

CHIARA NARDONE

Università di Bologna/Heinrich-Heine-Universität Düsseldorf

Chiara.nardone2@unibo.it

**Abstract**

*In recent years, interest towards CADS has significantly grown but cross-linguistic analysis still represents a small niche in this research field, as emphasized in Partington et al. 2013. Furthermore, Schafroth (2015) highlighted that, in Romance and German Studies, the combination of Corpus Linguistics and Discourse Analysis has not yet received much attention. This contribution aims at investigating, from a comparative corpus-assisted perspective, how the discourse on 'women and work' is constructed in German and Italian. The analysis has been carried out on two large corpora (itTenTen and deTenTen). First, lexical collocates of sequences containing the combinations 'women, work', 'work, women', 'men, work' and 'work, men' have been collected both in Italian and in German. Then, a sample of concordance lines has been studied in further detail. Results of the collocational analysis show that the discourse on 'women and work' is constructed around different semantic areas in German and Italian (i.e. respectively around the topic of equal pay and that of equal opportunities). On the grounds of these results, it can be argued that in both countries the issue of 'women and work' is still considered 'problematic', even though the size and kind of 'problems' to be solved are different.*

**Key words:** *cross-linguistic corpus assisted discourse studies, women and work, German, Italian*

**1. Introduction and Theoretical Background**

The present research weaves together language, gender and work. This approach stems simply from a mundane observation: almost every day we are presented with data, statistics, reports and various kinds of articles pointing out the stereotypes and discriminations that women face in their working life, as well as the different obstacles that prevent them from joining the labour market. To quote one of these documents, namely the report *Work in the EU: women and men at opposite ends*, published by the EU agency EIGE, which deals with gender related issues:

Gender segregation creates and reinforces gender inequalities in and beyond the world of work. It affects the EU economy by making the labour market less competitive and harder for companies to fill high-demand positions in IT and engineering. Segregation also creates differences in pay, higher risk of poverty and lower economic independence for women (European Institute for Gender Equality 2017).

The aim of this study is to investigate, from a comparative corpus-assisted perspective, how the discourse on ‘women and work’ is constructed in Germany and in Italy through the analysis of two general corpora: deTenTen for German and itTenTen for Italian; both corpora are part of the TenTen corpus family (Jakubíček et al. 2013). Therefore, this study situates itself within the research field of cross-linguistic corpus assisted discourse studies (term used in Partington et al. 2013 and in Taylor 2014), which combine methods of Corpus Linguistics and of Discourse Analysis to investigate multilingual, and often multicultural, data.

The combination of Corpus Linguistics and (Critical) Discourse Analysis has gained popularity in more recent years<sup>1</sup> and the several strengths of this integrated approach have been summarized by Baker et al. (2008: 297):

the combination of methodologies traditionally associated with CDA (DHA [discourse-historical approach] and CL in research projects, and their potential theoretical and methodological cross-pollination, seem to benefit both CDA and CL. Combining methods strengthens the theoretical basis of both DHA and CL. CL, in general, and concordance analysis, in particular, can be positively influenced by exposure and familiarity with CDA analytical techniques, and the theoretical notions and categories of DHA can inform the quantitative CL analysis. Also, CL needs to be supplemented by the close analysis of selected texts using CDA theory and methodology. CDA, in turn, can benefit from incorporating more objective, quantitative CL approaches, as quantification can reveal the degree of generality of, or confidence in, the study findings and conclusions, thus guarding against over- or under-interpretation (O’Halloran and Coffin, 2004).

It is worth mentioning that, within the huge amount of publications that draw on the above-mentioned combined methodology, only a few have extended the analysis to multilingual and multicultural data (see e.g. Freake 2012; Jaworska and Krishnamurthy 2012; Partington et al. 2013; Taylor 2014; Vessey 2013).

It goes without saying that, in comparison with monolingual CADS, researchers do face additional challenges when working with bi- and/or multilingual corpora because the data they have to analyse and interpret are not only produced in different languages, but also, by different cultures and in different national contexts. In addition to providing interesting case studies, some of the publications cited above have presented a few of these methodological challenges. For instance, Taylor 2014 and Vessey 2013 brought to the fore the issue of identifying comparable items in the different languages. As quoted in Taylor (2014: 347):

to translate means (1) to identify a specific function together with its formal realisations in L1, (2) to compare it with another set (function + formal realisation), or other sets, in L2 and finally, in the light of the previous stage, (3) to attempt to encode the given function into a chosen formal realisation in the target language (Tognini-Bonelli & Manca 2004: 371).

Based on this definition, identifying functional equivalence for search terms ‘involves looking at meanings and meanings in context’ while also taking into account the ‘evaluative value of the terms in each language’ (Taylor 2014: 374).

Another relevant challenge for cross-linguistic corpus assisted discourse studies lies in ‘the choice of how to report the findings’ (Taylor 2014: 374) which means having to choose between literal or functional translation into English of the terms analysed when presenting them in an international academic paper.

In relation to this issue, the main feature that differentiates this paper from most cross-linguistic CADS, which investigate data in English and in another language<sup>2</sup>, is that the present study involves not two but three languages: while the analysis has been carried out in German and in Italian, the paper is written in English.

The vast majority of CADS focus either solely on English or, in the case of cross-linguistic CADS, on English plus another language. In the fields of Romance Studies and of German Studies, this combined approach is scarcely adopted. As a matter of fact, among the numerous German Studies publications that focus on Discourse Analysis, only a tiny minority draws on Corpus Linguistics (see e.g. Bubenhofer 2008, 2009, 2013, 2013a; Ziem 2013) and, within this small group, only a niche investigates political and/or social issues (e.g. Bubenhofer 2013a; Ziem 2013).

Given these premises, this study aims at increasing the available knowledge in the CADS area and, more specifically, in cross-linguistic CADS that investigate languages other than English, in this case German and Italian. The urge of expanding the integrated application of Discourse Analysis and Corpus Linguistics within German and Romance Studies, especially from a comparative point of view, has been highlighted, among others, by Schafroth 2015.

One of the key features of CADS, as described in Partington et al. (2013: 206), is to look ‘at other sources of information outside our corpus’. Partington et al. stated that ‘in discourse analysis, we are analysing language in context and therefore to treat the corpus as an isolated black box is often methodologically unsound and unfruitful’. Furthermore, they added that looking at external data could also be ‘a means of identifying areas for analysis from a more objective and replicable starting point’ (Partington et al. 2013: 206). For these reasons, some external data related to the topic of ‘women and work’ in Germany and Italy will be illustrated, namely the Global Gender Gap Index<sup>3</sup> (Hausmann et al. 2016), the women’s unemployment rate<sup>4</sup> and the Gender Pay Gap (Eurofound 2016).

The Global Gender Gap Index was first introduced by the World Economic Forum in 2006 as a framework to assess gender-based disparities in the world. The Index benchmarks national gender gaps on economic, political, education and health criteria<sup>5</sup>. Since the core of this study is ‘women and work’, I will discuss the rankings that indicate Economic participation and Opportunity in the two selected countries. In terms of Economic participation and Opportunity, in 2016, Germany<sup>6</sup> ranked 57<sup>th</sup> and Italy<sup>7</sup> ranked 117<sup>th</sup> out of 144 countries; while in 2006, Germany ranked 32<sup>nd</sup> and Italy ranked 87<sup>th</sup>.

It is worth mentioning that both countries rank in higher positions in the Global Gender Gap Index, which is an average value of all criteria taken into account (i.e. Economic Participation and Opportunity, Educational Attainment, Health and Survival and Political Empowerment). In fact, in 2016 Germany ranked 13<sup>th</sup> and Italy ranked 50<sup>th</sup> out of 144 countries. In addition, it

appears relevant to list the indicators employed to calculate the above-mentioned Economic Participation and Opportunity ranking. Table 1 shows these data in detail.

	<b>Germany (2016)</b>	<b>Italy (2016)</b>
<b>Labour force participation</b>	41	89
<b>Wage equality for similar work (survey)</b>	95	127
<b>Estimated earned income</b>	41	98
<b>Legislators, senior officials, and managers</b>	75	79
<b>Professional and technical workers</b>	1	87

**Table 1.** Economic participation and opportunity – Germany and Italy

Concerning women's employment rate, Eurostat reported that in 2015 73.6% of women in Germany and 50.6% of women in Italy were employed. The average employment rate in the 28 EU member states was 64.3%.

On the other hand, in 2015 the average rate in the 28 EU member states was 75.9%, while in Germany it was 82.3% and in Italy it was 70.6%.

Data show that men's employment rate in Italy was 5.3 points lower than the average rate of the EU-28, whereas it lagged 13.7 points behind the average rate of the EU concerning women's employment rate. On the other hand, the rate of employed women in Germany was 9.3 points higher than the European average.

Since the present analysis on 'women and work' discourse has been carried out on corpora that were collected in 2010 (itTenTen) and in 2013 (deTenTen 2013), the statistical data collected in those years are also reported hereafter (see Table 2)<sup>8</sup>.

	<b>Germany</b>	<b>Italy</b>	<b>EU (28 countries)</b>
<b>Women's employment rate</b>	69.7% (2010)	49.5% (2010)	62.1% (2010)
	72.5% (2013)	49.9% (2013)	62.6% (2013)
	73.6% (2015)	50.6% (2015)	64.3% (2015)
<b>Men's employment rate</b>	80.4% (2010)	72.7% (2010)	75.1% (2010)
	82.1% (2013)	69.7% (2013)	74.3% (2013)
	82.3% (2015)	70.6% (2015)	75.9% (2015)

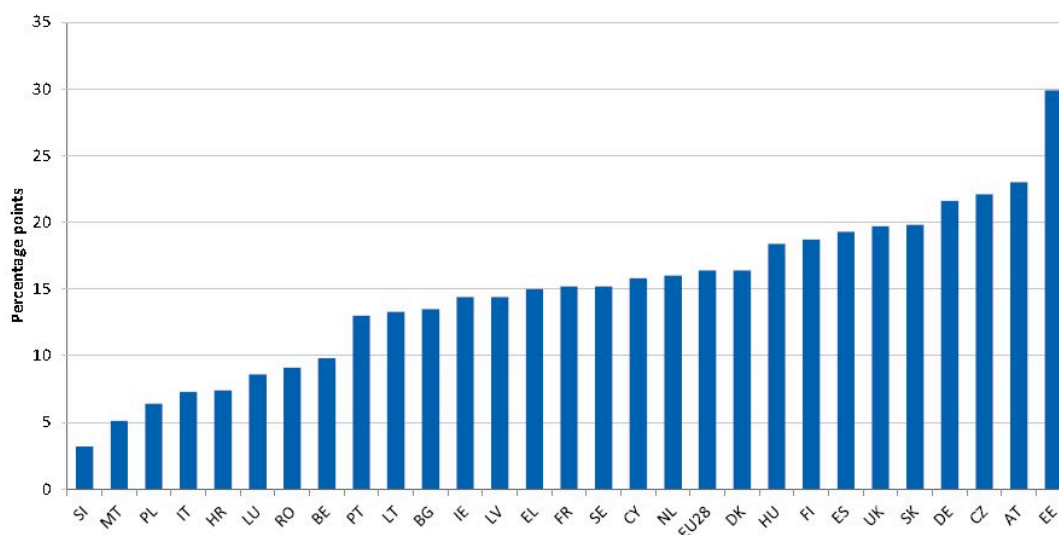
**Table 2.** Women's and men's employment rate

It is indeed interesting to note that, over 5 years (from 2010 to 2015) the gap between men's and women's employment rate has been consistently around 10% in Germany, and around 11-13% in the EU, whereas, in Italy, it has been always higher, around 20-22%.

To conclude the presentation of external data relevant to this research, it is important to mention also the gender pay gap. The gender pay gap is the last one of the external data taken into account within this research. As stated in Eurofound (2016):

In spite of more than 30 years of equal pay legislation, the gender pay gap has remained persistent across all Member States, regardless of the overall level of female employment, national welfare models or equality legislation. A gender-segregated labour market, the difficulty of balancing work and family life, and the undervaluation of women's skills and work are some of the complex causes of the persistent gender pay gap. On average, in 2014, women in the EU earned 16.1% less per hour than men, according to Eurostat. The gender pay gap exists even though women do better at school and university than men.

As Figure 1 shows, Germany has one of the highest gender pay gaps in Europe, while Italy has one of the lowest.



**Figure 1.** Percentage point gap in pay between men and women, EU Member States, 2013 (Eurofound 2016)

As presented above, the Women's employment rate, the Economic Participation and Opportunity index and the Gender Pay Gap show substantial discrepancies between Germany and Italy in these areas. On the grounds of these differences, this paper aims at investigating whether, in the two countries, the discourse on 'women and work' focuses on different issues or it presents similarities and common points despite the differing values of the external indicators.

## 2. Methodology

In order to investigate the discourse on 'women and work' in Italy and in Germany, I carried out an analysis on two general corpora from Sketch Engine

(2018): itTenTen<sup>9</sup> and deTenTen<sup>10</sup>. Both corpora belong to the TenTen generation corpora: they include texts of different genres drawn from the Internet and they are available in the Sketch Engine (Kilgarriff et al. 2014). ItTenTen and deTenTen were created by Web crawling in 2010 and contained respectively 2.5 and 2.3 billion words. In time, both corpora were further developed: the latest version of deTenTen, created in 2013, contains 16.5 billion words, while the 2016 version of itTenTen contains 4.9 billion words<sup>11</sup>.

Investigating two large general corpora rather than smaller specialized ones implies having less control over the text genres included in the analysis. On the other hand, due to the large number of available data, rare linguistic patterns are also likely to emerge and may become the object of analysis, although they may not be representative of general patterns.

The starting point of the analysis was collecting sequences containing the following combinations of words: ‘women, work’ and ‘work, women’ in both corpora. As Baker (2006: 64) pointed out, ‘one way of making sense of a discourse of something is to compare it to a discourse that is not’. Therefore, in order to compare the results in a more solid and reliable way, I collected not only structures containing ‘women, work’ and ‘work, women’, but also the ones containing ‘men, work’ and ‘work, men’, both in German and in Italian<sup>12</sup>.

Rather than looking for single words, I used the Corpus Query Language CQL<sup>13</sup> to retrieve all sentences containing the following combinations of words, divided by a gap from 0 to 5 words:

- *donne* []{0,5} *lavoro* within <s/>
- *lavoro* []{0,5} *donne* within <s/>
- *lavoro* []{0,5} *uomini* within <s/>
- *uomini* []{0,5} *lavoro* within <s/>
- *Frauen* []{0,5} *Arbeit* within <s/>
- *Arbeit* []{0,5} *Frauen* within <s/>
- *Männer* []{0,5} *Arbeit* within <s/>
- *Arbeit* []{0,5} *Männer* within <s/>

The following step of the analysis was to look for the ten most frequent lexical collocates (window span -5, +5)<sup>14</sup> of the above-listed sequences ranked by the log-likelihood score<sup>15</sup>. I used Baker (2006a: 95) as reference for the definition of collocates:

When a word regularly appears near another word, and the relationship is statistically significant in some way, then such co-occurrences are referred to as collocates and the phenomena of certain words frequently occurring next to or near each other is *collocation*.

I decided to include only lexical collocates, thus excluding grammatical or function words such as articles, prepositions, conjunctions and pronouns because, as Baker (2006a: 100) stated,

one of the problems with using a frequency-based technique to calculate collocates is that these high frequency words generally tend to be function

words – which does not always reveal much of interest, particularly in terms of discourse.

The analysis focused on collocates because recurrence is often considered a salient feature in Discourse Analysis. As Stubbs (1996: 158, reported in Vessey 2013: 7), among others, notes:

frequent phrases and discourse patterns are vital to discourse communities: they enable community members to understand connected discourse, attitudes, values and even ideologies within their communities. In other words, considering the function of frequent, and repeated, words and phrases allows a researcher to gain insights into the discourse and culture of specific communities.

Furthermore, Baker (2006a: 114) highlighted the strong connection between lexical collocates and the power of the message they convey:

A collocation analysis [...] gives us the most salient and obvious lexical patterns surrounding a subject, from which a number of discourses can be obtained. When two words frequently collocate, there is evidence that the discourses surrounding them are particularly powerful – the strength of collocation implies that these are two concepts which have been linked in the minds of people and have been used again and again.

A further step of the analysis was to observe a sample of concordance lines of some lexical collocates in more detail. I chose to focus on the words equal (*gleich* for German and *pari* for Italian) and family (*Familie* for German and *famiglia* for Italian) since, as shown in the findings section, these terms occur among the most frequent collocates both in the itTenTen analysis and in the deTenTen one.

In the present study I combined a quantitative and a qualitative approach. As Baker stated (2006a: 71), ‘a concordance is simply a list of the occurrences of a particular search term in a corpus, presented within the context they occur in; usually a few words to the left and right of the search term’. According to him, concordance analysis may be useful to identify linguistic patterns, and ‘identifying such patterns may help us to note discourses, particularly if the patterns are relatively common’ (Baker 2006a: 77). In addition, Baker (2006a: 89) pinpointed that:

a concordance analysis is one of the more qualitative forms of analysis associated with corpus linguistics. While concordance programs allow researchers to sort and therefore view data in a variety of different ways, it is still the responsibility of the analyst to recognize linguistic patterns and also to explain why they exist.

In the following section lexical collocates and some concordance lines will be listed, discussed and analysed in relation with the extra-textual information presented in the previous section, i.e. the Global Gender Gap Index, the women’s employment rate and the Gender Pay Gap.

### 3. Findings

Before presenting the research findings, it is important to clarify a methodological point concerning the choice of how to report such findings.

Partington et al. (2013: 190) and Taylor (2014: 374) affirmed that one of the challenges of cross-linguistic discourse analysis lies in how to report the research findings. Since this paper is written in English and its audience may be international, I decided to present the results in a more accessible way for readers who are unfamiliar with Italian and German and, therefore, I translated the findings into English<sup>16</sup>. During the translation process, it has been difficult sometimes to choose between a literal and functional translation. In order to provide an accurate translation, I carried out an in-depth analysis of the context and of the collocates' concordance lines, as, for instances, in the examples reported below.

IT: *lavoro*; DE: *Arbeit*; EN: work, job;

DE: *Arbeit leisten*, *Arbeit verrichten*; EN: do the job, do work, carry out their work, etc.

Moving on to the results, I will first mention some quantitative aspects, even though this research does not focus on a purely quantitative analysis. It is noteworthy that sequences containing 'women, work' and 'work, women' occur more frequently than sequences containing 'men, work' and 'work, men', both in the Italian and in the German corpus, as the frequency per million words (pmw) reported in tables 3 and 4 shows.

Structure containing/ English translation	Occurrences	Frequency pmw
<i>donne, lavoro</i> (women, work)	7,991	2.60 pmw
<i>lavoro, donne</i> (work, women)	6,766	2.20 pmw
<i>lavoro, uomini</i> (men, work)	2,311	0.80 pmw
<i>uomini, lavoro</i> (work, men)	2,338	0.80 pmw

**Table 3.** Quantitative analysis (itTenTen)

Structure containing/ English translation	Occurrences	Frequency pmw
Frauen, Arbeit (women, work)	10,819	0.54 pmw
Arbeit, Frauen (work, women)	15,600	0.80 pmw
Arbeit, Männer (men, work)	2,950	0.15 pmw
Männer, Arbeit (work, men)	4,297	0.20 pmw

**Table 4.** Quantitative analysis (deTenTen 2013)

Furthermore, the combinations of words *donne, lavoro* [women, work] and *lavoro, donne* [work, women] are the most frequent structures in the itTenTen corpus. As Baker pointed out, a reason for this could be that



minority identities, in this case women (in relation to work), are ‘more unusual and/or problematic and therefore more likely to be remarked upon’ (Baker 2006: 66). In fact, as the women’s employment rate shows (see Table 2) the relation between women and work seems to be quite problematic, especially in Italy.

### 3.1 Lexical Collocates

In this section, the ten most frequent lexical collocates of the above-mentioned structures will be presented and a corresponding English translation will be provided.

Table 5 presents the lexical collocates of the sequences *donne, lavoro* [women, work] and *lavoro, donne* [work, women] collected in the itTenTen corpus (window span -5, +5).

<b>Lexical collocates: <i>donne, lavoro</i></b>	<b>English translation</b>	<b>Lexical collocates: <i>lavoro, donne</i></b>	<b>English translation</b>
<i>uomini</i>	men	<i>uomini</i>	men
<i>partecipazione</i>	participation	<i>mercato</i>	market
<i>opportunità</i>	opportunities	<i>mondo</i>	world
<i>trattamento</i>	treatment	<i>lavoro</i>	work
<i>accesso</i>	access	<i>opportunità</i>	opportunities
<i>lavoro</i>	work	<i>vita</i>	life
<i>pari</i>	equal	<i>donne</i>	women
<i>ingresso</i>	entry	<i>partecipazione</i>	participation
<i>cura</i>	care	<i>condizioni</i>	conditions
<i>famiglia</i>	family	<i>posti</i>	jobs

**Table 5.** Lexical collocates of *donne, lavoro* and *lavoro, donne*

Table 6 provides the ten most frequent lexical collocates of the sequences containing *uomini, lavoro* [men, work] and *lavoro, uomini* [work, men] collected in the itTenTen corpus.

When applying corpus approaches to discourse studies, Taylor (2013: 83) suggested focusing not only on differences, but also on similarities; she stated that:

By focussing on difference, we effectively create a 'blind spot'; this means that, rather than aiming for a 360-degree perspective of our data, we are actually starting out with the goal of achieving only a 180-degree visualisation. Therefore, the search for similarity can add a new range of starting points into our data and allow us to begin with a more ambitious aim regarding the 'completeness' of the analysis.

<b>Lexical collocates: <i>uomini, lavoro</i></b>	<b>English translation</b>	<b>Lexical collocates: <i>lavoro, uomini</i></b>	<b>English translation</b>
<i>donne</i>	women	<i>donne</i>	women
<i>parità</i>	equality	<i>opportunità</i>	opportunities
<i>trattamento</i>	treatment	<i>mercato</i>	market
<i>opportunità</i>	opportunities	<i>lavoro</i>	work
<i>discriminazione</i>	discrimination	<i>pari</i>	equal
<i>pari</i>	equal	<i>divisione</i>	division
<i>dignità</i>	dignity	<i>Polizia</i>	Police
<i>uguaglianza</i>	equality	<i>mondo</i>	world
<i>lavoro</i>	work	<i>vita</i>	life
<i>sostanziale</i>	substantive	<i>condizioni</i>	conditions

**TABLE 6:** Lexical collocates of *uomini, lavoro* and *lavoro, uomini*

Following this suggestion, I looked for lexical collocates that ranked amongst the most frequent ones in more than one of these four sequences. In other words, I highlighted those collocates that the different combinations of words analysed have in common and I collected them in table 7.

Specifically, I highlighted in blue the words that collocate with all of the analysed sequences: i.e. *opportunità* [opportunities] and *lavoro* [work]. I highlighted in green the collocates shared by three out of four sequences: *donne* [women] and *pari* [equal]. Finally, I highlighted in yellow the collocates that occur with two sequences: *uomini* [men], *partecipazione* [participation], *mercato* [market], *mondo* [world], *trattamento* [treatment], *vita* [life] and *condizioni* [conditions].

By looking at table 7, it is clear that the lexical collocates shared by two or more analysed structures outnumber by far the ones that occur with only one sequence.

It is also interesting that many of the above-listed collocates are related to the issue of equal opportunities for men and women in the labour market, as shown by the following collocates: *opportunità* [opportunities], *partecipazione* [participation], *trattamento* [treatment], *condizioni* [conditions], *uguaglianza* [equality], *parità* [equality], *pari* [equal], *ingresso* [entry], *accesso* [access].

This recurrent pattern will be further emphasized in the following section, in which some concordance lines of these lexical collocates are reported.

<b>Lexical collocates: <i>donne, lavoro</i></b>	<b>Lexical collocates: <i>lavoro, donne</i></b>	<b>Lexical collocates: <i>uomini, lavoro</i></b>	<b>Lexical collocates: <i>lavoro, uomini</i></b>
men	men	women	women
participation	market	equality	opportunities
opportunities	world	treatment	market
treatment	work	opportunities	work
access	opportunities	discrimination	equal
work	life	equal	division
equal	women	dignity	Police
entry	participation	equality	world
care	conditions	work	life
family	jobs	substantive	conditions

**Table 7.** Lexical collocates of *donne, lavoro*; *lavoro, donne*; *uomini, lavoro*; and *lavoro, uomini*<sup>17</sup>

After reporting the lexical collocates of the sequences analysed in the itTenTen corpus, the following tables will list the most frequent lexical collocates retrieved from the deTenTen corpus for the sequences containing *Frauen, Arbeit* [women, work], *Arbeit, Frauen* [work, women], *Männer, Arbeit* [men, work] and *Arbeit, Männer* [work, men].

<b>Lexical collocates: <i>Frauen, Arbeit</i></b>	<b>English translation</b>	<b>Lexical collocates: <i>Arbeit, Frauen</i></b>	<b>English translation</b>
<i>Männer</i>	men	<i>Männer</i>	men
<i>weniger</i>	less	<i>Lohn</i>	pay
<i>Lohn</i>	pay	<i>gleich</i>	equal
<i>Frauen</i>	women	<i>Frauen</i>	women
<i>verdienen</i>	earn	<i>unbezahlt</i>	unpaid
<i>leisten</i>	do (work)	<i>gleichwertig</i>	equivalent
<i>Familie</i>	family	<i>Kinder</i>	children
<i>verrichten</i>	do (work)	<i>Mädchen</i>	girl/girls
<i>bezahlt</i>	paid	<i>geleistet</i>	(work) done
<i>bekommen</i>	receive	<i>Arbeit</i>	work

**Table 8.** Lexical collocates of *Frauen, Arbeit* and *Arbeit, Frauen*

<b>Lexical collocates: <i>Männer, Arbeit</i></b>	<b>English translation</b>	<b>Lexical collocates: <i>Arbeit, Männer</i></b>	<b>English translation</b>
<i>Frauen</i>	women	<i>Frauen</i>	women
<i>weniger</i>	less	<i>gleich</i>	equal
<i>jung</i>	young	<i>Lohn</i>	pay
<i>mehr</i>	more	<i>gleichwertig</i>	equivalent
<i>Lohn</i>	pay	<i>verdienen</i>	earn
<i>verrichten</i>	do (work)	<i>unbezahlt</i>	unpaid
<i>verdienen</i>	earn	<i>hart</i>	hard
<i>bezahlt</i>	paid	<i>weniger</i>	less
<i>gleich</i>	equal	<i>Arbeit</i>	work
<i>leisten</i>	do (work)	<i>mehr</i>	more

**Table 9.** Lexical collocates of *Männer, Arbeit* and *Arbeit, Männer*

Following the same methodology described above and thereby focusing not only on differences but also on similarities, Table 10 puts forward the collocates that occur among the most frequent ones with more than one analysed structure.

<b>Lexical collocates: <i>Frauen, Arbeit</i></b>	<b>Lexical collocates: <i>Arbeit, Frauen</i></b>	<b>Lexical collocates: <i>Männer, Arbeit</i></b>	<b>Lexical collocates: <i>Arbeit, Männer</i></b>
men	men	women	women
less	pay	less	equal
pay	equal	young	pay
women	women	more	equivalent
earn	unpaid	pay	earn
do (work)	equivalent	do (work)	unpaid
family	children	earn	hard
do (work)	girl/girls	paid	less
paid	(work) done	equal	work
receive	work	do (work)	more

**Table 10.** Lexical collocates of *Frauen, Arbeit*; *Arbeit, Frauen*; *Männer, Arbeit* and *Arbeit, Männer*

As already stressed in the Italian results, a substantial majority of collocates occur with more than one structure. Two collocates *Frauen* [women] and *Lohn* [pay], highlighted in blue, occur with all the combinations of words examined. *Weniger* [less], *gleich* [equal] and *verdienen* [earn], in green, occur with three out of four sequences. Lastly, the collocates shared by two of the analysed sequences, highlighted in yellow, are the following ones: *Arbeit*

[work], *leisten* [do work], *mehr* [more], *bezahlt* [paid], *gleichwertig* [equivalent], *verrichten* [do work], *unbezahlt* [unpaid], *Männer* [men].

Table 10 indicates that many collocates are related to the specific issue of equal pay for women and men, as for instance *Lohn* [pay], *weniger* [less], *mehr* [more], *gleich* [equal], *unbezahlt* [unpaid], *verdienen* [earn], *bekommen* [receive], *bezahlt* [paid].

This recurring pattern will be confirmed by looking at the concordance lines that will be reported in the following section.

In the light of the results presented so far, it is also worth mentioning that, both in the German corpus and in the Italian one the structures containing ‘men, work’ (and vice versa) and ‘women, work’ (and vice versa) presented no significant differences in terms of collocates.

To conclude, a general overview of the lexical collocates retrieved from both corpora is provided. The violet boxes in Table 11 contain the collocates that occur among the most frequent ones both in the itTenTen analysis and in the deTenTen one. These are: *uomini*, *Männer* [men]; *donne*, *Frauen* [women]; *lavoro*, *Arbeit* [work]; *pari*, *gleich* [equal] and *famiglia*, *Familie* [family].

<i>donne, lavoro</i>	<i>lavoro, donne</i>	<i>uomini, lavoro</i>	<i>lavoro, uomini</i>	<i>Frauen, Arbeit</i>	<i>Arbeit, Frauen</i>	<i>Männer, Arbeit</i>	<i>Arbeit, Männer</i>
men	men	women	women	men	men	women	women
participation	market	equality	opportunities	less	pay	less	equal
opportunities	world	treatment	market	pay	equal	young	pay
treatment	work	opportunities	work	women	women	more	equivalent
access	opportunities	discrimination	equal	earn	unpaid	pay	earn
work	life	equal	division	do (work)	equivalent	do (work)	unpaid
equal	women	dignity	Police	family	children	earn	hard
entry	participation	equality	world	do (work)	girl/girls	paid	less
care	conditions	work	life	paid	(work) done	equal	work
family	jobs	substantive	conditions	receive	work	do (work)	more

**Table 11.** Overview of the lexical collocates

In the following section, some concordance lines of the shared collocates ‘equal’ (*gleich* in German and *pari* in Italian) and ‘family’ (*Familie* in German and *famiglia* in Italian)<sup>18</sup> will be presented and discussed, in relation to the analysis of the discourse on ‘women and work’ in German and in Italian.

### 3.2 Concordance Lines

As previously mentioned, ‘equal’ (*pari* in Italian and *gleich* in German) and ‘family’ (*famiglia* in Italian and *Familie* in German) are two of the most frequent collocates shared by the structures analysed both in the Italian corpus and in the German one. In this section, a sample of concordance lines of these collocates will be presented in order to outline similarities and

differences concerning the way the two words occur within the discourse on 'women and work' in the two corpora (Figures 1 and 2).

In the itTenTen corpus, the word *pari* [equal] occurs 250 times (0.08 pmw) as collocates of the sequence *donne, lavoro* [women, work]. By analysing a sample of 100 occurrences, the word *pari* refers 82 times to *opportunità* [opportunities].



Figure 2. A sample of concordance lines of *pari* [equal]

Furthermore, *pari* [equal] occurs 99 times (0.03 pmw) with the string *uomini, lavoro* [men, work] and 60 times (0.02 pmw) with the string *lavoro, uomini* [work, men]. Considering a sample of 50 occurrences per string, i.e. 100 occurrences in total, the word *pari* refers 84 times to *opportunità*.

In the most of these occurrences, *pari opportunità* [equal opportunities] for women and men in the labour market are depicted as a goal that has to be promoted and achieved. This implicitly proves that, up to now, the situation of women in the labour market is still discriminatory and unequal, since the discourse on equal opportunities for women and men is projected in the future and not situated in the status quo.

Moving on to the two forms (*gleiche* and *gleicher*) of the collocates *gleich* [equal] in the deTenTen corpus, they occur respectively 489 times (0.02 pmw) and 206 times (0.01 pmw) with the string *Arbeit, Frauen* [work, women]; and 454 times (0.02 pmw) and 124 times (0.01 pmw) with the sequence *Arbeit, Männer* [work, men].

For each string, I observed in more detail a sample of 100 occurrences, in which *gleich* [equal] occurs as a collocates. In the sequence 'Arbeit, Frauen', *gleich* refers to *Arbeit* [work] 95 times and in all these occurrences it correlates with the concept of *Gleicher Lohn für gleiche Arbeit* [equal pay for equal work].

Moreover, considering a sample of 100 concordance lines in which *gleich* occurs with the string *Arbeit, Männer*, in all of them *gleich* refers either to

*Arbeit* [work], to *Qualifikation* [qualification] or to *Lohn* [pay], thus expressing again the concept of equal pay for equal work.

Most of the above-listed concordance lines either present the issue of ‘equal pay for equal work’ as a goal yet to be achieved or, alternatively, point out that in the status quo women and men do not receive equal pay for equal work. So, the discriminatory situation that women experience within the labour market in terms of different retribution for the same job is at the core of the discourse.

Query *Arbeit, Männer* 2,950 > Positive filter (excluding KWIC) *gleicher* 124 (0.01 per million) ⓘ

Page 1 of 7  [Next](#) | [Last](#)

theuropea...	Schnitt natürlich immer noch, bei <i>gleicher</i>	<i>Arbeit wie Männer</i>	, versteht sich. </p><p> Was also nun? Jungs
sigmar-gab...	. Dazu gehören <i>gleicher</i> Lohn für gleiche	<i>Arbeit für Frauen und Männer</i>	sowieso ein gesetzlicher Mindestlohn. Als
spd-reiche...	angemessene Bezahlung, <i>gleicher</i> Lohn bei gleicher	<i>Arbeit für Frauen und Männer</i>	. Die arbeitsmarktpolitischen Maßnahmen
dkp-darmst...	Abgesehen davon, dass Frauen bei <i>gleicher</i>	<i>Arbeit weniger verdienen als Männer</i>	. Abgesehen von prügelnden Polizisten und
winteltoosh...	Geschlechterproblem. Auch Männer verdienen bei <i>gleicher</i>	<i>Arbeit weniger als andere Männer</i>	. Oder es gibt Manager, die sich eher rumlummeln
postmaster...	1980er und 1990er Jahren trotz <i>gleicher</i>	<i>Arbeit nicht dem der Männer</i>	entsprech. </p><p> Die Zahlung basiert auf
spd-oberve...	und ist: <i>gleicher</i> Lohn für gleichwertige	<i>Arbeit für Männer</i>	und Frauen", erklärt die Vorsitzende der
blick.de	Frauen bei <i>gleicher</i> Qualifikation die gleiche	<i>Arbeit wie Männer</i>	leisten, erhalten sie im Schnitt 13 Prozent
landtag.It...	dass es den gleichen Lohn für die gleiche	<i>Arbeit gibt - derzeit verdienen Männer</i>	Statistiken zufolge bei <i>gleicher</i> Tätigkeit
cdulandesg...	Gehalt bei <i>gleicher</i> bzw. gleichwertiger	<i>Arbeit für Frauen und Männer</i>	ist nach wie vor keine Selbstverständlichkeit
equalpayda...	ungerecht ist, dass Frauen bei <i>gleicher</i>	<i>Arbeit schlechter bezahlt werden als Männer</i>	." </p><p> Andrea Nahles, SPD-Generalsekretärin
spd-oberve...	dem das Prinzip <i>gleicher</i> Lohn für gleiche	<i>Arbeit für Frauen und Männer</i>	festgeschrieben werden soll. </p><p> EU-weit
ekhn.de	Lohn für Frauen und Männer bei <i>gleicher</i>	<i>Arbeit, stärkere Beteiligung der Männer</i>	bei der Kindererziehung und im Haushalt
blog.eva-h...	verdienen in bestimmten Bereichen bei <i>gleicher</i>	<i>Arbeit noch immer weniger als Männer</i>	und werden durch familienpolitische und
m-media.or...	, da Frauen beispielsweise bei <i>gleicher</i>	<i>Arbeit weniger verdienen als Männer</i>	. "Frauen, die dagegen demonstrieren und
hamburger...	Umsetzung des Prinzips <i>gleicher</i> Lohn für gleiche	<i>Arbeit für Frauen und Männer</i>	eingefordert. Vergleichende Lohnstestverfahren
spd-dobera...	Gleichstellung. Das heißt: <i>gleicher</i> Lohn für gleiche	<i>Arbeit für Frauen und Männer</i>	. Das heißt, dass Frauen künftig 40 Prozent
europarl.a...	Blauflossenthunfische und gleichen Lohn bei <i>gleicher</i>	<i>Arbeit für Frauen und Männer</i>	. Diskutiert wurden auch die politische
siegmund-e...	und Frauen: Frauen verdienen bei <i>gleicher</i>	<i>Arbeit 22 % weniger als Männer</i>	in vergleichbarer Tätigkeit! Das ist ein
powerfraue...	bei <i>gleicher</i> Qualifikation für die gleiche	<i>Arbeit weniger verdienen als Männer</i>	. In Deutschland liegt die Differenz teilweise

Page 1 of 7  [Next](#) | [Last](#)

**Figure 3.** A sample of concordance lines of *gleich* [equal]

Moving on to the second collocate in question, a sample of concordance lines of the collocate ‘family’ (*famiglia* in Italian and *Familie* in German) will be reported in order to observe how this word situates itself within the discourse on ‘women and work’ (Figures 3 and 4).

In the itTenTen corpus, *famiglia* [family] occurs 207 times (0.07 pmw) as collocate of the string *donne, lavoro* [women, work]. As previously explained, I analysed in more detail a sample of 100 concordance lines so as to determine how the relationship between women, work and family is portrayed in the two corpora. The main issue that emerges from the concordance analysis is the work-family balance that many women strive to achieve, despite all the obstacles. The focus is on the multiple difficulties that women have to face in order to work and, at the same time, be part of a family. In fact, another recurrent pattern shows that women still spend more time taking care of their children (sometimes also of their elderly parents) and of the chores.

Query *donne, lavoro* 7,991 > Positive filter (excluding KWIC) *famiglia* 207 (0.07 per million) ⓘ

First | Previous Page 4 of 11 Go Next | Last

home.rifon...	dei servizi pubblici, e, nel caso delle	<i>donne, il doppio peso del lavoro</i>	e della <i>famiglia</i> . Anni ed anni di esperienza
users.libe...	che impegna tutta la <i>famiglia</i> , uomini e	<i>donne, nessuno escluso; un lavoro</i>	, infine, che comporta troppi rischi: una
donne-lavo...	queste associazioni, infatti, sono molte le	<i>donne che lasciano il lavoro</i>	per occuparsi della <i>famiglia</i> , e che trovano
donne-lavo...	valli altoatesine. Sono molte infatti le	<i>donne che lasciano un lavoro</i>	regolare per occuparsi della <i>famiglia</i> ,
libreriade...	caso il tono da sergente di molte giovani	<i>donne impegnate sul lavoro</i>	e in <i>famiglia</i> . Partire dalla relazione
dica33.it ...	una sostanziale uguaglianza tra uomini e	<i>donne in relazione al lavoro</i>	e alla <i>famiglia</i> . Anche in materia sanitaria
dols.it	entusiasmo, la passione e le esperienze di tre	<i>donne alle quali lavoro</i>	, <i>famiglia</i> e figli non bastano. </p><p> Come
luca-volon...	di promuovere la <i>famiglia</i> di sostenere le	<i>donne nel lavoro</i>	e negli altri ruoli sociali. Perché, aveva
donne-lavo...	femminile nelle valli altoatesine. Molte	<i>donne, infatti, lasciano un lavoro</i>	regolare per occuparsi della <i>famiglia</i> ,
delegazion...	giudicato spesso una soluzione positiva per le	<i>donne perché riesce a conciliare lavoro</i>	e <i>famiglia</i> ; occorre però valutarne bene
ninniradic...	Die Zeit, a proposito dei modi in cui le	<i>donne cercano di conciliare il lavoro</i>	e la <i>famiglia</i> , ha detto che comunque "
stpauls.it...	sempre più difficile, poi, soprattutto per le	<i>donne, conciliare lavoro</i>	e <i>famiglia</i> . </p><p> Queste, non il riconoscimento
forum.mila...	e mansione! Eppure la maggioranza delle	<i>donne svolge ancora un lavoro</i>	doppio per la <i>famiglia</i> , organizzando,
dols.it	time, etc...? Ritieni che consentire alle	<i>donne di coniugare facilmente lavoro</i>	e <i>famiglia</i> debba divenire una priorità
sinistra-d...	interpretato da troppi come il dovere delle	<i>donne di farci stare tutto, lavoro</i>	, <i>famiglia</i> , lavoro di cura, come richiesta
jobtel.it ...	'evidentemente ancora difficile, per le	<i>donne, conciliare il lavoro</i>	con la <i>famiglia</i> , e la legislazione di
ondaoserv...	in una società iper-competitiva dove le	<i>donne devono conciliare lavoro</i>	, <i>famiglia</i> , cura personale senza avere
assemblea...	particolare interesse: la presenza notevole delle	<i>donne nel mondo del lavoro</i>	; la fine della <i>famiglia</i> patriarcale e la
women.it	quattro martedì, durante i quali si parlerà di	<i>donne, lavoro</i>	, welfare, <i>famiglia</i> ; politiche regionali
saperi.for...	dell'onestà delle donne. I sacrifici che le	<i>donne fanno per conciliare lavoro</i>	e <i>famiglia</i> e per non veder decurtato il

First | Previous Page 4 of 11 Go Next | Last

Figure 4. A sample of concordance lines of *famiglia* [family]

In addition, some of these concordance lines also report that some women decided to quit their job, because the work-family balance was too difficult to achieve.

In the deTenTen corpus, *Familie* [family] occurs 226 times (0.01 pmw) as collocate of the string *Frauen, Arbeit* [women, work]. Again, a sample of 100 concordance lines has been examined. Exactly as in the case of the Italian concordance analysis, also in the German one the most frequently discussed issue is the work-family balance that many women seek. However, interestingly enough, the texts, in this case, present not only the numerous obstacles that women must overcome, but also some measures that aim at facilitating the achievement of this balance.

Query *Frauen, Arbeit* 10,819 > Positive filter (excluding KWIC) *Familie* 226 (0.01 per million) ⓘ

First | Previous Page 10 of 12 Go Next | Last

igm-nuernb...	übernehmen, liegt unter 5%. So wie es für	<i>Frauen möglich sein muss Arbeit</i>	und <i>Familie</i> zu verbinden, muss es für Männer
balance-be...	für den Mittelstand". </p><p> Qualifizierte	<i>Frauen entscheiden sich für die Arbeit</i>	oder die <i>Familie</i> familienpolitisch und
neofa-auss...	</p><p> Der Einsatz für gleiche Rechte der	<i>Frauen in Bildung, Arbeit</i>	und <i>Familie</i> wird bekämpft. </p><p> Foto:
partner-fur...	für Arbeit und Sozialordnung, <i>Familie</i> und	<i>Frauen und der Bundesagentur für Arbeit</i>	München ins Leben gerufen worden, um Mädchen
verwaltung...	der Frauen ist mir ein Anliegen. Vor allem	<i>Frauen schieben durch die Mehrfachbelastung wie Arbeit, Familie</i>	und Haushalt ihre eigenen Bedürfnisse
e-hausaufg...	Bundesministerium für Jugend, <i>Familie</i> ,	<i>Frauen und Gesundheit nimmt die Arbeit</i>	auf. 1988 Das Erziehungsgeld und der Erziehungsurlaub
allmystery...	<p> Hast Recht! Zuerst dachte ich, dass die	<i>Frauen durch die Doppelbelastung "Arbeit</i>	und <i>Familie</i> " vielleicht als Ausgleich
rnz.de	Bundesministerium für <i>Familie</i> , Senioren,	<i>Frauen und Jugend fördert diese Arbeit</i>	mit monatlich 350 Euro pro Freiwilligem
stjosef.at...	Situation unserer Gesellschaft fällt es vielen	<i>Frauen nicht leicht, die die Arbeit</i>	in der <i>Familie</i> und einen außerhäuslichen
lportal.g...	Konservativität der <i>Familie</i> ab, inwieweit	<i>Frauen einer Arbeit</i>	nachgehen können. Bis zum Umsturz war es
emanzipati...	missen. 1982 sagten 60% der 20-40-jährigen	<i>Frauen aus, daß Arbeit</i>	und <i>Familie</i> für sie die gleiche Bedeutung
ingrid-fis...	dabei für die Bereiche <i>Familie</i> , Senioren,	<i>Frauen, Jugend, Arbeit</i>	und Sozialen, Arbeitnehmer und Kirchen
frauenbund...	Fraktionsvorstand die Bereiche <i>Familie</i> , Senioren,	<i>Frauen und Jugend, Arbeit</i>	und Soziales, Kirchen und Arbeitnehmer.
spd.de	Zahlen. Es gibt nämlich tatsächlich auch	<i>Frauen, die eine Balance zwischen Arbeit</i>	und <i>Familie</i> halten können bzw. generell
vera-hewen...	www.geocities.com/wissenboerse. </p> <gap/> <gap/> <p>	<i>Frauen leisten die meiste Arbeit</i>	in der <i>Familie</i> . Sie sind kommunikativer
abgeordnet...	Bundesministerium für <i>Familie</i> , Senioren,	<i>Frauen und Jugend die Arbeit</i>	einer bundesweiten /Gründerinnenagentur
dagmar-enk...	wenig Plätze zur Kinderbetreuung, damit	<i>Frauen Arbeit</i>	und <i>Familie</i> vereinbaren können. Frauen
bvs.de	der Leitstelle für die Gleichstellung von	<i>Frauen und Männern (Ministerium für Arbeit</i>	und Sozialordnung, <i>Familie</i> und Frauen)
adventure...	Bundesministerin für <i>Familie</i> , Senioren,	<i>Frauen und Jugend würdigte die vorbildliche Arbeit</i>	von Sabriye Tenberken. Der Wohltätigkeitspreis
gtg.tu-ber...	Landwirtschaft oder die Doppelbelastung von	<i>Frauen durch Arbeit</i>	und <i>Familie</i> . Dyhrenfurth verfolgte mit

First | Previous Page 10 of 12 Go Next | Last

Figure 5. A sample of concordance lines of *Familie* [family]

From a quantitative perspective, the results of the collocate *family* in itTenTen and in deTenTen corpora are difficult to compare. This is due to the fact that in the Italian corpus *famiglia* occurs 0.07 pmw as collocate of the sequence *donne, lavoro* [women, work], whereas in the German one *Familie* occurs only 0.01 pmw as collocate of the string *Frauen, Arbeit* [women, work].



## 4. Conclusions

As highlighted by the external data provided in the first section (i.e. women's employment rate, The Global Gender Gap Index and the gender pay gap), Italy and Germany have a different situation in terms of women and work. This different situation is also mirrored by the way the discourse on 'women and work' is constructed in the two analysed corpora.

On the one hand, results show that the major issue addressed in the German corpus is equal pay for equal work for women and men, as brought forward by collocates like: 'pay', 'earn', 'paid', 'receive', 'equal', 'more', 'less', 'unpaid'.

On the other hand, results indicate that the main issue regarding 'women and work' in the Italian corpus is that of equal opportunities for women and men in the labour market, as emphasized by collocates such as 'opportunities', 'participation', 'equal', 'equality', 'entry', 'conditions', 'treatment' and 'access'. The word 'equal' is emblematic in this regard. In fact, while in the German corpus 'equal' refers mostly to the concept of 'equal pay for equal work', in the Italian one, it is connected to 'opportunities' in 80% of the occurrences analysed.

Besides these differences, the findings obtained through this analysis have also revealed some similarities concerning how the discourse on 'women and work' is constructed in the two corpora. In fact, as the shared collocate 'family' shows, the relationship between family, women and work is part of the discourse both in the German and in the Italian corpora. Moreover, within this sub-discourse, the issue of work-family balance, which many women seek, is a recurring topic in both countries.

After providing an overview of the results achieved through this cross-linguistic corpus assisted-discourse study, some limitations of the present study, which concern mainly issues of comparison.

First, as already mentioned, the texts analysed in itTenTen were crawled in 2010, whereas the ones of deTenTen were crawled in 2013. Second, it was not possible to retrieve the geographical origin of the texts included in the corpora. This point could be particularly relevant for the German corpus, since it may contain texts produced not only in Germany but also in other German-speaking countries, such as Switzerland and Austria.

To conclude, the aim of this study was to provide an insight as to how the discourse on 'women and work' is constructed in two general corpora, drawing on the methodology of corpus-assisted discourse studies. The research findings proved that pressing issues in society become recurrent topics in the language. Concerning the subject of 'women and work', the external data provided above show that the gender pay gap is the main problem in Germany, whereas women's unemployment rate is the most critical issue in Italy. It is not a coincidence that in the German corpus the central theme within the discourse of 'women and work' revolves around the concept of 'equal pay for equal work.' And by the same logic, the most recurrent subject concerning 'women and work' in the Italian corpus is related to the opportunities that women should have in order to enter the labour market and the lack thereof.

On a personal note, as mentioned before, it is apparent that, as of today, cross-linguistic corpus-assisted discourse studies are still a niche and that those focusing on languages other than English are even more so. Thus, I wanted to contribute to expand this promising field of research.

## Notes

- 1 For an overview of the publications that use this methodology, see the list compiled by Gabrielatos: <http://www.gabrielatos.com/CLDA-Biblio.htm>.
- 2 As far as I could retrieve, the only exception is the analysis on prescriptivism in and about the media in Greece and in Germany (Moschonas and Spitzmüller 2010).
- 3 The Global Gender Gap Index is available at this page: <http://reports.weforum.org/global-gender-gap-report-2016/rankings/>. Last accessed 1 May 2018.
- 4 In order to provide comparable data collected and published by the same source, EUROSTAT data, which are collected by the European Union, have been preferred over data gathered by Germany and by Italy. The data presented here are available at: [http://ec.europa.eu/eurostat/tgm/graph.do;jsessionid=-S3VcHk9CSWBcOssevapmhC8\\_WdKv\\_jnQzYsVHcJkRfAl9rxt6ha!-1635108100?tab=graph&plugin=1&pcode=t2020\\_10&language=en&toolbox=data](http://ec.europa.eu/eurostat/tgm/graph.do;jsessionid=-S3VcHk9CSWBcOssevapmhC8_WdKv_jnQzYsVHcJkRfAl9rxt6ha!-1635108100?tab=graph&plugin=1&pcode=t2020_10&language=en&toolbox=data). Last accessed 1 May 2018.
- 5 For further information see: <https://www.weforum.org/reports/the-global-gender-gap-report-2017>. Last accessed 1 May 2018.
- 6 The data quoted above, as well as further information, are available at the page: <http://reports.weforum.org/global-gender-gap-report-2016/economies/#economy=DEU>. Last accessed 1 May 2018.
- 7 The data quoted above as well, as further information, are available at the page: <http://reports.weforum.org/global-gender-gap-report-2016/economies/#economy=ITA>. Last accessed 1 May 2018.
- 8 Since they are not available online, it was not possible to retrieve data collected in 2010 and in 2013 concerning the Global Gender Gap Index.
- 9 For further information see: <https://www.sketchengine.co.uk/ittenten-corpus/>. Last accessed 1 May 2018.
- 10 For further information see: <https://www.sketchengine.co.uk/detenten-corpus/>. Last accessed 1 May 2018.
- 11 It is important to note that the present research has been carried out on the 2010 version of ItTenTen.
- 12 As this is a bilingual study, the choice of the equivalent words (i.e. *Arbeit* and *lavoro*) has been limited to linguistic, cultural and translation issues. In fact, I am aware that finding a perfect semantic correspondence between two words belonging to different languages may be, in some cases, a utopian goal.
- 13 For further information about Corpus Query Language see: <https://www.sketchengine.co.uk/documentation/corpus-querying/>. Last accessed 1 May 2018.
- 14 Based on Baker 2006, I decided to include a window span of 5 words.
- 15 As Baker (2006: 102) pointed out in *Using Corpora for Discourse Analysis*, this statistical technique allows obtaining both high frequency and low frequency words.
- 16 Due to time constraints, I decided not to translate the screenshots of the concordance lines reported in this section.
- 17 Due to space constraints, in the overview tables (see Tables 7, 10, and 11) I decided to report only the English translation of the collocates, since the German and the Italian words have already been reported in Tables 5, 6, 8, and 9.

18 I decided to analyse the shared collocates 'equal' and 'family' rather than 'women', 'men' and 'work' because the latter are both shared collocates and search terms.

## References

- Baker, P. (2006). *Public Discourses of Gay Men*. London: Routledge.
- Baker, P. (2006a). *Using Corpora in Discourse Analysis*. London: A&C Black.
- Baker, P., C. Gabrielatos, M. Khosravini, M. Krzyzanowski, T. McEnery, and R. Wodak (2008). A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse & Society* 19(3): 273-306.
- Bubenhofer, N. (2008). Diskurse berechnen? Wege zu einer korpuslinguistischen Diskursanalyse. In J. Spitzmüller, and Warnke, I.H. (eds.), *Methoden der Diskurslinguistik: Sprachwissenschaftliche Zugänge zur transtextuellen Ebene*. Berlin: de Gruyter. pp. 407-434.
- Bubenhofer, N. (2009). Sprachgebrauchsmuster: Korpuslinguistik als Methode der Diskurs- und Kulturanalyse. Berlin: de Gruyter.
- Bubenhofer, N. (2013). Quantitativ informierte qualitative Diskursanalyse: Korpuslinguistische Zugänge zu Einzeltexten und Serien. In K.S. Roth, and C. Spiegel (ed.), *Angewandte Diskurslinguistik: Felder, Probleme, Perspektiven*. Berlin: Akademie Verlag. pp. 109-134.
- Bubenhofer, N. (2013a). Skandalisierung korpuslinguistisch. Ein empirisch-linguistischer Blick auf die Berichterstattung zur "Wulff-Affäre". *Linguistik Online* 61(4): 31-56.
- European Institute for Gender Equality. (2017). Work in the EU: women and men at opposite ends. Vilnius: European Institute for Gender Equality. Available: <http://eige.europa.eu/rdc/eige-publications/work-eu-women-and-men-opposite-ends>. Last accessed 1 May 2018.
- Eurofound. (2016). *The Gender Employment Gap: Challenges and Solutions*. Luxembourg: Publications Office of the European Union. Available: [http://www.astrid-online.it/static/upload/ef16/ef1638en\\_o.pdf](http://www.astrid-online.it/static/upload/ef16/ef1638en_o.pdf). Last accessed 1 May 2018.
- Freake, R. (2012). A cross-linguistic corpus-assisted discourse study of language ideologies in Canadian newspapers. In *Proceedings of the 2011 Corpus Linguistics Conference*. Birmingham University: 1-11.
- Hausmann, R., L.D. Tyson, Y. Bekhouche, and S. Zahidi (2016). The global gender gap index 2014. *World Economic Forum*. [online]. Available: <http://reports.weforum.org/global-gender-gap-report-2014/part-1/>. Last accessed 1 May 2018.
- Jakubíček, M., A. Kilgarriff, V. Kovář, P. Rychlý, and V. Suchomel (2013). The TenTen corpus family. In *7th International Corpus Linguistics Conference CL*: 125-127.
- Jaworska, S., and R. Krishnamurthy (2012). On the F word: A corpus-based analysis of the media representation of feminism in British and German press discourse, 1990-2009. *Discourse & Society* 23(4): 401-431.
- Kilgarriff, A., V. Baisa, J. Bušta, M. Jakubíček, V. Kovář, J. Michelfeit, P. Rychlý, and V. Suchomel (2014). The Sketch Engine: ten years on. *Lexicography* 1(1): 7-36.
- Moschonas, S., and Spitzmüller, J. (2010). Prescriptivism in and about the media: A comparative analysis of corrective practices in Greece and Germany. In S. Johnson, and T.M. Milani (eds.), *Language Ideologies and Media Discourse: Texts, Practices, Politics*. London: Bloomsbury Publishing. pp. 17-44.
- Partington, A., A. Duguid, and C. Taylor (2013). *Patterns and Meanings in Discourse: Theory and Practice in Corpus-Assisted Discourse Studies (CADS)*. Amsterdam and Philadelphia: John Benjamins Publishing.
- Schafroth, E. (2015). Vergleichende Diskursanalyse als romanistische Forschungspraxis: Zur Rezeption des Sarrazin-Diskurses in französischen, italienischen und spanischen

- Printmedien. In A. Heinemann and C. Schlaak (eds.), *Politische und mediale Diskurse. Fallstudien aus der Romania*. Berlin: Frank & Timme. pp. 57-82.
- Sketch Engine. (2018). *Sketch Engine: Learn how Language Works*. [online]. Available: <https://www.sketchengine.co.uk>. Last accessed 18 November 2017.
- Taylor, C. (2013). Searching for similarity using corpus-assisted discourse studies. *Corpora* 8(1): 81-113.
- Taylor, C. (2014). Investigating the representation of migrants in the UK and Italian press: A cross-linguistic corpus-assisted discourse analysis. *International Journal of Corpus Linguistics* 19(3): 368-400.
- Tognini-Bonelli, E., and Manca, E. (2002). Welcoming children, pets and guests. A problem of non-equivalence in the languages of Agriturismo and Farmhouse Holidays. *Textus* 15(2): 317-334.
- Vessey, R. (2013). Challenges in cross-linguistic corpus-assisted discourse studies. *Corpora* 8(1): 1-26.
- Ziem, A. (2013). Krise im politischen Wahlkampf: linguistische Korpusanalysen mit AntConc. In F. Liedtke (eds.), *Die da oben: Sprache, Wahlen, Partizipation*. Bremen: Hempen. pp. 69-90.