

A woman's tradition? Quantifying gender difference in the Child ballads

Duggan, Lucie

Published in:
Orbis Litterarum

DOI:
10.1111/oli.12400

Publication date:
2023

Document version:
Final published version

Document license:
CC BY

Citation for pulished version (APA):
Duggan, L. (2023). A woman's tradition? Quantifying gender difference in the Child ballads. *Orbis Litterarum*, 78(5), 384-400. <https://doi.org/10.1111/oli.12400>

Go to publication entry in University of Southern Denmark's Research Portal

Terms of use

This work is brought to you by the University of Southern Denmark.
Unless otherwise specified it has been shared according to the terms for self-archiving.
If no other license is stated, these terms apply:

- You may download this work for personal use only.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying this open access version

If you believe that this document breaches copyright please contact us providing details and we will investigate your claim.
Please direct all enquiries to puresupport@bib.sdu.dk

A woman's tradition? Quantifying gender difference in the Child ballads

Lucie Duggan 

Department of Language, Culture, History and Communication, University of Southern Denmark, Odense, Denmark

Correspondence

Lucie Duggan, Department of Language, Culture, History and Communication, University of Southern Denmark, Odense, Denmark.

Email: lucied@sdu.dk

Abstract

While the major ballad-collecting efforts of the eighteenth and nineteenth centuries were largely carried out by men, the academic discourse surrounding ballads gendered traditional balladry—orally transmitted narrative poetry—as a woman's tradition. Subsequent ballad scholarship of the twentieth century perpetuated the antiquarian notion of a specifically female ballad tradition, and yet it has remained unclear whether and how this women's tradition should be considered distinct from, for example, a male ballad tradition. This study suggests that a quantitative approach can be used to investigate the question of difference in women's and men's ballad repertoires in the period spanning the mid-eighteenth century to the close of the nineteenth. Employing Linguistic Inquiry Word Count (LIWC) software to examine the definitive ballad anthology—the Child collection—this study finds that the notion of a woman's tradition goes beyond stereotypical notions of repertoires gendered by genre; women's ballads are instead found to be characterised by distinctive linguistic patterns that convey a strong focus on female subjectivity.

KEYWORDS

Child ballads, Digital Humanities, *English and Scottish Popular Ballads*, gender, LIWC

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Author. *Orbis Litterarum* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

In the 1860s and 1870s, the American ballad scholar Francis James Child issued a series of public appeals for ‘all that is left, or all that can be found, of the genuine ballads of the people’, announcing his intention to complete an exhaustive collection of the English and Scottish ballads, which he intended to encompass every ‘popular ballad’ that he could find, including all known versions of it. In one such appeal, Child wrote:

Something must also still be left in the memory of men, or better, of *women*, who have been the chief preservers of ballad-poetry. May I entreat the aid of gentlewomen in Scotland, or elsewhere, who remember ballads that they have heard repeated by their grandmothers or nurses? May I ask clergymen or schoolmasters, living in sequestered places, to exert themselves to collect what is left among the people?

(Child, 2002, 150)

In his appeal for ‘genuine’ ballads, Child identified women as the most likely sources for the kind of songs that he had in mind. He referred to the well-known singers of eighteenth-century Scottish oral tradition, and he appealed for more such women to come to his aid:

[W]here are the Mrs. Farquhars, the Mrs. Browns, the Mrs. Arnots, the Miss Rutherfords themselves, and the nurses who taught them ballads? Small hope, we acknowledge, of finding such nurses any more, or such foster-children, and yet it cannot be that the diffusion of useful knowledge, the intrusion of railroads, and the general progress of society, have quite driven all the old songs out of country- women's heads—for it will be noted that it is mainly through women everywhere—‘The spinsters and the knitters in the son, And the free maids that weave their thread with bones’—that ballads have been preserved.

(Child, 1994, 33)

Identifying women as the chief preservers of ballad poetry, Child's appeal implicitly referenced a well-established association between women and oral singing traditions. Throughout the long eighteenth century, literary figures and gentlemen antiquarians in Scotland and England (such as Walter Scott, William Motherwell and Robert Jamieson) sought to collect and preserve traditional, orally transmitted ballads before they disappeared from memory, publishing their findings in popular ballad books and anthologies.

As they searched for ballads among singers in remote communities, ballad men (for most song collectors were male) recorded songs from predominantly female sources. While antiquarian scholarship hypothesised the origins of ballad poetry in an ancient (male), bardic culture, by the golden age of balladry in the late eighteenth century, women were considered the custodians of these ancient songs, preserving them in their memories, and transmitting them along maternal lines. Gentlemen antiquarians and ballad enthusiasts nostalgically recalled the ballads sung to them in early childhood by female members of household staff, whose singing traditions were frequently connected to a world of domestic labour, including spinning, sewing, and the care of children.¹

Crucially, women (particularly poorer, unlettered women) were seen as repositories for the songs that antiquarian collectors most valued: oral, and thus separate from what song collectors considered to be the correlative, commercial culture of the printed broadside ballad industry.² Singers such as Mary Barr and Amelia Harris emphasised the purely oral nature of the songs they sang, offering collectors assurances of the ancient pedigree of their songs as well as their separateness from broadside culture (Buchan, 1972, 67). The prevalence of women as recorded ballad sources is also frequently speculated to be related to lower rates of literacy among women than men historically (p. 76; Perry, 2008, 88). Yet the biographical details of individual singers (such as the highly educated Anna Gordon) suggest that literacy and ballad singing are less strongly correlated than has previously

been speculated; after all, as Ruth Perry has written of Scottish song culture: women and men had equal access (Perry, 2008, 87). The antiquarian focus on women as the preservers of oral ballad tradition might be better attributed to the German Romantic notion of the 'essential maternal femininity of orality', as Sigrid Rieuwerts has suggested (Rieuwerts, 2002, 151).

The association between women and oral ballad tradition has proved remarkably persistent. In his landmark study *The Ballad and the Folk*, the ballad scholar David Buchan described the songs of the well-known eighteenth-century Scottish singer Anna Gordon as 'a woman's tradition within the regional tradition' (Buchan, 1972, 76). Yet despite the long-standing associations between women's singing and traditional balladry, there has been little investigation of how a female ballad tradition is distinct from, for example, a male ballad tradition (Brown, 1997, 50). It has been suggested that gendered genre preferences mark the distinction between the ballad repertoires of female and male singers (Buchan, 1972, 76; Symonds, 1997, 40; Brown, 1997, 49–51). Women, Buchan suggested, were 'constitutionally more inclined to the marvellous than the martial'; Matthew Hodgart wrote that 'it is a reasonably safe generalisation that the heroic ballad is a masculine art, the pathetic folk-lyric largely a feminine one' (Buchan, 1972, 76; Hodgart, 1965, 13). Other scholars have observed that the oral ballad medium has traditionally afforded female singers a range of themes that speak directly to women's experiences (including unwanted pregnancy, childbirth and women's lack of control within a patriarchal society; see Wollstadt, 2002; Freedman, 1991; Symonds, 1997, 40; Brown, 1997, 51; Stewart, 1993). Yet as Mary Ellen Brown has observed, there remains a dearth of comparative repertoire studies of the kind that might indicate whether the relationship between men's and women's songs is indeed one of difference (Brown, 1997, 50).

More recently, the historian Katie Barclay has derived useful insights from an analysis of the ballads gathered by the collector William Motherwell in the 1820s by female and male singers, and she concludes that gender clearly influenced the choice of song in the repertoires that Motherwell encountered (Barclay, 2010, 353). Barclay's analysis suggests that it was less the genre of a ballad that influenced the singer's choice than 'the potential for the songs to portray [the singer's] concerns, experiences, or identities' (p. 353). She finds that Motherwell's female singers devoted much greater attention to the experiences of female characters in the songs they sang, and she argues that they consistently moved women from the background to the foreground in their songs. However, Barclay's comparative analysis of male- and female-sung ballads is limited to one song, 'Johnie Scot' (Child 99).³

2 | AIM AND METHOD

A large-scale, corpus-wide study of the question of difference between women's and men's ballads is lacking in the field. This article proposes that a quantitative approach can be used to investigate the question of difference in women's and men's ballad repertoires in the period spanning the mid-eighteenth century to the close of the nineteenth.⁴ Employing computational text analysis and statistical methods, I ask: in what ways were women's ballads and men's ballads distinct? And what is the evidence to validate the hypothesis that women singers consistently privileged female-oriented language in their songs?

My approach is inspired by what Timothy Tangherlini has termed 'the folklore macroscope' a means by which to 'model the complex dynamics of a [folklore] collection taking into account not only the texts but also, as Alan Dundes proposed, the context and texture'.⁵ For Child, the study of ballads necessitated a macroscopic approach that could account for the vastness of his data, as well as the myriad interconnections among—and the patterns across—them. His collection represents an attempt to keep the relationship between the part and the whole constantly in view; individual ballads were always part of a wider context of ballad traditions that stretched beyond the individual singer, collector, the local, the regional and even the national context (as Child's efforts to document foreign-language cognates attest).

The genre-defining Child collection remains an exceptionally rich scholarly resource with which to examine a multitude of ballads, gathered from a diverse array of sources (manuscript, broadside and oral) dating from the

medieval period right up until the time of publication in the late nineteenth century.⁶ The *English and Scottish Popular Ballads* is made up of a patchwork of source materials, carefully sifted through, deliberated over and assembled into five volumes (initially issued in ten parts). The collection encompasses not only the 305 ballad texts but also the units of information that further describe those texts, that is, the metadata that tells us something about the text, such as where it was sourced, by whom, what year, and so on. Inspired by the work of the Danish folklorist Svend Grundtvig and his definitive collection of Danish ballads (*Danmarks gamle Folkeviser*), Child was attentive to the fact that an individual ballad possessed multiple versions. The prevailing editorial practice of collating various texts to reconstruct the 'best' version of a ballad had for too long obscured the dynamic and complex process by which ballads were made distinct. For Child, such practices effaced the mark of tradition; he aimed to include all known variants where possible, allowing readers of his collection insights into to the workings of tradition—and, I argue here, the agency of the *singer*—on a ballad over time.

Each ballad entry in *The English and Scottish Popular Ballads* begins with a number, a title and the various versions of that ballad contained within the collection. An example of the entry for 'The Fair Flower of Northumberland' (9) is provided in [Figure 1](#).

Each version is accorded a letter (capitalised) which indicates the status, or rank, of that version, according to Child's inclusion criteria. The A-text is thus the ballad text that Child deemed the superior version of all extant material.

The following metadata are identified for each ballad within the corpus:

- The identificatory number and letter for each ballad entry.
- Whether the ballad was sourced from print or from manuscript.
- Whether the source of the ballad was identified as a printed (broadside) or a traditional (oral) source.
- If the latter, whether the singer was identified (and if so, whether the singer was male or female).
- Occupation of singer (if provided).
- Information on where (and from whom) the singer learned the song (if provided).
- Region of source material (England, Scotland, Ireland or USA).
- Year that the song was collected (if oral).
- Year of publication (if published in a collection).
- Collector, or collection from which the ballad was sourced.

I identify two distinct sub-corpora within the collection: broadside ballads (sourced from the major seventeenth-century collections such as those of Pepys and Roxburghe) and traditional ballads (sourced from the eighteenth- and nineteenth-century collections of collectors such as Walter Scott, Peter Buchan and William Motherwell). The traditional (or what Child termed *popular*) status of the latter is attributed to their provenance among singers from the middle of the eighteenth century onwards. In many cases, the names of these singers have been provided, in some cases with details of their livelihoods and from whom they learned their songs. Within this sub-corpus of

9

THE FAIR FLOWER OF NORTHUMBERLAND

A. a. Deloney's 'Jack of Newbury,' reprint of 1859, p. 61. b. 'The Ungrateful Knight and the Fair Flower of Northumberland,' Ritson's *Ancient Songs*, 1790, p. 169.

B. a. Kinloch MSS, V, 49. b. 'The Provost's Tochter,' Kinloch's *Ancient Scottish Ballads*, p. 131.

C. 'The Betrayed Lady.' a. Buchan's MSS, II, 166. b. Buchan's *Ballads of the North of Scotland*, II, 208.

D. Motherwell's MS., p. 102.

E. 'The Flower of Northumberland,' Mr Robert White's papers.

FIGURE 1 Extract from Child (1965, 111).

Source: [Gutenberg.org](http://www.gutenberg.org).

traditional ballads, I identify songs contributed by female singers and those contributed by male singers. First, I test the assumption that women's and men's repertoires differ according to gendered genre preferences. I then employ a lexicon-based word frequency measure to ascertain whether there are salient differences in the repertoires of female and male singers and if so, what those differences might be.

3 | GENDER OF SINGER AND GENRE

Where the gender of a performer is identified in Child's corpus, female singers outnumber male singers by quite some degree—307 and 75 cases, respectively. However, male singers also feature among the oral sources within *The English and Scottish Popular Ballads*. While outnumbered by female singers, they represent a sizeable share of the oral sources, with no significant difference in the average length of the ballad texts they provided (compared to the female singers). As noted, past scholarship has tended to assert differing genre preferences in the repertoires of female and male singers (Buchan, 1972, 76; Symonds, 1997, 40; Brown, 1997, 49–51). Comparing the female- and male-sung ballads in the Child collection, I find some evidence for gendered genre preferences in *The English and Scottish Popular Ballads*. 75 ballad versions (of various ballads) can be attributed to a male singer, out of which only 24 are identified as being sung only by male and not by female singers, as shown in Table 1.

TABLE 1 Ballads sung by male singer only.

Ballad id	Volume	Title
19A	I	King Orfeo
24B	I	Bonnie Annie
41C	I	Hind Etin
66A	II	Lord Ingram and Chiel Wyet
158C	III	Hugh Spenser's Feats in France
169C	III	Johnie Armstrong
193A	IV	The Death of Parcy Reed
195B	IV	Lord Maxwell's Last Goodnight
196A	IV	The Fire of Fren draught
197A	IV	James Grant
198A	IV	Bonny John Seton
223A	IV	Eppie Morrie
239A	IV	Lord Saltoun and Auchanachie
243D	IV	James Harris (The Daemon Lover)
243F	IV	James Harris (The Daemon Lover)
259A	IV	Lord Thomas Stuart
281A	V	The Keach i' the Creel
281B	V	The Keach i' the Creel
281D	V	The Keach i' the Creel
286B	V	The Sweet Trinity
289C	V	The Mermaid
289F	V	The Mermaid
295B	V	The Brown Girl
302A	V	Young Bearwell

While supernatural ballads as well as romantic courtship ballads feature, this group of ballads has an emphasis on masculine heroism ('Hugh Spenser's Feats in France', 'Johnie Armstrong'), feuding rivals ('Lord Maxwell's Last Goodnight', 'The Fire of Fren draught', 'James Grant') as well as a strong maritime focus ('The Sweet Trinity', 'The Mermaid'). 'The Keach i' the Creel' (with three versions provided by male singers) is a fabliau-esque ballad. Furthermore, many of the ballads in this group are narratives that privilege a predominantly male cast of characters ('The Death of Parcy Reed', 'Bonny John Seton', 'The Mermaid').

However, I find that the majority of Child ballads for which a male and female singer can be identified do not differ substantially in terms of genre, and more than two-thirds of the male-sung ballads are versions of a ballad for which female-sung contributions are also included.⁷ While individual repertoires such as that of Anna Gordon might privilege ballads of a certain genre-specific bent, the male singers documented in *The English and Scottish Popular Ballads* were just as likely to include courtship ballads as their female counterparts, who also knew and sang ballads of warrior heroes and border raiders. In so far as a women's ballad tradition can be identified in the Child corpus, it does not appear to be easily reduced to a preference for genres traditionally associated with women—the marvellous rather than the martial, as David Buchan suggested (Buchan, 1972, 76).

4 | QUANTIFYING BALLAD DIFFERENCE

Men and women do not appear to have had substantial differences in the kinds of songs they included in their repertoires. Yet as the multiple variants of each individual Child ballad attests, there are different ways of singing the same song, as each singer leaves his or her mark on a ballad in the singing of it. Often these changes may not be perceptible qualitatively at the level of the individual ballad text—a word substituted, a changed pronoun—but by adopting a vantage-point at the level of the corpus, I aim to identify whether there are broader gendered patterns of quantitative significance across the Child collection.

I employ a word count software that captures the various linguistic dimensions of a text, known as Linguistic Inquiry Word Count (LIWC); the program enables me to quantify dimensions of difference between women's and men's songs using general semantic categories that are derived independently of the dataset.⁸ The software scans a text and counts the percentage of words that reflect pre-defined word categories. To put it differently, the software produces a coefficient that indicates the number of words from a given category per 100 words of text. The software considers each word of the text by looking for a match with a set of internal dictionaries, which have been created and validated by human judges. As each target word is processed, the dictionary file is searched, looking for a dictionary match with the current target word. If the target word is matched with a dictionary word, the appropriate word category scale for that word is incremented. Word categories include linguistic features (percentage of words in a text that are pronouns, articles, verbs, etc.), psychological (affect, anxiety, anger), social (male, female, family, friends) and more. For example, if LIWC scanned a single ballad that was 400 words and compared them to the built-in LIWC dictionary, it might find that there were 40 pronouns and 16 positive emotion words used. It would convert these numbers to percentages, 10% pronouns and 4% positive emotion words.

A lexicon-based approach to large-scale text analysis offers a relatively effective and efficient means of testing for the presence of linguistic patterns across a large body of text samples. In this case, it provides evidence to support the claim that there are differences in the linguistic make-up of two sub-corpora of ballads and allows us to pinpoint more specifically areas in which we can expect these differences to occur. The LIWC semantic categories provide the means to build a picture of the distinctiveness of women's and men's ballads within the Child corpus by providing a word-measure that is independent of the data analysed. This provides a workable mechanism with which to validate the following hypothesis: there are statistically significant differences between the ballads contributed by female singers and those from male singers within the Child corpus. Furthermore, the LIWC 'female reference words' category—which includes pronouns as well as nouns—allows

us to identify whether the frequency of female-oriented words differs significantly if the singer is female as opposed to male.⁹

Isolating the sub-corpora of the dataset (female-sung and male-sung), the raw data (ballad texts) are analysed with LIWC, and an output dataset with percentages for all LIWC word categories, for each ballad and ballad version, is compiled. The mean averages of the various word categories for the samples are compared, and the significance of the differences that occur between them is assessed using a t-test.

5 | FINDINGS (I): EFFECTS OF SINGER GENDER

Analysis of the linguistic and lexical dimensions that characterise male and female songs in the Child collection uncovers less a picture of difference than of similarity. This is perhaps unsurprising; men's choices in songs have been noted to be broadly consistent with women's (Symonds, 1997, 40; Barclay, 2010, 343). Furthermore, given Child's apparent preference for songs from women singers (as noted in his appeal), we might expect a certain sample bias in the male-sung ballads that he admitted into his collection; that is, perhaps the male-sung ballads in his collection approximated something closer to the qualities of the traditional ballad that Child appeared to perceive more often in women's ballads. In which case, we might expect that the strength of difference identified in the Child collection is a conservative estimation of difference in male- and female-sung (non-Child) ballads in the period. However, despite these considerations, the repertoires of male and female singers differ significantly on a number of linguistic dimensions. Table 2 (see p. 391) demonstrates the language features for which significant differences are observed.

These differences appear most saliently at the level of function words. Most notably, female-sung ballads are defined by a higher use of personal pronouns (*I, we, her*) more generally, and first person singular (*I*) more specifically. Male singers, on the other hand, were much less inclined to emphasise a first-person perspective, showing instead a significantly higher frequency of third person plural (*they*) in their songs. Men's songs were also more verb-heavy and more likely to emphasise conjunctions (*and, but*). An emphasis on time and space orientation also characterised the male-contributed songs, with a particular focus on retrospection (past focus), and spatially focused words (*down, in*) and prepositions (*to, with, above*). In addition, the men's songs in the corpus had a significantly higher frequency of words referring to reward (*take, prize*).

Women's ballads contained higher frequencies of words relating to biological processes (*body, blood, hand, mouth*). In addition, women's songs emphasised social processes and connections, suggesting a greater focus on the sociality of the ballad world and the characters that inhabit it. Importantly, female singers were much more likely than male singers to use words relating to female characters (*girl, her, mother*).¹⁰ The frequency of female-reference words increased by just under one word on average when the singer was female as opposed to male.

Using a linear regression model (Table 3, see p. 391), I examine more closely the estimated relationship between the dependent variable (female reference words) and independent variable (whether the singer is female, as opposed to male) in order to quantify the effect that the gender of a singer has on the content of the ballad she or he sings. If the singer is female, the number of female reference words increases by 0.781 word (per 100 words of text). This is a considerable increase; in a ballad of 500 words, we can thus expect an increase of approximately four additional female reference words, on average. To be sure, this is not an increase that dramatically transforms a ballad beyond recognition; indeed, these are patterns that are not likely to be perceptible qualitatively to the ballad audience. Yet it is a clear pattern that—perceived quantitatively—transforms our understanding of the role that singers (in particular, *women* singers) have had historically in shaping and transforming ballad narratives over time.

The relationship estimated between variables deals with the overall number of ballads contributed by female and male singers within the Child corpus. Digging deeper into the data, I ascertain whether the relationship between gender of singer and female-oriented content is sustained when only those ballads that are versions of the same ballad number are considered; that is, do the numbers change when women and men sing 'the same song'?

TABLE 2 The difference in the frequency of words used between male and female singer.

Feature	Category	Male singer		Female singer		Difference	p-value
		(Mean)	(SD)	(Mean)	(SD)		
Reward	Drive	1.20	0.95	0.84	0.64	0.36	0.00009
Quote	Linguistic	1.60	1.46	0.96	1.29	0.65	0.00018
They	Linguistic	1.30	1.38	0.83	0.83	0.47	0.00019
Relativity	Time	12.55	4.15	11.13	3.17	1.42	0.00123
Negations	Linguistic	1.17	0.97	0.84	0.85	0.33	0.00389
Preposition	Linguistic	10.31	2.46	9.47	2.23	0.84	0.00447
Quantity	Linguistic	1.05	0.87	0.73	0.89	0.32	0.00490
Space	Time	6.49	2.17	5.73	2.10	0.77	0.00510
Past focus	Time	4.69	1.69	4.13	1.71	0.57	0.01022
Verb	Linguistic	13.43	2.77	12.52	2.73	0.91	0.01057
Female	Social	3.39	2.30	4.17	2.38	-0.78	0.01059
Ingest	Biological	0.53	0.66	0.78	0.87	-0.25	0.01904
Pronoun	Linguistic	16.43	2.98	17.34	3.01	-0.90	0.02021
Biological	Biological	2.53	1.59	2.99	1.54	-0.46	0.02228
Conjunction	Linguistic	6.88	2.18	6.26	2.29	0.62	0.03485
I	Linguistic	4.20	2.00	4.84	2.59	-0.64	0.04784
Friend	Social	0.27	0.39	0.40	0.55	-0.13	0.05318

Note: The table shows LIWC word categories that have a significant difference in frequency between male singer ballads ($n=75$) and female singer ballads ($n=307$). The difference in column (5) is measured as follows: Difference = LIWC value for Male Singer minus LIWC value for Female Singer. The degree of statistical significance of this difference is shown in the column (6).

TABLE 3 Regression results for LIWC female reference words and gender of singer.

	(1)	(2)
	LIWC female reference words	
Female singer (vs. Male singer)	0.781** (0.304)	0.701** (0.326)
LIWC dictionary match		0.0838*** (0.0207)
Constant	3.394*** (0.273)	-2.617* (1.584)
Ballad fixed effect		Yes
Observations	382	382
R-squared	0.017	0.074

Note: Linear regression results for a model with LIWC female reference words as dependent variable regressed, in column 1, on a dummy variable equal to one for female-sung ballads, and zero for male-sung ballads ($F[1,380]=6.60$). Column 2 includes a control variable for the quality of dictionary match and includes ballad fixed effects to estimate differences across versions of the same ballad ($F[2,227]=9.0$). 1 p -values are as follows: $\leq 0.01 = ***$, $\leq 0.05 = **$, $\leq 0.1 = *$. Standard errors reported in parentheses.

This is an important consideration since it makes a distinction from the possibility that men and women sang different types of ballads. Using the regression model to control for fixed effects of ballad number, I find that the result holds: whether the singer is female or not affects the content of a ballad, even where that ballad is ostensibly the same ballad narrative as one contributed by a male singer.¹¹

6 | FINDINGS (II): EFFECTS OF INDIVIDUAL SINGERS

Finally, I turn towards an analysis of the individual singer. For the oral sub-corpus, I identify the named individuals whose repertoires are a core part of the traditional ballads of *The English and Scottish Popular Ballads*. I provide a regression model which estimates LIWC female reference words as a function of a combination of individual singers (the baseline category is any other ballad where the singer is not known). Table 4 below presents an overview of the repertoires of individual named singers and illustrates the frequency in female reference words that can be expected depending on the identity of that singer.¹²

Two singers emerge as highly significant in their focus on female-oriented language: Anna Gordon and Agnes Laird. As can be seen in column 3, for a ballad contributed by Gordon, we can expect 1.43 more female reference words (per 100) than for any of the other versions where the singer is unknown. The following calculation serves to illustrate the significance: The constant of 3.5 indicates the average number of female reference words that characterise an average ballad with an unknown singer. If the ballad is sung by Anna Gordon this number increases by 1.43. Therefore, the expected number of female reference words in Gordon's ballads is equal to about 5 per 100 words. For Agnes Laird, that figure increases by 2.1 compared to ballads with unknown singers and is expected to be equal to about 5.7 female reference words per 100 words.

That the former of these two singers should demonstrate a higher frequency of female-centric ballad content in her repertoire is perhaps unsurprising: Anna Gordon's songs have long been considered definitive of a women's ballad tradition.¹³ Agnes Laird, however, has received comparatively little attention for her contribution to the ballad canon. One of the singers referred to by Motherwell as 'old singing women', Laird contributed 12 songs to Motherwell during the course of his song collecting in the parish of Kilbarchan in the summer of 1825. Six of Laird's songs were subsequently included in Child's corpus.¹⁴ In a comparative analysis of Agnes Lyle's (another Kilbarchan singer) and Agnes Laird's versions of 'The Cruel Mother', Flemming Andersen noted distinct differences in the progression of the women's respective ballad narratives, noting how the singers employed the same traditional ballad narrative tools to produce 'very different renditions, each with their own individual focus (which, we may speculate, may in some way reflect the two singers' different tastes, personalities, and world

TABLE 4 Regression results for LIWC female reference words by singer.

	(1)	(2)	(3)
LIWC female reference words			
Anna Gordon	1.362*** (0.423)	1.407*** (0.423)	1.435*** (0.423)
James Nicol		0.626 (0.543)	0.654 (0.542)
Agnes Lyle		0.988 (0.620)	1.017 (0.620)
Mary Barr		0.597 (0.715)	0.626 (0.714)
Mrs Harris		0.691 (0.746)	0.720 (0.746)
Mrs Thomson		0.476 (0.782)	0.504 (0.782)
Miss Jane Webster			0.846 (0.873)
Agnes Laird			2.143** (1.007)
Jenny Watson			0.162 (1.103)
Mrs Storie			1.392 (1.103)
Widow Mc Cormick			0.340 (1.103)
Constant	3.569*** (0.0749)	3.524*** (0.0774)	3.496*** (0.0784)
Observations	1,117	1,117	1,117
R-squared	0.009	0.014	0.020

Note: Linear regression results for a model with LIWC female reference words as dependent variable regressed on different sets of dummy variables indicating the singer. Column 1: $F(1, 1,115)=10.37$. Column 2: $F(6, 1,110)=2.65$. Column 3: $F(11, 1,105)=2.09$. p -values are as follows: $\leq 0.01=***$, $\leq 0.05=**$, $\leq 0.1=*$. Standard errors reported in parentheses.

views)' (Andersen, 2014). Andersen finds, for example, that Laird's version is characterised by an emphasis on the psychological development of the 'cruel mother' of the ballad title.

Barclay goes further in identifying a specific focus on female characters in Laird's ballads, though her analysis is also limited to one song: 'Johnie Scot' (99F). The ballad narrates the actions of its eponymous Scottish hero, who, after a period of employment at the court of the King of England, leaves the King's daughter pregnant; she is locked in a tower but manages to send word to Johnie Scot (sometimes named McNaughton), who eventually raises an armed force to come to her rescue. Barclay notes that Laird's version was typical of those contributed by other female singers, devoting much greater attention to the female experience within what was otherwise a narrative of masculine heroism. Barclay observes that Laird placed an increased emphasis on the pregnancy of the female character within the ballad, quantifying this emphasis: 'Laird mentioned the woman's pregnancy in eight separate verses of a ballad consisting of twenty-two. The smallest number of references to the woman's pregnant condition by a female singer was five' (Barclay, 2010, 344). Moving female characters from background to foreground, Barclay argues, 'Laird claimed space for women in a genre that was dominated by male experience' (p. 344).

The following extract from the ballad identified as 'Johnie Scot' is outlined below in order to provide a comparative demonstration of how a computer-driven, dictionary-based approach can observe patterns both similar to and distinct from those found through a close reading of a single text. 'Johnie Scot' was one of the better-represented ballads in oral tradition, with 16 recorded versions in total, six of them from singers identified by name. Five of those singers were sources of William Motherwell in the mid-1820s. The ballad version 99E was provided by the singer Thomas Risk, from Paisley, Renfrewshire. LIWC identifies 1.45% of the ballad text as female reference words. The ballad version 99F was contributed by Agnes Laird and has a considerably higher proportion of the words (4.59%) identified as female reference words. The first five verses of the ballads are given below (with female reference words highlighted in bold).

'Johnie Scot' 99E (stanza 1–5)	'Johnie Scot' 99F (stanza 1–5)
Motherwell's MS, from the recitation of T. Risk	Motherwell's MS, from the recitation of Agnes Laird, Kilbarchan, 21 June 1825.
McNaughton's unto England gane, The king's banner to bear: 'O do you see yon castle, boy? It's walled round about; There you will spy a fair ladye, In the window looking out.'	Word has to the kitchen gane, And word is to the ha, And word has to the king himsell, In the chamber where he sat, That his ae daughter gaes wi bairn To bonnie Johnie Scot.
'Here is a silken sark, fair lady , Thine own hand sewed the sleeve, And thou must go to yon green wood, To Johnnie thy true-love.'	Word has to the kitchen gane, And word has to the ha, And word has to the queen hersell, In the chamber where she sat, That her ae dochter gaes wi bairn To bonnie Johnie Scot.
'The castle it is high, my boy, And walled round about; My feet are in the fetters strong, And how can I get out?'	'O if she be wi bairn,' he says, 'As I trew well she be, We'll put her in a prison strang, And try her verity.'
'My garters o the gude black iron, And they are very cold; My breast plate's of the sturdy steel, Instead of beaten gold.'	'O if she be wi bairn,' she says, 'As I trew weel she be, We'll put her in a dungeon dark, And hunger her till she die.'
'But had I paper, pen and ink, And candle at my command, It's I would write a lang letter To John in fair Scotland.'	Now she has written a letter, And sealed it with her hand, And sent it unto Johnie Scot, To come at her command.

The text in bold demonstrates the limitations of the program where non-standardised spelling, dialect or archaisms are concerned; in the E-text, 'ladye' in the first stanza is not identified (though 'lady' in the second stanza is) and 'daughter' of the opening stanza of the F-text is identified (though the 'dochter' of the second stanza is not). The extract also illustrates the kinds of references that are missed by the program but which a close reading such as Barclay's is attentive to, such as the mention of the female character's pregnancy; the dictionary does not, of course, recognise the formulation 'gaes wi bairn' that describes her state of being-with-child. And yet, we can see how the increased emphasis on the female protagonist is captured in other ways, even by means of the relatively basic measure of gendered pronouns and nouns. Laird repeats the burden 'oh if she be wi bairn' across two stanzas, so that the she/her pronoun is sustained, appearing at least once in every line. The highlighted text thus demonstrates how a dictionary-based analysis is able to capture an accurate picture of where the singers are focusing their attention in the first five stanzas of the ballad alone.

A close reading undoubtedly enriches this picture further, revealing an almost exclusive focus on the daughter's pregnancy as the plot-driver in Laird's F-text, information that is not even once alluded to in Risk's version. However, my results advance the claim that Laird's repertoire (extended to encompass the entirety of her offerings to the Child corpus) privileges female experience. Like Anna Gordon, whose repertoire is considered to epitomise a female ballad tradition, Laird emerges as a key figure in a consistent pattern that sees female singers as drivers of female-oriented content in their ballad offerings.

7 | SUMMARY OF FINDINGS

The ballad worlds described by the women and men who sang them into being are not the same. Attending quantitatively to semantic categories that pattern differently in women's and men's ballads allows us to build a picture of how these ballad worlds are distinct. Female singers depicted a ballad landscape articulated by connections between individuals who tend to identify as 'I' and 'you' rather than 'we' and 'they'. The imagined community that features in women's songs is profoundly social and is characterised by the proximity of—and communication between—its members. Furthermore, the social world of female-sung ballads is one in which the corporeality of its characters is keenly felt. The findings of this study—based on a corpus-wide quantitative analysis—add weight to the claims of previous scholarship that ballads—with their emphases on themes of out-of-wedlock pregnancy, childbirth and sexual aggression—function as a vehicle for the expression of women's perspectives (Freedman, 1991; Stewart, 1993; Wollstadt, 2002).¹⁵ Female singers, I find, were more likely to employ words that emphasise the biological processes of characters that negotiate their surroundings by way of their mortal frame. The bodies that walk, run and move across the ballad landscape are articulated by arms, limbs and sides—physical extremities that are under constant threat of violation by external forces. The cup that poisons, the dagger that wounds, and the sword that deals a fatal blow are all standard devices in the ballad singer's armoury of storytelling tools. But this study reveals a significant difference in the frequency of words used to describe biological processes in the songs of women and men in the Child corpus, where women's songs demonstrated a heightened focus on the bodies of the ballad world than those of their male counterparts.

Adopting a corpus-wide, macroscopic approach to the question of difference between women's and men's ballads, this study finds patterns of difference that are impossible to quantify by close reading alone. This study shows how the language that female and male singers used to construct their ballad worlds brings certain characters into sharper focus than others. Women, I find, sing more about women. Previous speculation on the gender-specific genre preferences of ballad singers does not hold, since men and women appeared to sing the same songs, on average.¹⁶ It is not the genre or content of a received ballad that matters, but what that recipient (the singer) does with it.

8 | CONCLUSION

Every collection bears the mark of its compiler. The history of ballad scholarship has been propelled by the choices of individual collectors and editors to preserve material that would otherwise have disappeared, but also, to disregard material that did not fit the very often poorly defined criteria around which they structured their collections. Child intended his collection as a body of evidence—a record of the existence of every ‘genuine ballad’ at the time of compilation. Today it is recognised more as an aesthetic canon; the end product of a careful process of elimination by which only those materials considered authentic and attaining a sufficient degree of aesthetic acceptability by Child were included. In *Fakesong*, published in 1985, Dave Harker charged Child with the same crime as that of other ballad ‘mediators’: by rejecting material that did not fit their narrowly defined criteria, collectors and editors (Harker claimed) distorted and suppressed the truth of working-class musical life. Harker has stated:

we can fairly conclude that Child's magnum opus is simply its own definition—a product of a particular stage of academic scholarship. And so long as we are careful not to confuse its underpinning ideology with science, or to take at face value its comments on the cultures from which the ballads came, and through which they passed, we may continue to use the texts as highly mediated examples of the kinds of songs described by late-nineteenth century English-speaking literary scholars as ‘ballads’. About the lives, interests and general culture of the people who made, remade and used these songs, however, a compilation such as Child's can tell us almost nothing.

(Harker, 1981, 162–163)

To be sure, the Child collection (now in its digital iteration) does not tell us anything about the non-Child ballads that exist or that existed at one time—it is far from a transparent window onto the past, and very likely obscures much of the historical realities of women's and men's singing traditions.

Still, if all datasets are imperfect, many are useful. This article finds that the Child collection, shaped though it is by the choices of its editor, also bears witness to the choices of *singers*. It is by studying the collection quantitatively that we gain insights into what these choices were, and how the ballads of women singers in the Child corpus come to be distinct from those of male singers. The theory of oral-formulaic composition espoused by Milman Parry and Albert Lord in the twentieth century has been influential for subsequent ballad scholarship (see Foley, 1985; Lord, 1960). Composition-in-performance by means of a system of ballad formulae or commonplaces—recurring phrases or lines that express a part of the narrative detail—depended on an implicit understanding that the stability of the ballad story existed independently of what Buchan called ‘the stability of text, of the exact words of the story’ (Buchan, 1972, 64).¹⁷ A singer engaged in a process of oral re-creation with each new telling of the ballad story, so that no two versions of any given ballad—even when sung by the same singer—were identical.¹⁸ In the light of Parry and Lord's work on oral composition, ballad scholars have found in the transmission of ballads a highly dynamic and creative process in which the singer shapes the ballad of oral tradition.¹⁹ Many have observed how this process allows for the expression of individual subjectivities, where singers voiced personal values and concerns through their singing (Barclay, 2010; Long, 1973; McCarthy, 1990; Porter & Gower, 1995).

Yet while the methodologies employed to date (close readings of a small selection of ballads) continue to do important and valuable work, they are limited in what they can tell us about broad patterns of difference across a large corpus of ballads based on the independent variable of singer gender. By using a computational word frequency approach based on certain lexical categories, this study finds patterns of difference that are impossible to identify by close reading alone.²⁰ By consistently maximising certain features of any given ballad (including character gender and first-person focus), the singer expressed a poetic agency in her performance. The differences that I observe between male- and female-sung ballads imply a range of choices made in the retelling of the

story-song. These choices subtly altered the content of a ballad narrative to the extent that we can reasonably attribute the frequency of certain language features to the gender of a singer. These gendered language patterns then, sharpen our understanding of how women singers shaped oral ballad tradition over time. My finding that women's ballads are characterised by a higher frequency of first-person pronouns, for example, challenges the prevailing notion that the traditional ballad is devoid of subjectivity. As Ruth Perry writes of the ballads of Anna Gordon, 'There is no psychologizing, no individual subjectivity, no interiority developed for any of the characters and almost no explicit judgment on the part of the narrator' (Perry, 2012, 18). Perry's observations identify what has long been considered standard in the definition of the traditional ballad: its impersonality. Like fairy tales, ballads feature a cast of stock characters—noble lords, treacherous servants, cruel mothers—whose role in the narrative is pre-determined; they act as we would expect them to, time and time again.

However, the specific patterns of pronoun usage that differentiate male- and female-sung ballads respectively in the Child corpus invite further consideration.²¹ Female-sung ballads, I find, are a first-person phenomenon; women singers were more inclined to emphasise the first-person singular in their repertoires than men, which constitutes a significant finding that has a bearing on the extent to which we understand gender and subjectivity in the ballad performance process. The 'I' that speaks in a ballad narrative is notoriously difficult to pin down, and it would be a mistake to conflate the ballad 'I' with the singer. Furthermore, women singers frequently adopted the speaking position of a male character, just as men sang ballads in which female characters narrate. But they appear to have done so in different ways. This study identifies a positive relationship between female singer and first-person perspective. How should this finding be interpreted? The performance context of ballads—an experience shared between singer or reciter and audience—is an important consideration. Ballad characters are brought to life, time and time again, by ballad singers; while the relationship between the two is fluid and shaped by many factors, *identification* with aspects of ballad characters or themes is surely one of them.²² Writing about oral song traditions in an Irish context, the scholar Lillis O'Laoire has noted how subtle variations in an individual singer's version of a song present themselves not as errors but as integral components in the song-as-performance; he observed that performances of female singers in a Gaelic oral tradition were characterised by a practice of personalisation, in which the distance between the singer and the song's female narrator was foreshortened through the adoption of a first-person narrative voice (O'Laoire, 2004, 200–201).²³ While there are clear differences between the context of the Irish oral tradition that O'Laoire describes and the oral narrative ballad tradition recorded in the Child collection, the finding that female singers in the Child corpus used a first-person perspective at a consistently higher rate than male singers suggests the presence of similar gendered practices of personalisation in oral ballad tradition.

This article proposes that a quantitative analysis of a major ballad collection—*The English and Scottish Popular Ballads*—yields findings that are useful and meaningful for our understanding of the socially embedded transmission processes of ballads, and the importance of gender in those processes. Although Child articulated the view that ballads were indicative of a universal, collective voice—'an expression of the mind and heart of the people as an individual, and never of the personality of individual men'—his collection tells, in many ways, a different story (Child, 1995, 59). If women were the chief preservers of ballad poetry, as Child felt they were, then they were also creative innovators who substantially shaped and worked the form of the ballad genre that emerges from the volumes of *The English and Scottish Popular Ballads*. In identifying quantitatively patterns of difference in the songs women and men sang, I formulate the specific ways in which women singers made tradition their own. Dave Harker has claimed that the Child collection 'can tell us almost nothing' about 'the lives, interests and general culture of the people who made, remade and used these songs' (Harker, 1981, 162–163). The quantitative, computational approach to the collection outlined in this article suggests otherwise.

ORCID

Lucie Duggan  <https://orcid.org/0000-0001-8900-3889>

ENDNOTES

- ¹ For example, Oliver Goldsmith recalled being 'sung into tears by our old dairymaid', Sir Walter Scott credited his knowledge of 'Broom blooms bonnie' to his 'old nursery maid', and Robert Louis Stevenson described the old songs sung by his nurse, Alison Cunningham (quoted in Perry, 2006, 289–307).
- ² For Child's notorious pronouncement on broadsides (as 'veritable dunghills'), see Child's letter to Grundtvig (25 August 1872) in Hustvedt (1930, 254).
- ³ Following scholarly convention in ballad studies, the ballads in this article are identified by their Child number and letter (where appropriate) e.g., Child 99A.
- ⁴ This article builds on material presented in the author's unpublished PhD dissertation (Duggan, 2021). I am aware of only one corpus-wide, computational analysis of the Child corpus to date (Preston, 1989, 323–332). While Preston's study demonstrated the potential of computational methods for furthering understanding of stylised ballad language, it did not consider singer gender as a category of analysis. The current study significantly expands the horizons for computational analysis of the Child ballads by making available a dataset that includes all Child ballad texts (including multiple versions) as well as contextual metadata (including singer gender).
- ⁵ Tangherlini borrows the concept of the macroscope developed by Katy Broder, for whom it provides 'a "vision of the whole," helping us "synthesize" the related elements and detect patterns, trends and outliers while granting access to myriad details. Rather than make things larger or smaller, macrosopes let us observe what is at once too great, slow, or complex for the human eye and mind to notice and comprehend' (Tangherlini, 2013, 11).
- ⁶ Among the earliest sources used in the collection are the Sloane Manuscript 2593 (British Museum), circa 1450, from which two ballads are taken: 'St Stephen and Herod' (22) and 'Robyn and Gandeley'n' (115). Among the latest are those ballads recorded in the 1890s by Child's collaborator William Macmath from the recitation of Miss Jane Webster; 'Trooper and Maid' (299D) is one such example. For Child's bibliography, see Child (1965).
- ⁷ Deborah Symonds has observed this in her study of Scottish infanticide ballads, *Weep Not for Me: Women, Ballads, and Infanticide in Early Modern Scotland* (1997, 40). This paper complements her observation by providing a more exact measurement: 51 ballads out of 75 male-sung ballads are also sung by female singers.
- ⁸ For an introduction to LIWC, see <https://www.liwc.app>. For uses of the software in computational text analysis, see Koolen and van Cranenburgh (2017), Piper (2018), Tatlock, Erlin, Pentecost and Knox (2018).
- ⁹ There are, of course, limitations to a lexicon-based word count software and there are important insights to be gained from examining the performance of the LIWC word categories at the level of the text. The LIWC language categories are derived from standard, modern English dictionaries, and therefore do not capture all words in a ballad text drawn from Scots oral tradition (and its concomitant idiolects and sociolects). The limitations of the software are discussed in Pennebaker, Booth, Boyd and Francis (2015), Pennebaker, Boyd, Jordan and Blackburn (2015) and Tausczik & Pennebaker (2010, 24–54). However, there is a relatively high average LIWC dictionary match of 71% for ballad texts classed as traditional. In comparison, the typical dictionary match disclosed by LIWC, when applied in a variety of tests conducted by the authors of the software, ranges between 75% and 86%.
- ¹⁰ Male-sung ballads had a slightly higher mean frequency of male reference words than female-sung ballads, however the difference between the two did not prove statistically significant.
- ¹¹ I provide an additional control for dictionary match, finding that the relatively low dictionary match for the oral ballads is negligible for the accurate measure of frequency of female reference words.
- ¹² I have considered only those singers whose repertoires exceed a total of five ballad texts.
- ¹³ Much has been written about the ballads of Anna Gordon following Child's appraisal of her repertoire as paradigmatic of the ballad tradition. See for example, Bronson (1945), Perry (2013, 297), Pettitt (1984), Rieuwerts (2011).
I have provided two separate controls for the potential skewing effects of Anna Gordon's ballads: (1) using a dummy variable for Anna Gordon ballads, and (2) dropping all Anna Gordon ballads. In both cases the results of the linear regression for LIWC female reference words (see Table 3) remain consistent.
- ¹⁴ Agnes Laird's ballads are: 'The Cruel Mother' (20H), 'Lord Thomas and Annet' (73C), 'Prince Robert' (87C), 'The Gay Goshawk' (96D), 'Johnie Scot' (99F) and 'Willie o Winesberry' (100F).
- ¹⁵ These themes are also discussed in relation to Icelandic women's ballad traditions in Acker (2015).
- ¹⁶ At least where the Child corpus is concerned. My observations do not extend to those ballads that may or may not have been performed in a wider context beyond the Child corpus, therefore I do not make assumptions about an unknown sample of ballads in a pre-1890s context.

- ¹⁷ See Andersen (1985) for an analysis of ballad formulae in *The English and Scottish Popular Ballads*; for a feminist reading of specific ballad formulae, see Martínez García (2017).
- ¹⁸ The differences between versions of the same ballad recorded from Mrs Brown of Falkland are discussed in Buchan (1972) and Pettitt (1984).
- ¹⁹ For a recent discussion of oral tradition and the principles of continuity, variation and selection, see Atkinson (2018, 2–7). ‘Variation can be ascribed’, Atkinson writes, ‘to the agency of individual performers, elevating the ballad singer in some instances to a position more or less akin to that of the literary author or the composer of art music’ (Atkinson, 2014, 98).
- ²⁰ The aim of this study is not to discredit or diminish the importance of close reading. However, the significant limitations of sample bias, the subjective frameworks of value with which we read, the lack of a robust method of generalisation, and the absence of a standard of reproducibility pose problems for the study of ballads that have not been adequately addressed to date. For a discussion of how quantitative methodologies work to address these problems, see Jockers (2013), Piper (2018), Underwood (2019).
- ²¹ My findings take up the concern articulated by the American folklorist Louise Pound in 1921: ‘The English and Scottish ballads are not so wholly impersonal as one is often assured. The ballad “I” may not often refer to the individuality of the author, but the “I” of the singer or reciter is frequently present’ (Pound, 1995, 89).
- ²² Bruce Smith has termed the singing of a ballad ‘an act of impersonation’, where ‘the singer takes on the identity of a distant third person [...] and endows that distant third person with first-person immediacy’ (Smith, 2005, 286). I prefer to use the term *identification* here, as it captures the role of empathy that previous scholarship has identified in song performance. See for example Heisel (2015), Lynch-Thomason (2018).
- ²³ In a comparative analysis of renditions of a song by a well-known Tory Island singer (Séamus Ó Dúgáin) and his mother-in-law (Hannah Shéamais Bháin) respectively in the late 1980s, O’Laire noted perceptible shifts in the pronoun usage of the two singers. Hannah’s version of the love song—in which a woman laments her abandonment by a lover whom she simultaneously praises and castigates—adopted a first-person perspective that was absent from her son-in-law’s rendition of the same song.

REFERENCES

- Acker, P. (2015). Performing gender in the Icelandic ballads. In J. Turco (Ed.), *New Norse studies: Essays on the literature and culture of medieval Scandinavia* (pp. 301–319). Cornell University Library.
- Andersen, F. G. (1985). *Commonplace and creativity: The role of formulaic diction in Anglo-Scottish traditional balladry*. Odense University Press.
- Andersen, F. G. (2014). Voices from Kilbarchan: Two versions of ‘The cruel mother’ from south-west Scotland, 1825. *Oral Tradition*, 29, 47–68. <https://doi.org/10.1353/ort.2014.0000>
- Atkinson, D. (2014). *The Anglo-Scottish ballad and its imaginary contexts*. Open Book Publishers.
- Atkinson, D. (2018). *The ballad and its pasts: Literary histories and the play of memory*. Boydell & Brewer.
- Barclay, K. (2010). Composing the self: Gender, subjectivity and Scottish balladry. *Cultural and Social History*, 7, 337–353. <https://doi.org/10.2752/147800410X12714191853300>
- Bronson, B. H. (1945). Mrs. Brown and the ballad. *California Folklore Quarterly*, 4, 129–140. <https://doi.org/10.2307/1495675>
- Brown, M. E. (1997). Old singing women and the canons of Scottish balladry and song. In G. Douglas & M. Dorothy (Eds.), *A history of Scottish women’s writing* (pp. 44–57). Edinburgh University Press.
- Buchan, D. (1972). *The ballad and the folk*. Routledge & Kegan Paul.
- Child, F. J. (1965). *The English and Scottish popular ballads* (repr. Dover ed.). Dover Publications. Retrieved from Gutenberg.org.
- Child, F. J. (1994). Ballad books. *The Nation*, 1868. In S. Rieuwerts, 33 ‘The genuine ballads of the people’: F. J. Child and the ballad cause. *Journal of Folklore Research*, 31(1/3), 3.
- Child, F. J. (1995). Ballad poetry. In D. Dugaw (Ed.), *The Anglo-American ballad: A folklore casebook* (Vol. 8, pp. 58–67). Garland Publishing.
- Child, F. J. (2002). Old ballads. Prof. Child’s appeal. In S. Rieuwerts, Women as the chief preservers of traditional ballad poetry. *Acta Ethnographica Hungarica*, 47(1–2), 150. <https://doi.org/10.1556/AEthn.47.2002.1-2.16>
- Duggan, L. (2021). *‘I think the distinction easier to feel than to formulate.’: Defining the Child ballad in the digital age*. Syddansk Universitet. Det Humanistiske Fakultet.
- Foley, J. M. (1985). *Oral-formulaic theory and research: An introduction and annotated bibliography*. Garland.
- Freedman, J. F. (1991). With child: Illegitimate pregnancy in Scottish traditional ballads. *Folklore Forum*, 24(1), 3–18.
- Harker, D. (1981). Francis James Child and the ‘Ballad consensus’. *Folk Music Journal*, 4(2), 146–164.

- Heisel, E. (2015). Empathy as a tool for embodiment processes in vocal performance. *Empirical Musicology Review*, 10, 104–110.
- Hodgart, M. (1965). *The Faber book of ballads*. Faber & Faber.
- Hustvedt, S. B. (1930). *Ballad books and ballad men: Raids and rescues in Britain, America, and the the Scandinavian north since 1660*. Harvard University Press.
- Jockers, M. L. (2013). *Macroanalysis: Digital methods & literary history*. University of Illinois Press.
- Koolen, C., & Cranenburgh, A. van (2017). These are not the stereotypes you are looking for: Bias and fairness in authorial gender attribution. In *Proceedings of the First ACL Workshop on Ethics in Natural Language Processing* (pp. 12–22). Association for Computational Linguistics.
- Long, E. R. (1973). Ballad singers, ballad makers, and ballad etiology. *Western Folklore*, 32, 225–236.
- Lord, A. B. (1960). *The singer of tales*. Harvard University Press.
- Lynch-Thomason, S. (2018). 'I've always identified with the women.' *How Appalachian women ballad singers' repertoire choices reflect their gendered concerns*. East Tennessee State University. Electronic Theses and Dissertations. Paper 3488. Retrieved from <https://dc.etsu.edu/etd/3488>
- Martínez García, A. B. (2017). Formulaicity in Child ballads as a means to express assertiveness. *Folklore (London)*, 128, 175–188.
- McCarthy, W. B. (1990). *The ballad matrix: Personality, milieu, and the oral tradition*. Indiana University Press.
- O'Laoire, L. (2004). The right words: Conflict and resolution in an oral Gaelic song text. *Oral Tradition*, 19, 187–213. <https://doi.org/10.1353/ort.2005.0008>
- Pennebaker, J., Booth, R., Boyd, R., & Francis, M. (2015). *Linguistic inquiry and word count: LIWC 2015 Operators' Manual*.
- Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. G. (2015). *The development and psychometric properties of LIWC2015*. The University of Texas at Austin.
- Perry, R. (2006). Brother trouble: Incest ballads of the British Isles. *The Eighteenth Century*, 47, 289–307.
- Perry, R. (2008). 'The finest ballads': Women's oral traditions in eighteenth-century Scotland. *Eighteenth-Century Life*, 32(2), 81–97.
- Perry, R. (2012). The famous ballads of Anna Gordon, Mrs. Brown. In E. Pollak (Ed.), *A cultural history of women in the age of Enlightenment* (pp. 187–208). Michigan State University.
- Perry, R. (2013). The ballad repertoire of Anna Gordon, Mrs Brown of Falkland. *Eighteenth-Century Music*, 10, 297–299. <https://doi.org/10.1017/S1478570613000122>
- Pettitt, T. (1984). Mrs. Brown's 'lass of Roch royal' and the Golden age of Scottish balladry. *Jahrbuch für Volksliedforschung*, 29, 13–31. <https://doi.org/10.2307/849286>
- Piper, A. (2018). *Enumerations: Data and literary study*. The University of Chicago Press.
- Porter, J., & Gower, H. (1995). *Jeannie Robertson: Emergent singer, transformative voice*. University of Tennessee Press.
- Pound, L. (1995). Ballads and the illiterate [1921]. In D. Dugaw (Ed.), *The Anglo-American ballad: A folklore casebook* (Vol. 8, p. 89). Garland Publishing.
- Preston, C. L. (1989). The way stylized language means: Pattern matching in the Child ballads. *Computers and the Humanities*, 23, 323–332.
- Rieuwerts, S. (2002). Women as the chief preservers of traditional ballad poetry. *Acta Ethnographica Hungarica*, 47(1–2), 149–159. <https://doi.org/10.1556/AEthn.47.2002.1-2.16>
- Rieuwerts, S. (Ed.). (2011). *The ballad repertoire of Anna Gordon, Mrs. Brown of Falkland*. Boydell Press.
- Smith, B. R. (2005). Female impersonation in early modern ballads. In P. Brown & P. A. Parolin (Eds.), *Women players in England, 1500–1660: Beyond the all-male stage* (pp. 284–301). Ashgate.
- Stewart, P. (1993). Wishful, wilful, wily women: Lessons for female success in Child ballads. In J. N. Radner (Ed.), *Feminist messages: Coding in women's folklore culture* (pp. 54–75). University of Illinois Press.
- Symonds, D. A. (1997). *Weep not for me: Women, ballads, and infanticide in early modern Scotland*. Pennsylvania State University Press.
- Tangherlini, T. R. (2013). The folklore macroscope: Challenges for a computational folkloristics. *Western Folklore*, 72(1), 7–27.
- Tatlock, L., Erlin, M., Pentecost, S., & Knox, D. (2018). Crossing over: Gendered reading formations at the Muncie Public Library, 1891–1902. *Cultural Analytics*, 3, 1–31.
- Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: Liwc and computerized text analysis methods. *Journal of Language and Social Psychology*, 29, 24–54.
- Underwood, T. (2019). *Distant horizons: Digital evidence and literary change*. The University of Chicago Press.
- Wollstadt, L. (2002). Controlling women: Reading gender in the ballads Scottish women sang. *Western Folklore*, 61, 295–317.

AUTHOR BIOGRAPHY

Lucie Duggan (lucied@sdu.dk) is a post-doctoral researcher and PI of the Augustinus-funded research project *Reading Women. Karen Brahe and Female Book Ownership (1609–1736)* at the Department of Language, Culture, History and Communication at the University of Southern Denmark. She holds a PhD in Literary Studies and has a research background in Digital Humanities, women's literary and cultural history, ballad collections and antiquarian collecting practices. She has been awarded the Leah Leneman Essay Prize (2021) by Women's History Scotland and the SDU PhD Award for the Faculty of Humanities (2021) for her doctoral research on ballads.

How to cite this article: Duggan, L. (2023). A woman's tradition? Quantifying gender difference in the Child ballads. *Orbis Litterarum*, 78, 384–400. <https://doi.org/10.1111/oli.12400>