *fudge-packer*, *rug-muncher*, *bull-dyke*, *pansy*. Referring to Table 7, the slurs higher in familiarity for younger respondents are more concerned with possessing an LGBTQ+ identity than they are with displaying behaviours associated with being LGBTQ+; the latter references may therefore be more outdated, so less familiar to younger speakers. Another possibility concerns the fact that many of the slurs in Table 6 are formed as compounds. If gaining familiarity with compound slurs necessitates gaining familiarity with each of their constituent words, this process may take longer, resulting in lower familiarity with compound slurs among younger speakers. This might be especially relevant where one of the constituent words is itself a slur, as in *bull-dyke*. However, if these particular slurs are only ever encountered as compounds, their meaning may be a better explanation of their declining familiarity than their structure.

5.3. Gender Identity Differences and the Gendering of Slurs

Women consistently reported lower familiarity, personal use and reclamation behaviour ratings. Indeed, the majority of the most statistically significant gender differences were differences in reclamation behaviour, although all of the slurs with gendered differences in reclamation specifically targeted LGBTQ+ men (see Table 10). Women also tended to rate LGBTQ+ slurs as more arousing, more negative, more offensive and more taboo (see Tables 8 and 9). Combined, these results reflect the broader relationships between ratings in Figure 1. This was the case for most slurs which explicitly target LGBTQ+ women, but for some, there were no differences between women and men whatsoever. Given the broader pattern, this suggests two possibilities. Either gender identity simply did not affect responses to these particular slurs, or the cause of the general difference in women's responses may not be as present when women are the target of the slur.

As to what this cause may be, it is possible that these results speak to a broader social expectation regarding women and taboo language. A number of well-known studies in language and gender research have identified a particular burden on women to use forms of language which are felt to be more polite and indirect (e.g. Brown 1980; Coates 1987; Fishman 1980; Holmes 1995). There has been much debate over the course of recent decades as to why this is the case, but more contemporary arguments stress that where gendered speech norms do develop, they are often a result of pervasive, patriarchal and heteronormative socialisation from an early age (Cameron 2006). This body of work offers an explanation as to why women may have generally reacted less favourably to LGBTQ+ slurs. However, it does not explain why women no longer exhibited this more negative response for some slurs which explicitly target them. Here, I note that more of the slurs in Table 2 typically target LGBTQ+ men rather than LGBTQ+ women. I also note that my sexual identity comparisons revealed those not targeted by LGBTQ+ slurs provided lower familiarity, personal use and reclamation behaviour ratings, but higher arousal, negativity, offensiveness, tabooness and AoA ratings. This was the same pattern that gender identity differences followed, suggesting that differences in women's ratings may also have been reflecting their judgements of slurs which do not typically target them. This would help explain why, for some of the slurs which did explicitly target women, these gender identity differences disappeared.

5.4. Sexual Identity Differences and Reclamation

As raised in section 5.3., my sexual identity comparisons revealed that compared to non-heterosexual participants, heterosexual participants provided lower familiarity, personal use and reclamation behaviour ratings, but higher arousal, negativity, offensiveness, tabooness and AoA ratings. This pattern again mirrors the overall correlation pattern in Figure 1, and suggests that those not targeted by LGBTQ+ slurs reported less experience with and greater objection to them. This is an extremely interesting finding, as it suggests that the group which reports the least experience of and strongest objections to slurs is the group which is not targeted by them, raising a question about who most reinforces the taboos surrounding marginalised social groups and the language which targets them.

Of the sexual identity differences found in the data, 132 (70.56%) were differences in reclamation behaviours. It is unsurprising that it was always the non-heterosexual group reporting higher reclamation behaviour for LGBTQ+ slurs (see Tables 12 for example), given that the reclamation of slurs is a process which operates specifically within the targeted group (Bianchi 2014: 36; Herbert 2015: 131). This might also explain why the majority of differences between the two groups reached high levels of statistical significance, with 80 (61.07%) reaching either $p \leq .01$ or $p \leq .001$ level.

Unlike the gender differences, sexual identity did appear to have affected responses to sexual identity-based slurs in particular. I suggest this finding indicates a much clearer issue of group membership and hegemonic power, which justifies the emphasis on these concepts in existing literature regarding their relationship to slurs and reclamation (see section 2).

It is interesting that other-self reclamation was the most common sexual identity difference in reclamation behaviour, rather than any of the reclamation behaviours focused on individuals using the slurs themselves (see Table 11). With the relevance of ingroup/outgroup membership in mind, I suggest as a possible explanation that in-group access may itself be required in order to experience other LGBTQ+ people using an LGBTQ+ slur to describe themselves positively.

5.5. Differences in Ratings for the Transphobic Slur *Tranny*

A final set of findings concerned the transphobic slur *tranny*, for which ratings often misaligned with the correlation and demographic patterns described thus far. As illustrated in Table 4, *tranny* had one of the highest average familiarity ratings in the data, and was typically reported as being acquired earlier by respondents in the 16-29 age group, which was consistent with other age group differences. However, none of the other age differences for *tranny* reflected the general correlation pattern in terms of how they related to AoA differences. The pattern in Figure 1 suggests that an earlier AoA should have meant that ratings for personal use and self-reclamation would be higher in the 16-29 group, while ratings for arousal, negativity and offensiveness would be lower in the same. However, Table 5 shows the opposite was true. Despite first encountering the word at an earlier age, younger respondents appeared to object more to its use--the only slur for which this was the case.

A similar breaking of pattern was found in the sexual identity comparisons for *tranny*. Referring to Table 13, while the non-heterosexual group generally reported higher levels of familiarity with the slur, they also gave it higher ratings for arousal, negativity and offensiveness. Again, this is not consistent with the results described in Figure 1 and section 5.3. These results suggest that despite the non-heterosexual group being more familiar with the word, they too objected to it more strongly.

With reference to the demographics of my respondents, it is worth noting at this point that just over 90% of the respondents in this study were cisgender, as shown in Table 1. These results therefore cannot speak to the perceptions of *tranny* among those who the slur targets most. A similar study to this with a statistically valid sample of both cisgender and transgender respondents could and should explore whether membership of the trans community affects these findings. However, I argue that these results suggest two perceptions of transphobic slurs among cisgender people. First, that there may be a greater awareness of and subsequent caution toward explicit transphobia among younger cisgender people. This possibility is supported by recent literature suggesting more unified recognition of and opposition to transphobic discrimination among younger people (e.g. Crissman, Czuhajewski, Moniz, Plegue & Chang 2019) compared to older people (Callahan & Zukowski 2019; Elischberger, Glazier, Hill & Verduzco-Baker 2016; Parent & Silva 2018). Second, these results suggest that among cisgender people, membership of the LGBTQ+ community seems to simultaneously increase awareness of and opposition to slurs like *tranny*.

6. Conclusion

The principal goal of this study was to obtain a set of normed ratings for LGBTQ+ slurs, both for a number of word properties relevant to the processing of taboo language, and for various measures of linguistic reclamation. In this, the study succeeded: it generated a combined total of 34,949 individual ratings, which were collectively used to produce 14 sets of normed ratings for all 41 LGBTQ+ slurs in the study. The resulting database is a novel contribution to the field in two regards: it is the only such database of word norms specifically for LGBTQ+ slurs, and it is the only one to include slurs which also includes normed measures of slur reclamation.

However, this was not the only successful outcome of the study: further analysis of the data produced by the study yielded important findings of its own. I consider the most significant of these to concern the

relationships found between word properties and reclamation behaviours. Every rating type in the study was found to have a statistically significant correlation with every other rating type, and there was a clear pattern to these correlations. Familiarity, personal use and all types of reclamation behaviour all correlated positively with each other. Additionally, arousal, negativity, offensiveness, tabooness and AoA all correlated positively with each other. Importantly, rating types in one group always correlated negatively with rating types in the other group. Although these findings cannot be used to argue causation, I argue they provide an early indication that increased reclamation of slurs might be associated with reduced arousal, negativity and tabooness judgements, which can be tested via experimental study.

My study also identified a number of important demographic differences in responses. Results generally indicated that the age of acquisition of LGBTQ+ slurs was generally earlier in younger (16-29) respondents, which I argue may be attributable to increasing visibility of LGBTQ+ people, and by extension the words used to describe them. Several age group differences were also found for familiarity ratings, providing insight on which LGBTQ+ slurs may be increasing or decreasing in use cross-generationally.

Differences in word property ratings between women and men, as well as between non-heterosexuals and heterosexuals, were found to have aligned with the general word property correlation pattern. Typically, both women and heterosexual respondents provided lower familiarity, personal use and reclamation behaviour ratings, but higher arousal, negativity, offensiveness, tabooness and age of acquisition ratings. Most notably, these gender identity differences typically only manifested when LGBTQ+ women were not the target of the slur, the same was necessarily true of sexual identity comparisons, and the most numerous sexual identity differences concerned reclamation behaviours. Although there are several possible explanations for these findings, I particularly consider that these results reflect ingroup/outgroup differences, as well as different gendered expectations regarding use of taboo language more broadly.

Finally, my results suggested that patterns of demographic differences in responses to transphobic slurs may not be the same as those to slurs targeting other groups within the LGBTQ+ community, at least for cisgender individuals. I suggest that for cisgender LGBTQ+ people, both age group and membership of the LGBTQ+ community affects recognition of and opposition to transphobic language.

7.   References

Allan, Keith & Kate Burridge. 2006. *Forbidden words: Taboo and the censoring of language*. Cambridge: Cambridge University Press.

Anderson, Adam K. & Elizabeth A. Phelps. 2002. Is the human amygdala critical for the subjective experience of emotion? Evidence of intact dispositional affect in patients with amygdala lesions. *Journal of Cognitive Neuroscience* 14(5). 709-720. doi:10.1162/08989290260138618

Baker, Paul. 2008. *Sexed texts: Language, gender and sexuality*. London: Equinox.

Bayer, Mareike, Werner Sommer & Annekathrin Schacht. 2012. P1 and beyond: Functional separation of multiple emotion effects in word recognition. *Psychophysiology* 49. 959-969. doi:10.1111/j.1469-8986.2012.01381.x

Bianchi, Claudia. 2014. Slurs and appropriation: An echoic account. *Journal of Pragmatics* 66. 35-44. doi:10.1016/j.pragma.2014.02.009

Bradley, Margaret M. & Peter J. Lang. 1999. *Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings (Technical Report C-1)*. Gainesville, FA: University of Florida.

Brontsema, Robin. 2004. A queer revolution: Reconceptualizing the debate over linguistic reclamation. *Colorado Research in Linguistics* 17. doi:10.25810/dky3-zq57

Brown, Penelope. 1980. How and why are women more polite: Some evidence from a Mayan community. In Sally McConnell-Ginet, Ruth Borker & Nelly Furman (eds.), *Women and language in literature and society*, 111-136. Santa Barbara, CA: Praeger Publishers.

Brown, Roger & James Kulik. 1977. Flashbulb memories. *Cognition* 5. 73-99. doi:10.1016/0010-0277(77)90018-X

Brysbaert, Marc & Michael J. Cortese. 2011. Do the effects of subjective frequency and age of acquisition survive better word frequency norms? *Quarterly Journal of Experimental Psychology* 64. 545-559. doi:10.1080/17470218.2010.503374

Callahan, Matthew P. & Kyle T. Zukowski. 2019. Reactions to transgender women and men in public restrooms: Correlates and gender differences. *Journal of Homosexuality* 66(1). 117-138. doi:10.1080/00918369.2017.1395661

Cameron, Deborah. 2006. *On language and sexual politics*. London: Routledge.

Chen, Melinda Y-C. 1998. '"I *am* an Animal!": Lexical Reappropriation, Performativity, and Queer'. *Engendering Communication: Proceedings from the Fifth Berkeley Women and Language Conference.* 128-140.

Coleman-Fountain, Edmund. 2014. Lesbian and gay youth and the question of labels. *Sexualities* 17(7). 802-817. doi: 10.1177/1363460714531432

Coates, Jennifer. 1987. Epistemic modality and spoken discourse. *Transactions of the Philological Society.* 110-131.

Crissman, Halley P., Christina Czuhajewski, Michelle H. Moniz, Missy Plegue & Tammy Chang. 2019. Youth perspectives regarding the regulating of bathroom use by transgender individuals. *Journal of Homosexuality.* doi:10.1080/00918369.2019.1618646

Croom, Adam M. 2011. Slurs. *Language Sciences* 33(3). 343-358. doi:10.1016/j.langsci.2010.11.005

Gramsci, Antonio. 1971. *Selections from Prison Notebooks.* Trans. Quintin Hoare, & Geoffrey Nowell-Smith. London: Lawrence and Wishart.

Gramsci, Antonio. 1985. *Selections from Cultural Writings.* Eds. David Forgacs, & Geoffrey Nowell-Smith, trans. William Boelhower. London: Lawrence and Wishart.

Elischberger, Holger B., Jessica J. Glazier, Eric D. Hill & Lynn Verduzco-Baker. 2016. "Boys don't cry" - or do they? Adult attitudes toward and beliefs about transgender youth. *Sex Roles: A Journal of Research* 75(5-6). 197-214. doi:10.1007/s11199-016-0609-y

Field, Andy. 2013. *Discovering statistics using IBM SPSS Statistics*, 4th edn. London: SAGE.

Fishman, Pamela M. 1980. Conversational insecurity. *Language.* 127-132. doi: 10.1016/B978-0-08-024696-3.50024-7

Gaucher, Danielle, Brianna Hunt & Lisa Sinclair. 2015. Can pejorative terms ever lead to positive social consequences? The case of SlutWalk. *Language Sciences* 52. doi:10.1016/j.langsci.2015.03.005

Hart, Anna. 2001. Mann-Whitney test is not just a test of medians: Differences in spread can be important. *BMJ* 323. 391-393. doi:10.1136/bmj.323.7309.391

Herbert, Cassie. 2015. Precarious projects: The performative structure of reclamation. *Language Sciences* 52. 131-138. doi:10.1016/j.langsci.2015.05.002

Holmes, Janet. 1995. *Women, men and politeness.* New York: Longman.

Janschewitz, Kristin. 2008. Taboo, emotionally valenced, and emotionally neutral word norms. *Behavior Research Methods, 40*(4), 1065-1074. doi:10.3758/BRM.40.4.1065

Jay, Timothy., Catherine Caldwell-Harris & Krista King. 2008. Recalling taboo and nontaboo words. *American Journal of Psychology, 121*, 83-103. doi:10.2307/20445445

Jay, Timothy B. 2019. The psychology of expressing and interpreting linguistic taboos. In Keith Allan (ed.), *The Oxford handbook of taboo words and language*, 77-96. Oxford: Oxford University Press.

Juhasz, Barbara J., Melvin J. Yap, Joanna Dicke, Sarah C. Taylor & Margaret M. Gullick. 2011. Tangible words are recognized faster: The grounding of meaning in sensory and perceptual systems. *The Quarterly Journal of Experimental Psychology* 64(9). 1683-1691. doi:10.1080/17470218.2011.605150

Kensinger, Elizabeth A. & Suzanne Corkin. 2003. Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory & Cognition* 31(8). 1169-1180. doi:10.3758/BF03195800

LaBar, Kevin. S., & Elizabeth A. Phelps. 1998. Arousal-mediated memory consolidation: Role of the medial temporal lobe in humans. *Psychological Science* 9(6). 490-493. doi:10.1111/1467-9280.00090

Lehmann, Erich L. 2006. *Nonparametrics: Statistical methods based on ranks.* New York: Springer.

Madan, Chris R., Andrea T. Shafer, Michelle Chan & Anthony Singhal. 2017. Shock and awe: Distinct effects of taboo words on lexical decision and free recall. *Quarterly Journal of Experimental Psychology* 70(4). 793-810. doi:10.1080/17470218.2016.1167925

Mann, Henry. B. & Donald R. Whitney. 1947. On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Sciences* 18. 50-60.

Parent, Mike C. & Kevin Silva. 2018. Critical consciousness moderates the relationship between transphobia and "bathroom bill" voting. *Journal of Counseling Psychology* 65(4). 403-412. doi:10.1037/cou0000270

Russell, Stephen. T., Thomas J. Clarke, & Justin Clary. 2009. Are Teens "Post Gay"? Contemporary Adolescents' Sexual Identity Labels. *Journal of Youth and Adolescence* 38(7). 884-890. doi: 10.1007/s10964-008-9388-2

Spearman, Charles. 1904. The proof and measurement of association between two things. *The American Journal of Psychology* 15(1). 72-101.

Stevens, Hugh. 2011. Normality and queerness in gay fiction. In Hugh Stevens (ed.) *The Cambridge companion to gay and lesbian writing*, 81-96. Cambridge: Cambridge University Press.

White, E. 1980. The Political Vocabulary of Homosexuality. In Lucy Burke, Tony Crowley, & Alan Girvin (Eds.), 2000, *The Routledge Language and Cultural Theory Reader,* 189-196. London: Routledge.