Word2Vec study on the new KubHist corpus

Lawin Khalid

Term paper for Language Technology Resources
University of Gothenburg
`guskhala@student.gu.se`

Abstract

This study presents neural word embedding experiments on a new version of the Swedish historical newspaper corpus, KubHist. We investigate whether the new, extended corpus can overcome some issues that the older version suffered from when used with Word2Vec. Findings show a similar behaviour on the new version as on the old version of the corpus, when used with Word2Vec. However, the larger size, together with our approach, solves some of the issues, but instead amplifies issues regarding OCR errors in the vectors.

1 Introduction

Historical corpora are, especially when available in digitalized format, an important source for automatic word sense change studies. Most previous studies have focused on English, largely for availability reasons. Naturally, English offer more accessible data than other languages. Nevertheless, Språkbanken offer a historical corpus, KubHist, with a first version publicly available at Korp, Språkbanken[1], and a second version being under process and not yet publicly published (Adesam et al., 2018). Tahmasebi (2018) performed a study on the first version of KubHist, investigating whether neural embeddings with Word2Vec could create stable enough vectors to be efficiently used on the KubHist corpus, which suffers from low volume, and low quality regarding OCR errors. The low volume was a result of splitting the existing corpus into 1-year chunks, as a necessary step towards automatic word sense detection. One assumption in this study was that word meaning

does not change radically from one year to another. Hence, a high quality corpus should create similar vectors for consecutive years. The study found low similarities between vectors, indicating unstable vectors. Moreover, the similarity scores were correlated to the frequency of the words, and more so for the infrequent words, indicating that the stability of the vectors are dependent of the frequency. (Tahmasebi, 2018)

In this study we investigate the new version of the KubHist corpus in a similar manner as Tahmasebi (2018), but with some changes in approach. The aim is to examine whether the larger and higher quality corpus can withstand the issues that appeared in the previous study, but we also hope to get new findings in general, since the newer version of KubHist has previously not been used in a study of this type. We observe our results both isolated and in contrast to the previous study.

2 Method

The data is split into smaller chunks and we train models with Word2Vec (Mikolov et al., 2013a) separately for each chunk. To be able to compare vectors of different models, we take 12 words and look at the 30 most similar words in the vector spaces for these words in each model. We then calculate the degree of overlap of the 30 words between consecutive years. Following the assumption in Tahmasebi (2018), that word meaning does not change radically in a short period, we interpret the degree of overlapped words as a stability measure for the vectors.

Tahmasebi (2018) suggests that SVD should be better suited than Word2Vec for historical texts. However, we opted to use Word2Vec in the present study for several reasons, mainly due to time constraints and the fact that our models include bigger data. A comparison between Word2Vec and

---

[1] `https://spraakbanken.gu.se/korp/?mode=kubhist`

SVD on a corpus of 2.8 billion tokens showed that Word2Vec outperformed SVD on almost all tasks (Baroni et al., 2014). However, since we split our data into smaller chunks, each model will be trained on smaller sizes. Nevertheless, Altszyler et al. (2016) conclude that Word2Vec outperforms LSA/SVD at corpus sizes as small as 2-3 million tokens. Most of our corpora are between 10 million and 500 million tokens. Hence, we expect no negative consequences of using Word2Vec over count based word embedding methods.

Another factor is time constraints. One big benefit of Word2Vec is that it is much faster on larger corpora. Test runs on the smaller sets (1700 and 1750) confirmed the extended time to build SVD models compared to Word2Vec models.

2.1 The data

The digital corpus of Swedish newspapers, KubHist, was scanned and OCR processed by the Royal Swedish Library, and thereafter annotated by Språkbanken. The new version of the KubHist corpus ranges over more years than the previous version. Version 2 ranges from 1645 to 1926, compared to 1749 to 1925 for version 1. Moreover, more data has been added also for the years that were included in version 1. Version 2 includes 5.5 billion tokens in raw data (Adesam et al., 2018), compared to 1.1 billion in version 1.

The KubHist corpus were split into 10-year chunks, with a few exceptions. Some years in the first 100 years does not contain any data. Furthermore, the distribution of data is far from constant across the years, peaking in the late 19th century, and with very little data for the 17th century and the first half of the 18th century. For this reason, the first two models were built with data from 1645-1699 and 1700-1749. Also, the last model included only seven years, 1920-1926, since the corpus doesn't contain data after 1926.

We make use of raw data, without infrastructure, in contrast to xml files with POS-tags, sentence boundaries and lemmas. We preprocess the data by lowercasing, removing punctuation, and replacing numbers with <num>. Punctuation was replaced with empty string, rather than whitespace. The significant difference in tokens, between Adesam et al. (2018) and my count is likely a result of these preprocessing steps.

An entry like *3:e* (cf. Eng. 3rd), would with many tokenizers result in the three tokens *3, :*

and *e*, while my method only produces the token *<num>e*. The same applies to numbers in thousands, such as 12,000. Other similar abbreviations were common in historical Swedish, such as *D:r* for *doktor* (Eng. doctor), *d:o* for *dito*. The aim was to maintain these abbreviations as one word, which led to the choice of replacing punctuation with empty string rather than whitespace.

Furthermore, compounds in historical Swedish were in some cases created with a hyphen, such as *Regements-Pastor*, *Landtbruks-Akademien* and *Pariser-kabinettet* while other times created similar to modern Swedish, without hyphen, such as *hospitalskyrka* and *landsortsprenumerant*. In order keep these compounds intact when removing punctuation, we had to replace them with empty string. This did also produce a uniform format for compound words, which was preferred compared to separating them to two words in some cases, but keeping them compounded for others. We judged that these measures were especially important in a task that includes word embeddings, and consequently builds its model based on neighbouring words.

2.2 The models

The implementation is done in python. After preprocessing, we split the data into 10-year bins. The reason for this is twofold. One reason is that the data is more easily manageable, to fit the scope of this project. Secondly, and more importantly, this creates more data for each model. Since we train each model separately, splitting them into 1-year chunks would not leave enough data to each model. We compare our results with Tahmasebi (2018), and we are interested in if more data will solve some of the issues encountered in this previous study. Splitting in larger chunks, as well as using the newer, larger version of KubHist, allowed us to train models with a size of up to 500 million tokens, compared to at most 14 million tokens in the previous study (Tahmasebi, 2018).

There were a few exceptions to the 10-year chunks. The 17th century did not offer much data, so all these years were included in one chunk. For the same reason, we included the first half of the 18th century in one chunk. The barrier for this was set to 1748, in order to include 1749 in the next chunk, which would give a more fair comparison with the previous study, due to the fact that the previous version of KubHist started at 1749. Lastly,

Time span	# tokens
1645-1699	1.0 M
1700-1748	4.3 M
1749-1759	2.7 M
1760-1769	5.8 M
1770-1779	15.0 M
1780-1789	22.7 M
1790-1799	27.0 M
1800-1809	35.5 M
1810-1819	42.1 M
1820-1829	55.9 M
1830-1839	91.0 M
1840-1849	143.7 M
1850-1859	204.0 M
1860-1869	311.5 M
1870-1879	400.3 M
1880-1889	512.1 M
1890-1899	358.3 M
1900-1909	34.3 M
1910-1919	36.2 M
1920-1926	9.9 M

Table 1: Time span and size of the different models. Number of tokens per chunk.

the last chunk ranges from 1920 to 1926, because there are no more data after this year. A complete table of the time spans used to build the models, and their corresponding number of tokens, are shown in Table 1. The labels used in the plottings and the Appendix are the first year of the chunk, and can be read for example "1820s" for the label 1820, that includes the years 1820-1829. Exceptions are label 1600 for 1645-1699 and 1750 for 1749-1759.

We feed the data to Gensim's Word2Vec, building separate models for each chunk. Word2vec requires the data to be fed one sentence at the time. However, for unknown reason, our data did not include sentence boundaries, so instead we fed Word2Vec one paragraph at the time, using an empty newline to identify paragraph boundaries. It is unclear to what extent this split affected the result.

We build vectors with 300 dimensions in our experiments. Mikolov et al. (2013a) noted a significant improvement when raising dimensions from 100 to 300, for both medium and large corpora, but only a moderate improvement from 300 to 600 dimensions. Again, time constraints forced us to focus on achieving the best model possible without exhaustive load.

The skip-gram architecture was used in favour of continuous-bag-of-words (CBOW). Skip-gram predicts surrounding words given the current word. This architecture takes longer time to train than CBOW, that predicts the current word given its surrounding words. In a previous study, skip-gram performed significantly better in semantic tasks, and slightly worse in syntactic tasks (Mikolov et al., 2013a). Another study showed slightly better performance for both semantic and syntactic word analogy tasks with skip-gram compared to CBOW (Bojanowski et al., 2016).

Negative sampling was set to the default value of 5. Although previous studies showed a slight improvement in accuracy when using the higher value 15, this came with more than twice as long training time (Mikolov et al., 2013b). The rest of the hyper-parameters were also left at their default values. This include subsampling of 0.001, window size of 5, and minimum count of 5. Training all our models took roughly 3.5 days on a mid-end laptop.

We observe the same words as in the previous study (Tahmasebi, 2018). These are *man* (man), *kvinna* (woman), *qvinna* (old spelling of woman), *glad* (happy), *resa* (travel), *telefon* (telephone), *telegraf* (telegraph), *nyhet* (news), *tidning* (newspaper; magazine), *politik* (politics), *musik* (music), and *retorik* (rhetoric). In contrast to the previous study, we were able to build vectors also for the word *retorik*. However, the frequency is very low (see Figure 3 and Appendix B). We opt to present the plots and the lists of most similar words for *retorik*, but due to low frequency, and to be consistent with previous study, we exclude it from the summary tables (Tables 3 and 4). The total frequency of the word *retorik* was 622 in the present study, compared to 78 in the previous study.

2.3 Similarity measures

The previous study used Jaccard similarity as a similarity measure between vectors. Jaccard is calculated by taking the number of overlapped words, divided by the number of unique words, or more formally, the size of the intersection divided by the size of the union of two sample sets, as shown in (1).

$$Jaccard(A, B) = \frac{|A \cap B|}{|A \cup B|} \qquad (1)$$

This measure does, however, not take rank into account. Thus, an exact overlap of the top word has the same weight as an overlap of the top word in the first set and the bottom word of the second set. We judge this to be a poor measure for a study of this kind, as we believe that rank plays an important role when comparing the most similar words of word embeddings. That is, stable vectors should not only be determined by number of overlapped words between years, but also the positions of these overlaps. We tried two other similarity measures, Average Overlap (AO), presented as Average Accuracy in Wu and Crestani (2003), and Rank-biased Overlap (RBO), presented in Webber et al. (2010). These measures place more emphasis on disorder at the top of the rankings, than disorder at the bottom, which is highly preferred when comparing ranked lists, as in the present study.

A simple comparison of the three similarity measures is shown in Table 2. AO calculates the overlap at each depth and takes the average score to represent the measure. RBO has a more complex calculation; for a deeper understanding of this measure, see Webber et al. (2010). To summarize, RBO is top-weighted, handles non-conjointness satisfyingly, and also estimates similarity between incomplete ranks. Its parameter p determines the top-weighting strength, and was set to the default value of 0.98 throughout this study.

The examples in the first group of Table 2 all have one overlap with the first word, but with at different ranks. The next group have all words overlapped, but with one disorder at various ranks. The last example also has complete overlap but the words are in reverse order, resulting in the middle word to be exact overlap. We can see that Jaccard has the same score for all examples in each group, disregarding the rank of the overlap. AO weights overlap at the top more heavily than at the bottom, resulting in a wide range of scores between 0.04 and 0.46 in the first group. However, the strength of the top-weight is so high that the first example, with only one overlap, gets a higher score than the last example, where all words overlap, but in reverse order. We judge this to be too heavily top-weighted for the present study. We want a measure that weights the number of overlapped words higher than the order, but still take order into account. RBO fulfils this as the scores does change with disorder at different ranks, but

[a, b, c, d, e]	Jacc	AO	RBO
[**a**, x, y, z, r]	0.1111	0.4567	0.2254
[x, **a**, y, z, r]	0.1111	0.2567	0.2054
[x, y, **a**, z, r]	0.1111	0.1567	0.1956
[x, y, z, **a**, r]	0.1111	0.0900	0.1892
[x, y, z, r, **a**]	0.1111	0.0400	0.1845
[a, b, c, d, e]	1.0000	1.0000	1.0000
[a, b, c, *e*, *d*]	1.0000	0.9500	0.9953
[a, b, *d*, *c*, e]	1.0000	0.9333	0.9936
[a, *c*, *b*, d, e]	1.0000	0.9000	0.9902
[*b*, *a*, c, d, e]	1.0000	0.8000	0.9800
[*e*, *d*, c, *b*, *a*]	1.0000	0.4167	0.9429

Table 2: Comparison of different similarity measures. Rank-biased Overlap was preferred over Jaccard similarity (not top-weighted), and Average Overlap (too heavily top-weighted).

not in a radical manner. The values of the first group, with 20% overlap, are all close to 0.2, but shifting somewhat depending on rank. Similarly, the values of the second group, with 100% overlap, are all close to 1.0. Hence, we opt to use RBO for this study. However, since the previous study made use of Jaccard, we also use Jaccard in addition to RBO to be able to do a fair comparison with the previous study.

3 Results and discussion

In this section we present and discuss the results. We study the results in different ways. First, we will look at the general behaviour of the results. Next we look closer at each word over the years, by analysing graphs of RBO and frequency. We then move on to a brief study of the 30 most similar words of each word across the years. Finally, we compare our results with the previous study.

3.1 Overall behaviour

A summary of the results are shown in Table 3. In this table, we only include vectors from year 1749, to be consistent with the previous study. Also, the word *retorik* is excluded from the table. From here, we can observe several findings. First, we see that RBO is increasing with increasing frequency. The first half of the table, including *telegraf*, all have RBO scores below 0.20 and normalized frequency below 200. Apart from *nyhet*, which shows a completely different behaviour, these words also have a high Spearman correlation between RBO and frequency. The second half of

word	avg RBO (A)	norm freq (B)	abs freq (C)	corr (A, B)	corr (A, C)
qvinna	0.119	53.12	16946	0.877	0.877
kvinna	0.063	73.86	6518	0.797	0.852
politik	0.092	78.79	14488	0.544	0.822
nyhet	0.150	129.54	21577	-0.005	0.022
telefon	0.031	142.47	23882	0.774	0.800
telegraf	0.085	164.52	40125	0.735	0.847
glad	0.240	306.75	40177	0.828	0.328
musik	0.329	347.62	61517	0.591	0.650
resa	0.385	977.66	146212	0.074	0.233
tidning	0.252	980.57	172273	-0.174	0.314
man	0.243	9923.89	1383250	-0.010	0.647
Average	0.181			0.457	0.581

Table 3: Average RBO, average normalized frequency, average absolute frequency, and Spearman correlation. Sorted by increasing normalized frequency. Calculated on top 30 most similar words, for the years 1749-1926. The values are slightly different compared to the values in the plots, due to different time spans.

the words, starting with *glad*, show an RBO score above 0.20 and a normalized frequency above 300. Although *glad* and *musik* show relatively high correlation between RBO and normalized frequency, when the frequency comes closer to 1000, as for the last three words, the correlation is very low and close to zero. This is reasonable, and in line with previous study. When the models can be trained with many instances of the words, the vectors are no longer as dependent on frequency.

We also look at absolute frequency. We hypothesize that absolute frequency is also an important factor for stable vectors, as this measure also reflects the size of the corpus, which undoubtedly is of high importance. Here we see some differences compared to normalized frequency. We can see that the average correlation is higher between RBO and absolute frequency (0.581 vs. 0.457). This indicates that the stability of the vectors are stronger dependant on absolute frequency than normalized frequency. Moreover, we observe that this dependency holds stronger for higher frequencies. For example, the word *man* with by far the highest frequency, still shows a relatively strong correlation between RBO and absolute frequency, while this is not true for normalized frequency. Observing the plotting of this word in Figure 3, we see that the highest normalized frequency is the first point in the graph, that is, the years 1645-1748. However, the size of the corpus was very small for this period (see table 1), and likewise also the RBO is small. Instead the RBO peaks at

the 1880s, which is also the peak of the absolute frequency of the word *man*, as well as the peak of the corpus size. Plottings with absolute frequency are shown in Appendix B. Absolute frequency for most words show a similar pattern as corpus size in general, with a low frequency during 17th and 18th century, rising during the 19th century, peaking in the 1880s and then drops to a low frequency again for the 20th century.

3.2 Plotting

Next, we plot the RBO score and the normalized frequency for each word. In contrast to the previous study, we find no reason to smooth the RBO before plotting, since each chunk already includes 10 years. We plot the RBO score of the 30 most similar words between two consecutive chunks of years. That is, for every point on the x-axis the RBO value is calculated on the 30 most similar words in the model of that year, and the 30 most similar words in the model of the previous year. In the same manner, we plot the average of the normalized frequency of year *T* and *T-1*. For unknown reason, the previous study did not average the normalized frequency between the two years which vectors were compared. We argue that, since we compare the vectors of two years, calculate the similarity between them, and also calculate the correlation with the frequency, it is important to include the frequency of both years in the calculation.

The plots are presented grouped by frequency.

Figure 1: Plotting of RBO (left axis) and normalized frequency (right axis) for words with high frequency. Calculated on top 30 words. Time span 1645-1926.
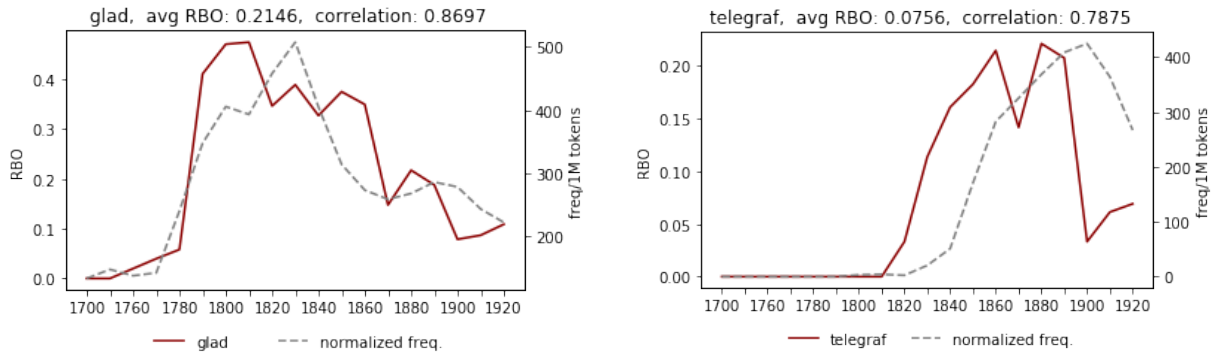
Figure 2: Plotting of RBO (left axis) and normalized frequency (right axis) for words with moderate frequency. Calculated on top 30 words. Time span 1645-1926.

Figure 1 shows RBO and normalized frequency for the words with the highest frequency, plotted for the time span 1645-1926. We also present the average RBO and the Spearman correlation between RBO and normalized frequency at the top of each graph. Note that due to different time spans, the values slightly differ from the ones presented in Table 3. Characteristics of these words are that they show a relatively high RBO across years - these are the four words with the highest RBO, ranging from 0.2356 to 0.3726. Moreover, we see a low correlation between RBO and

frequency. Similar to the previous study, an exception to this, is the word *musik*, which shows a high correlation. We can at this point not conclude what the reason for this is, but we note that the behaviour is similar across the two studies, and that this is the word in this group with the lowest average frequency. Apart from this, the Spearman correlation is very close to zero, indicating that these words have reached a high enough frequency that they are no longer dependent on frequency to build stable vectors. While the previous study showed similar behaviour, the correlation was still signifi-
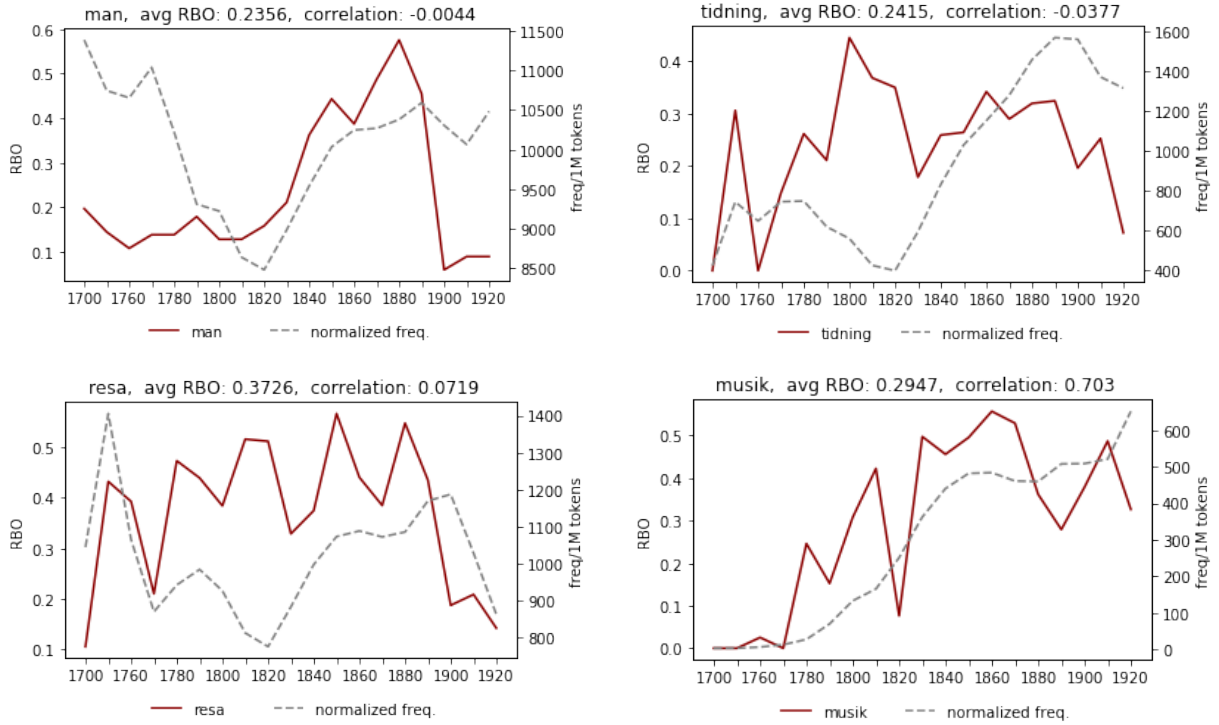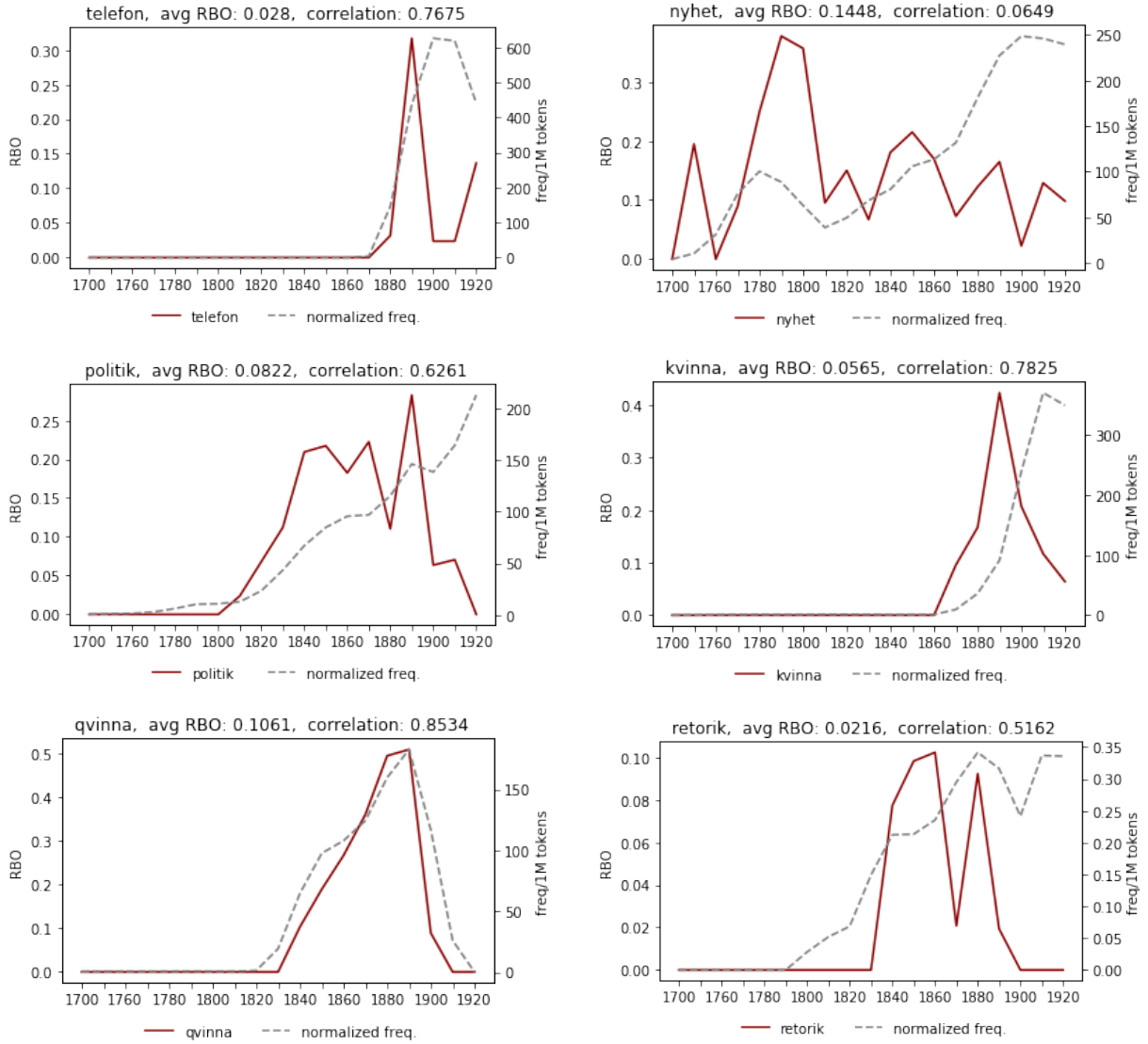
Figure 3: Plotting of RBO (left axis) and normalized frequency (right axis) for words with low frequency. Calculated on top 30 words. Time span 1645-1926.

cant for most words, indicating that the corpus size was an important factor for the stability of the vectors. In the present study, we can instead assume that the instability of the vectors in this group is instead due to the OCR errors. Important to note, though, is that also these words still show a correlation with absolute frequency, as shown in Table 3 and Appendix B.

We move on to Figure 2, which shows two words with moderate frequency. Regarding RBO, these two words are very different. While *glad* is showing an RBO similar to the previous group, *telegraf* is showing a low RBO, similar to the words with low frequency in Figure 3. However, the average normalized frequency is almost doubled between the words. Looking at correlation, it

is high for both words. These two words lie in between the other two groups, and do not offer much valuable information. We can however note that *glad* was an outlier also in the previous study.

Next, focusing on the words with lowest frequency in Figure 3, we see similar patterns between the words. Again, we find an outlier, *nyhet*, which shows a moderate RBO and a low correlation. This was also visible in the summary table 3, and also true for correlation with absolute frequency. Apart from this, all words show a very similar behaviour. Characteristics are low RBO and high correlation. The word *retorik* has a so low frequency that we should not put too much emphasis on it. Nevertheless, also this word follows the pattern. The low RBO and high corre-

lation, could however, partly be explained by the fact that the frequency for early years was so low that either no vectors could be built, or this results in no overlap which likely is due to the very low frequency. The graphs of all words, apart from *nyhet*, starts with several years of zero score for both RBO and frequency. This both lowers the average RBO, and raises the correlation.

3.3 Most similar words

Finally, we observe the most common words for the vectors, as shown in Appendix A. Time constraints did not allow us to do an exhaustive analysis of the vectors, but we can nevertheless get some findings from a brief study. One very evident finding is that the vectors are full of misspellings, caused by misclassification by the OCR engine. These OCR errors have been previously observed by Adesam et al. (2018). A large amount of the most similar words are misclassified instances of the word itself, resulting in very unreliable vectors. We concentrate mainly on OCR errors of the original word, e.g. *mufik* in the list of most similar words for *musik*, but we do not focus on OCR errors of other words found in the lists, e.g. *tretlig* → *treflig* (nice) in the lists of *glad*. We judge that OCR errors of the original word have a much greater impact on the reliability of the vectors, as these words should not appear in the lists. These errors are found at various ranks among the most similar words, but they seem to be more frequent in the top of the lists. This is especially true for words with high frequency, which is reasonable. Larger corpora will let the models find the OCR errors in more similar context as the correct word. Looking at the word *man*, with the by far highest frequency, we see that the top 1 word for almost all years are *nia*. Other words in the top of the lists are *mau*, *inan*, *niau* and *msn*, all most likely OCR errors of the original word *man*. Moving down the rank in the lists, we find less OCR errors of this type. Other words with high frequency, *tidning*, *resa* and *musik*, show similar patterns, although *tidning* and *musik* seems to have more OCR errors consistently throughout the ranks. Focusing on words with lower frequency, we see less OCR errors. One reason could be that the OCR errors were not of high enough frequency to pass the min count boundary of 5. The word *retorik*, for example, with by far the lowest frequency, show almost no OCR errors in the lists at all. For the word *poli-*

tik with a relatively low frequency, we find less OCR errors in the early years when the frequency was lower. Generally, we see most OCR errors in the 19th century. We believe that this is due to various reasons. The earlier years have lower frequency, excluding some of the OCR errors. The 20th century have both moderate frequency and a font that is more similar to modern fonts, allowing the OCR engine to correctly classify these words.

The vectors of Tahmasebi (2018) did not show frequent OCR errors. One reason might be that she set a higher boundary for min count when training the models, so that many of the OCR errors did not find their way into the vectors. One other factor might be that the previous study build their vectors on much smaller corpora, again not allowing the OCR errors to pass the min count boundary. However, since the previous study did not reveal any details, we cannot draw any certain conclusions of these differences.

Finally, we also see examples where the automatic lemmatizer failed to assign a lemma, e.g. plural forms of the words. Another expected finding is that the words *kvinna* and *qvinna* are frequently found in each other's vectors. This is especially true for the second half of the 19th century, when both spellings were used interchangeably.

3.4 Comparison with previous study

The previous study looked at only the top 10 words and used Jaccard as a similarity measure. For this reason we also present the results of the present study in this format, along with the normalized frequency and Spearman correlation. Table 4 shows a summary of the results of the previous study (ver 1) and the present study (ver 2). Looking at Jaccard scores, we see significantly higher scores for the present study, although this is not high enough to represent stable vectors. This is true for both the average and the max score. The average Jaccard score for all words, all years are more than 50% higher for the present study. The representation of *woman* (kvinna/qvinna) differs between the studies, as we separate the different spellings in the present study, while the previous study only had one representation of the word. Taking this into account when calculating the average Jaccard score for the present study, raises the average score to 0.132. We see the biggest difference for the word *glad*, which was an outlier in the previous study with the lowest Jaccard score, while show-

word	avg Jaccard (A)		avg norm freq (B)		correlation (A, B)	
	ver 1	ver 2	ver 1	ver 2	ver 1	ver 2
qvinna	*	0.080	*	52.12	*	0.876
kvinna	0.119	0.050	39.98	73.86	0.665	0.820
politik	0.042	0.040	31.81	78.79	0.472	0.408
nyhet	0.048	0.092	33.02	129.54	0.475	0.038
telefon	0.032	0.018	62.13	142.47	0.461	0.416
telegraf	0.029	0.054	16.59	164.52	0.466	0.657
glad	0.019	0.156	56.24	306.75	0.182	0.739
musik	0.179	0.248	68.68	347.62	0.607	0.617
resa	0.134	0.258	312.00	977.66	0.129	0.396
tidning	0.191	0.163	381.54	980.57	0.327	-0.159
man	0.105	0.223	2084.98	9923.89	0.067	-0.072
Average	0.080	0.126			0.385	0.430

Table 4: Jaccard, normalized frequency, and Spearman correlation for previous and present study. Calculated on top 10 most similar words, for the years 1749-1926. * The different spellings of woman (kvinna and qvinna) were kept separate in the present study, while the previous study only showed one representation.

ing a relatively high score in the present study. Observing the most similar words in Appendix A, we can see that at least part of this could be explained by overlaps of OCR errors of the original word in the present study. This could also partly explain the fact that the highest achieved Jaccard scores are higher in the present study. E.g. looking at the highest overlap found in the present study, *musik* between 1820s and 1830s, we have an overlap of 9 out of the top 10 words in the present study, all of which are OCR errors of the original word. However, the same does not apply to the second highest overlap, *man* between 1880s and 1890s, where half of the 8 overlapped words are true words, and half are OCR errors of the original word. We also see a radical difference in the frequencies between the studies. As mentioned, we believe that this is a result of different preprocessing methods.

Next, looking at correlation, we notice two outliers. *Nyhet* shows a low Jaccard score, a low frequency and a low correlation for the present but not the previous study. Similarly, *glad* shows the same pattern for the previous but not the present study. Reasons for these behaviours are left for future studies. Apart from this, we see similar patterns in the average Spearman scores. One interesting difference is that the behaviour is amplified in the present study, showing more extreme values, both for the higher correlation reaching above 0.80, and the lower scores reaching to negative values. The average Spearman for all words, all

years is higher for the present study. However, when adjusting for the various representations of *woman*, we get a score of 0.389, which is on par with the previous study.

4 Conclusions and future work

Despite using a newer and extended version of the corpus, as well as different methods in the experiments, including different preprocessing techniques, different similarity measures, a higher number of most similar words to compare, and building models of 10-year chunks instead of 1 year, the present study confirmed the findings of the previous study. The results are remarkably similar, and to some extent amplified. Specifically, we see an even stronger correlation for the words with lowest frequency, and an even weaker correlation for the words with highest frequency. These extremes seem to cancel out each other - after adjusting for the various spellings of *woman*, the average correlation between Jaccard and normalized frequency are almost identical between the two studies.

One aim was to examine how the newer version of the KubHist corpus would behave when used on Word2Vec with more data for each model. We conclude that the larger size of the corpora used for the models only solve some of the issues. This is visible for the three most frequent words, with a frequency starting at around 1000 per 1 million tokens, showing almost zero correlation be-

tween similarity score and frequency. This seems to be enough for eliminating frequency as a factor for the stability of the vectors. However, we also noted that there was still a correlation with absolute frequency, especially visible for the early years where models were small. Hence, a high normalized frequency must be accompanied with a corpus of large size in general, to expect stable vectors. The value of using RBO instead of Jaccard as similarity measure is not clear after this study.

Another finding that indicates some improvement from the previous study is that the newer version of KubHist shows more than 50% higher degree of overlap across the years, compared to the older version. Reasons for this is most likely the increased size of the models, resulting in both better predictions by the models, but also a higher frequency of OCR errors. A large fraction of the overlaps were overlaps of OCR errors of the original words. This makes us question if a similarity measure between the most similar words really is a good measure of the stability of the vectors, since a high score could simply mean that there are a lot of identical OCR errors among the most common words across years. Observing the results closer than the numbers presented thus far in this study, we also saw a higher similarity in top 10 than in top 30, which indicates that there are more overlaps in the top ranks of the lists. This, accompanied by the fact that more OCR errors occur in the top of the lists, strengthen our assumption that high similarity scores can often be overlaps of OCR errors, and thus, a similarity measure would be a poor measure of the stability of the vectors.

We judge that the most important factor of the unstable vectors are undoubtedly the presence of frequent OCR errors. For this reason, the main focus of future work should be to improve the OCR quality. One workaround could be to set the min count boundary to a higher value and/or limit the size of the vocabulary when building the Word2Vec models, and thus exclude some of the OCR errors from the models. The findings indicate that the larger models include more OCR errors than the smaller models, and we suspect that this is due to that the OCR errors of the smaller models does not reach the min count boundary. This is also in line with Adesam et al. (2018) who hypothesize that every specific OCR error should not be of high frequency. Nevertheless, more stud-

ies are needed to confirm this hypothesis.

Another way of overcoming the OCR issue is to directly address the errors with post-correction. In this case, we find indications that word embeddings could be a good candidate for identifying the OCR errors, since they naturally appear in the same context as the original words. For this to work optimally, we would want an as big corpus as possible to improve the quality of the model, and preferably with no vocabulary limit nor min count boundary, to make sure that all OCR errors are represented.

References

Yvonne Adesam, Dana Dannélls, and Nina Tahmasebi. 2018. Exploring the quality of the digital historical newspaper archive kubhist. In *Seventh Swedish Language Technology Conference (SLTC), Stockholm, 7-9 November 2018*.

Edgar Altszyler, Mariano Sigman, and Diego Fernández Slezak. 2016. Comparative study of LSA vs word2vec embeddings in small corpora: a case study in dreams database. *CoRR*, abs/1610.01520.

Marco Baroni, Georgiana Dinu, and Germán Kruszewski. 2014. Don't count, predict! a systematic comparison of context-counting vs. context-predicting semantic vectors. volume 1, pages 238–247.

Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. 2016. Enriching word vectors with subword information. *CoRR*, abs/1607.04606.

Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013a. Efficient estimation of word representations in vector space. *CoRR*, abs/1301.3781.

Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013b. Distributed representations of words and phrases and their compositionality. *CoRR*, abs/1310.4546.

Nina Tahmasebi. 2018. A study on word2vec on a historical swedish newspaper corpus. In *CEUR Workshop Proceedings. Vol. 2084. Proceedings of the Digital Humanities in the Nordic Countries 3rd Conference, Helsinki Finland, March 7-9, 2018. Edited by Eetu Mäkelä, Mikko Tolonen, Jouni Tuominen*. University of Helsinki, Faculty of Arts.

William Webber, Alistair Moffat, and Justin Zobel. 2010. A similarity measure for indefinite rankings. *ACM Trans. Inf. Syst.*, 28(4):20:1–20:38.

Shengli Wu and Fabio Crestani. 2003. Methods for ranking information retrieval systems without relevance judgments. pages 811–816.

Appendix

A. Most similar words

glad

1600 : wädret, förborga, innesluta, provi, fährn, befruchta, llaredan, druncknade, hagel, effect, prijset, häftig, ulle, öfwerfallen, tioner, stadha, diup, givrdt, doth, worde, linu, partiel, revier, vkhan, bresch, hotande, kordt, afstä, landtmannen, tilfredz

1700 : blistver, gören, nästko, uplt, nastkommaude, sammar, hwak, gores, lst, kundglordt, upwist, turst, gladi, allsi, bljfwer, pame, upoä, enl, päbegynt, parkie, ährtal, ycket, bliiwer, hasiver, mingen, roij, förmäld, sunn, bank, wsl

1750 : therpä, spä, thea, fada, jig, wrä, tutti, elid, sällstap, hwav, fogel, scholén, bär, herberge, jämn, hvar, borr, stum, öpning, frägd, ockla, heline, nöjsam, ful, dcrpä, bitr, frögd, tempel, havé, lykta

1760 : frögd, munter, swears, blifwe, mlg, qwal, glädja, glädje, hedersman, boch, ach, farwäl, vig, ljuflig, wälkommen, trone, innerligen, tin, gråta, gjör, hjertlig, derför, mången, gudinna, ert, önskan, wälsigne, skapa, emedlartid, lycksalig

1770 : glädjefull, glader, nögda, önskning, innerlig, glädjedag, frögden, glädjetärar, wälsignelsc, glada, ätstildes, omge, hjärteliga, wälsigne, hugkomst, svear, nögder, manhems, välsignelse, fägna, gladi, förnögda, hjerteligen, vase, srögd, wördnads, frögd, tilönstan, wälsig, walgörare

1780 : alada, glada, glava, ulsigt, rniu, glata, kjusa, utfigt, utflgt, ulfigt, ljufa, ljus, lsusa, ljnfa, warina, warma, ntssgt, warnia, gladi, behagelig, dragfri, ntflgt, ntfigt, luisa, tfigt, uisigt, ulstgt, wockra, ijnfa, wackerl

1790 : alada, ulsigt, ljus, ulstgt, utffgt, glava, gladi, uisigt, ulflgt, utstgt, utfigt, glda, utflgt, utssgt, utsigt, ljula, ulsizt, liufa, uiflgt, uistgt, warma, salsrum, wanna, glada, menblerade, nieublerat, unkarlsrum, glata, utstgi, ljuta

1800 : ljus, alada, kjusa, utstgt, gladi, liusa, uifigt, utffgt, uisigt, glava, uiflgt, liuia, ungkarlsrum, utflgt, ulfigt, ulstgt, ljufa, liufa, glda, utfigt, ungkarlsruin, garberobe, zlada, tjusa, gaiderob, utsigr, lmsa, meubierade, ulssgt, llngkarlsrum

1810 : alada, gladi, liusa, utstgt, giada, utfigt, olada, utflöt, kjusa, qlada, ungkarlrum, ofwerrum, glava, utfiat, llusa, ungkarlsrum, miblerade, uifigt, ljus, ulfigt, ljufa, utffgt, utflgt, glavt, glaba, nymälade, qfwerrum, gladr, ungkarlsru, ulflgt

1820 : glava, gladi, treflig, trestiga, ungkarlsrum, ntfigt, liusa, ntsigt, utfigt, ljus, ungkarlsrnm, utflgt, ntflgt, utstgt, tresiiga, ljnsa, garderob, alada, tjusa, ntstgt, ungkarlsruin, ivarina, inöbler, utffgt, glada, ijufa, tfigt, glata, ulfigt, nlsigt

1830 : treflig, trestiga, trcfliga, irefliga, giada, tjusa, alada, ntfigt, trelliga, trefiiga, warina, tresiiga, trcsiiga, ungkarlsrum, ungkarisrum, solsida, ljus, urstgt, ulsigt, utsiat, ivarina, gladi, olada, warma, glava, alkow, nytapetsera, glata, garderob, kjusa

1840 : treflig, treiliga, trelliga, tretliga, trcfliga, trefligä, jjusa, trefiiga, gladi, lrestiga, trefligr, välmöblerade, giada, trcstiga, landllig, nymalade, ofwerrum, trifliga, tresliga, irefliga, ljuft, tretligt, trebiga, instiga, ivarina, trestiga, tresiigt, nyfapetserade, lrefliga, irefiiga

1850 : treflig, gladi, irefliga, trelliga, tresliga, treslligt, trifliga, treiliga, glava, refliga, treffiga, qlada, ljuft, trcfligt, lrefliga, krefliga, tiefliga, trrfliga, hyggeliga, trestiga, glädtig, skrattsalfwa, glavt, tjusa, trcfliga, irelliga, treflizt, ungkarlsvåning, jlada, trefligl

1860 : treflig, munter, gladi, tretligt, glädligt, glava, glädtig, trefliaa, trefiiga, fraterniserande, trellligt, glättig, upptägsmakaren, tresligt, alada, tretiigt, lrefligt, ljuft, tjusa, sällskapsmenniska, treffiga, glnda, tresliga, sällstapsmennista, treniga, qladt, trestiga, högtidsqvällen, wiudskammare, linvad

1870 : glädtig, munter, giada, uppstuppen, mnnter, pratstund, symposium, angonäma, gladi, helsofrämjande, sröjdesull, hjerteglad, borgensköld, välkomstleende, russellsquare, fröjd, glådje, gladt, ungdomslust, toujoura, bröllopsklockornas, angenäm, kärleksvår, barnaflock, ljnft, födslodag, lynno, välkommna, ungkarlslifvet, gladasto

1880 : munter, glädtig, umgängeskarl, sällskapslik, barnasorl, glamm, barndomshågkomst, deglada, bortpratade, gladi, barnaskratt, barnskratt, bägarklang, infallsrik, ungdomslek, glädje, hjärtinnerlig, fröjd, hnrtiga, glädtlgt, hallelujafest, vårflägt, juvenaliska, barnaleende, öppenhjertade, liumör, angenämste, sjukkammaren, förnöjd, gladasto

1890 : munter, glädtig, umgängesmenniska, barnansigten, hjärtinnerligen, öfverrabkad, glädligt, blåsinstrumenternas, solskenslynne, solskensleende, ungkarlsfest, stockholmsgemyt, töresfält, gladi, giädtigt, augustikväll, hjerteglad, prestfruns, gryningstimme, stålsmidd, ljufveligt, marknadshumör, barnaanleten, barnajubel, fröjdfullt, mnnter, umgängsamma, fölungsgnäggning, uppsluppen, småherrskap

1900 : studentdag, uppspelt, gladi, lifsglada, uppsluppen, hjertans, lifsmod, hjärtinnerlig, oförbrännerliga, skämtande, barndomsår, lefnadslustig, solskenslynne, barnansikte, ljufvare, dansen, uppsluppenhet, lefnadsfriskhet, förväntningsfull, skalka, pesth, vinterny, glädjerus, skojfrisk, glädtighet, ritzooms, solbrynta, sällskapsbroder, pratstund, samspråkande

1910 : förnöjd, glädtig, glättighet, operetthumör, lättsam, sprittande, lefnadsfrisk, pojkaktig, språksam, sällskapsbroder, hetmes, överlycklig, oefterhärmlig, solbrynta, berätterskan, gemyt, gladi, stinta, rödkindad, munter, raljeri, skojfrisk, barndomskamrat, ungersven, kacka, guibal, odyssevs, nojs, abelard, sällskapsmänniska

1920 : fryntlig, gläda, sällskapskrets, gamman, nyter, skälmsk, språksam, barnaöga, överlycklig, innerligen, uppsluppen, glättig, glada, hjärtevärme, gladlynt, upprymd, belevad, rosig, håvor, återseende, ungmö, hjärte, upptåg, humör, glädtig, levnadsglad, dlu, förnöjd, fågelkvitter, gudagåva

kvinna

1770 : sällja, underwisnjng, correspondanee, fladig, tilbju, hennr, bcj, eör, rjcnt, ldas, löra, gersta, drjfwa, konimen, milj, sgg, arhete, gisle, twäit, tiäi, ivst, enbast, wzn, tilträsas, exer, fruntinmier, niio, otiu, cwi, ärbete

1780 : prestcrne, hälma, aifigt, erfarenbet, hcdcr, acksa, förelager, ridiculera, matmodren, waktat, funte, slafwen, zsr, osörständig, ortn, wedeibörandes, päfvrdras, nämligast, ftickelighet, ssje, answarlg, fnlkomligen, bsrdt, mucka, fjelsi, tjufaktig, honoii, sjrlfwe, förestrifwen, sontien

1790 : sermi, struka, tiänll, fvflor, kuimig, hflr, lwät, förjlär, slätiöm, lirstär, kammi, löistär, fijka, svsslor, hörkamning, ffrifwi, hutpig, tjfna, svtzlor, svflor, betjän, fruniimmersstögder, kunnia, ffrifroo, lpinna, spölor, gprzket, snäz, fpgior, lwätla

1800 : tilhdra, bewio, gsromil, upsorande, kunjkaper, srsgd, isrsedd, anlnäla, wittsord, upke, ivorc, sticklligdet, tjener, landlhushällningen, dcwi, sticklighei, betcg, bciyg, upfvrande, uvfirande, prastera, ftifroa, knnstaper, ärnä, upszrande,

försedöe, uptörande, bervg, jger, fädon

1810 : elaverspelning, spräkr, handastöjder, underwisving, underirirsa, wetenstaoer, undcrwiste, granläts, nvttige, undcrwista, tilldör, underwisiung, fruntimmersflöjder, öusta, erbluder, ffnfwa, undcrwisnlng, skrisna, wctcnstaper, strjfwa, zranffan, strisiva, cfrcen, söreträdeswis, cchatt, nstaper, undeiwisa, underwisnirg, nderwisa, kärdom

1820 : wijar, vieekonungens, zslamismen, wägarnc, hälsotillständ, consiantinopels, underwisningcns, nnderwisningsanstalt, neapels, äffven, xlats, föranlätna, nodwändigt, npstland, bcredwillig, turkjsta, förswär, äfwerr, systelsalt, sspdd, äinnade, sigart, kramgäng, äsrocn, läkarena, åtmindstone, bchöfdc, röfwarcband, syselsätta, middagsbord

1830 : sedr, afgifra, behöwer, svola, nåo, nåsra, förtä, anwan, sådme, tvärd, dagbladskvntorck, nåjot, anhå, ntigon, afvälta, vaua, raara, iagot, iäier, wärba, negot, detsrmma, landcr, inågot, rdentligt, hwcid, liagon, antinacn, muutliga, aniaga

1840 : llifva, öliigt, reyneries, beliölver, qvarslannande, jäfvadt, ätkomlige, bvilkn, vigligast, kdli, visssrligen, däiaf, tändig, uppmärksamnet, uträlta, sådaii, myoket, skjdd, företeg, bchöswes, canska, öfvcrflödiga, imelleilid, meriniskor, faktisia, riiot, keldst, ritva, hvilkft, musiklära

1850 : ointankan, bedvmma, ändamalek, wälbehöflig, kniiua, modervård, kunita, amhe, förhäfwa, bekymmei, betsföre, barmhcrtighctcn, prejerisystem, bunnil, utmeckla, hösiva, sonidet, månligt, kroppsarbeta, elån, bräcklighcter, roä, rcera, moddockor, inälsinän, sommarvecka, föitjent, siuile, lönandc, inneiranare

1860 : kvast, qivinnan, qvinnnn, gwinnan, romanifolk, qwiiiua, awinnor, qvinna, qvinuan, käklek, qvillnan, nödrar, gwinna, molifere, enkoståndet, qminnan, qwiiiiia, gvssen, qwinuaus, vinnorna, jvinna, ivärnlösa, qwlnnan, owinnau, qwianan, framkrattades, qwrrna, gwinnas, qwinua, sremmandes

1870 : qvinna, qwinnor, qwinna, gwinnor, qwinnan, qwiunor, bildflön, qviunor, fländades, grindvakterskor, qivinna, qwinuor, gwinna, skogsberghs, qinnan, givinnor, qvinnnn, qvält, strur, tjensteklassen, dalkulledrägt, qwinnorna, qwinnvr, qivinnor, kwinnor, qvinnomas, hycklerskor, qvinuans, qyinnan, bildstön

1880 : qvinna, qvinuor, qvinuan, qviunan, qvilina, qviunau, qvinnaus, qviunans, qvinuans, qrinnor, qvinuorna, qvinnana, qvinnau, qvinnorua, qviunorna, vinnorna, qvinuas, karaktärsstark, qvinner, qyinnan, qvinran, qviunors, qvinuors, qwinnor, qrinnorna, självförsörjande, kvinliga, qvinnuns, qvinaor, qviunor

1890 : qvinna, qvilina, qvinuor, qviuna, qvinnau, man, kvinno, qviunans, qrinnor, kvin, qvmnan, qviuuor, kvinua, hon, qvinnorua, qvinuan, qvionor, qviunan, kviuna, whitechapelmördarens, emancipationsvurmen, qvinnaus, stamfordstreet, qvinnnn, qvinoan, raan, qvinuans, borttvinar, qvinnao, bögljndt

1900 : qvinna, man, tiilsam, luuull, llluull, kvin, mäa, kvinno, mäu, självförsörjande, prostituerad, nig, rosträtt, kvinnosakskvinna, mbn, inän, mäll, sextontusen, karlkläder, rösträttsrörelse, ⟨num⟩åringarna, lösaktig, dalamän, kupékamrat, skattbetalande, välklädt, föräidrarne, öker, friarebref, perlande

1910 : självförsörjande, rösträttspetition, väljen, rösträttsmöte, självförsörjande, kvinno, rösträttsförening, fullmyndig, småbruksinstitut, hålahult, länspatienter, sedeslös, fosterlandsälskande, irländska, valbarhet, österåsen, namninsamling, fredssöndag, folkskollärarinneförbund, mäa, landtmannadoter, välfräjdad, målsmanskap, atgiftsfritt, aftonbesök, förvärfsarbete, barnaföderska, mäll, avelsagg, efterlemna

1920 : mäll, välklädd, kvinno, passionerad, neslig, statstjänst, uppskakad, idon, ⟨num⟩årsåldem, manhaftig, sammanbo, frånskild, lion, samaritiska, ofärdig, beklagansvärd, hon, bondflicka, grovsyssla, enke, tvillingsyster, föräldralös, kvin, välskapad, sällskapa, bonddräng, föräldrarne, hemfall, husföreståndarinna, uppriktighet

man

1600 : nia, matt, mauz, mehra, säker, dhct, mana, akman, mchra, meri, höra, medra, mera, mvcket, elliest, nogsamt, restera, befattning, mann, ahr, regn, wij, lagh, wist, nog, inttt, tänka, mani, manskap, wrd

1700 : nia, mauz, mani, nan, ilian, inan, ruan, raan, mau, mall, mangt, ninn, nion, matt, til⟨num⟩, somlige, mait, gemene, mans, egentcligen, derwtd, ehnruwal, änu, oczakows, fyllest, meddelande, mern, cgenteligen, fruchta, sevennes

1750 : nia, masterne, harman, stadd, öfwerhand, siuke, tilhielp, fortplantande, raan, fänge, akman, likwal, hemföra, haydamacker, mennistior, preuststa, blesterade, pandur, behageliga, medta, mott, tulle, flagna, lurendrägare, lyckeligit, archipelago, fäledes, befästad, uptäckt, inan

1760 : nia, inan, svalla, käring, döinma, landthushållare, alttd, lättel, flrifwa, anstucken, darutaf, wiht, otman, wederfäs, bchöfde, päliteliga, zag, krympling, gifmildhet, hwarthän, pöbel, uplystare, utwägen, liftvet, lcke, ofwanföre, werkställa, tancka, saran, beswarad

1770 : inan, nia, miswit, mangs, stvara, bestalt, hjelpt, öswergiswa, urspridt, bliftoit, beklagel, nägoir, dagellgen, flycktiga, yan, gemene, ulwäg, tilgädt, spanjorernes, nimrod, ostyldigaste, näaen, leiwa, hwarthan, spratt, scribent, dälde, ntwag, handtwcrks, mangc

1780 : nia, inan, mau, ilian, niau, lrkwäl, gisia, storvitzirens, förwisto, wikhet, wisihct, flygting, oöfwerwinnerlig, tilbaka⟨num⟩, trowardige, kammerater, fredsförflag, förester, dctzntom, laudonsta, förwirringen, wijjt, bortskicka, wärkeligeu, dcröswcr, imedlerlid, ollena, filan, owiht, öswertygat

1790 : nia, inan, wihhet, wiflhet, fraiissa, lranffe, otman, patriolerne, fangna, ropalister, belägringsstycken, ianitscharer, artilleritrain, bankrut, hothamsta, anfallsplan, rykre, ehouanerne, öfwergifwas, tipposaid, afflär, gisiningar, upsökt, arwee, wisihct, hmaremot, terroristista, helvetissa, underhandlingarnc, buonaparke

1800 : nia, mn, volontairer, härwio, han, cavallri, mau, inan, metodist, engelsmännen, klmde, blegerade, olälighet, armes, speeulation, maltesare, desiukom, ofstcerare, spaniorcrnes, en, fjolla, mcnnissor, kompanien, uppbördssätt, mennistsr, rpsiar, smedan, fbrtjcnar, fränsta, contustoner

1810 : nia, inan, wällustiga, aktman, wisshet, mqn, förflaffa, wägir, msn, ultras, uppflä, päkostande, sörtienar, afhandlingcn, öfwerwägas, minnct, afundsam, destomcra, jaggt, nödsa, ögonwittne, dernnd, ldrig, mckn, tidningsartikel, mychet, öswade, ärfara, ädagaläggas, tilbringa

1820 : nia, inan, ilian, mau, tammeligen, ma, niäns, temmelig, klentrogen, nnderstödja, ättninstone, wederläggningen, opartiff, fosterlandsvän, alswarlig, bryggerirörelsen, swärlige, redcwisning, näjda, fredswilkor, sörtwiflade, niedborgare, mail, redakt, handelsklass, atthan, undbära, bibelversion, sbrst, allvarligen

1830 : nia, mau, raan, mäll, inan, msn, mati, mannea, manntn, han, mäu, fordringsägarena, åliiöjas, kattunstryckare, man⟨num⟩, inans, qwiniia, ni, n, manuen, jag, en, rorsman, mänj, ilian, mannell, krcditore, ramgång, beluta, allmunt

1840 : nia, mäll, ilian, raan, inan, mau, ruan, märi, han, wiflhet, manneu, men, männcn, mändc, ansög, lverkligen, förlovning, wittja, rältstaffens, förstål, inän, öswerdädig, huslund, gumnian, manneii, manncn, hmnes, mati, qwinnaii, officcrarncs

1850 : nia, mäll, inan, ilian, mau, märi, mäu, rnannen, han, en, mait, niännen, mati, hushållsam, inän, fräninande, sörmögen, utcbliftva, randct, ma, sränwaranbe, månnen, men, tlllstä, ahans, jians, wifthet, rättstaffcns, fränrvarande, anspräkölös

1860 : nia, inan, mäll, han, inans, person, märi, maus, men, denna, mau, ock, kunna, ntannen, niän, den, hon, en, hustrnr, elektricerande, mani, öfwerilad, enteda, mäu, mänuen, raan, ingep, äiskwärd, niannens, nannens

1870 : nia, han, inan, raan, mau, men, person, någon, tro, ung, kvinna, hon, denna, kunna, en, tillfallon, nöjd⟨num⟩, medålders, qvjnnan, sotnikow, de, mäll, grindvakterskor, inans, kämpastark, oiåste, blott, porkmans, stimpsons, qvionas

1880 : raan, nia, mau, han, qvinna, ung, men, kvinna, också, kunna, någon, inan, hon, person, självförsörjande, mäll, blott, en, ocb, löregifna, då, tro, hvein, sorn, saniga, gardines, qvmnorna, välbelåtenhet, maus, juxou

1890 : raan, han, kvinna, men, qvinna, också, någon, mau, lörband, ung, dagsbröd, dallmers, begrafningsmin, ailägsnat, prätorii, någonstädea, henneen, walkirks, må⟨num⟩e, konungabörd, idgenting, själfmordskandidaterna, tro, vargpappa, bnstrur, mannnn, mordstankar, en, besknfvas, fremorol

1900 : nia, kvinna, luuull, tiilsam, llluull, emedau, försonadt, butiks, läshufvud, mbn, gnabb, öker, nig, mäu, käana, långtid, obarmhertighet, vagnförare, humbeit, idealmänniska, landkrabba, djgt, förbryterska, torpargumma, lierre, osjälfständigt, förlofningstiden, mlera, ämnessven, slimsa

1910 : nia, men, raan, förokommit, öfvergåfve, vissliet, mau, rösträttskraf, trätgirig, allmänl, åtti, självbedrägeri, motbevisa, svekos, liket, spännt, sägos, omöjligt, kännei, knappast, djärfvas, trollhare, tiilbaka, hoppasi, någlon, tilllåtet, fbå, sonr, öfvare, rösträttsutvidgning

1920 : nia, sällskapa, knepig, manhaftig, välklädd, nedvända, mannell, platonisk, tidningsredaktion, kroppskonstitution, jaktlycka, ibara, iland, näraliggande, gagnefsbonden, lamslagen, valbluff, bli⟨num⟩, folkledare, också⟨num⟩, mian, ålan, isyntes, djet, omvittna, sjukdomstillstånd, pratande, konstighet, pojk, självmordsplan

musik

1750 : läsning, utivägar, ytlig, flärd, folcket, wrckan, äring, stäme, lilsa, hål, sedolärare, upläta, autor, werck, satler, nagt, vswcr, vidga, stri, aigot, ftj, förenlmde, twenoe, bestält, mennistors, widi, stiörd, nerhet, eilet, oan

1760 : uprepade, andacht, skriftenes, wistes, hedniste, önstningar, sansad, förnedring, glädie, els, närwarelfe, löjeligt, hiertans, stickelse, witzte, omgå, ropte, folcket, ätankan, betygelfer, kyrker, prägtigt, uphöga, upbyggeligt, emedlarttd, snabbhet, ätbörder, ängest, häpna, romarena

1770 : musiaue, muffque, musigue, mustque, muftque, siunger, musiquc, musigne, sångstycke, fulstämmig, npfördes, muflque, bocal, instrumentalmusique, upsörd, afsöngs, musique, musiqne, orchester, upfvrd, vocalmusique, sängstycken, mnsiqne, verö, dants, vocal, trompeler, söngos, skädeplatsen, vokal

1780 : mustque, musiqne, musique, musiquc, komponera, mnsique, arierne, mnsigue, muflque, glucks, accompagnerad, componerat, seene, mufigue, obligataria, musigue, musiqven, trompetter, musiqucn, componerad, afsöngos, piccen, musigne, celestini, thurner, trumpet, siunger, soloaria, aircr, spelår

1790 : musiquen, musiguen, musique, gretrp, musiqve, gretry, eomponerad, musiqucn, mustquen, mustguen, mufiquen, duni, buffa, musigue, piccini, mustken, dezedes, devienne, gretr, monsigny, musiqven, mustk, envalltzon, chorer, envallon, mufiguen, grelry, cimarosa, aleyrae, komponera

1800 : musiquen, mufiquen, mufiken, mustken, gretry, daleyrac, daleprac, mchul, daleyrae, mehul, mustquen, musiqucn, gaveaux, oella, champein, piccini, mehus, ballett, mufikcn, grelrp, muflken, gretrp, mufik, divertisiement, muflquen, muftquen, mnsiken, grctry, dalcyrac, divertissement

1810 : mufiken, muflken, mustken, daleyrac, musiquen, daleyrae, mufik, dalayrae, zelindas, dalayrac, musikn, ballett, boyeldieu, berton, oaleyrae, mehul, nieolo, gretry, nicolo, divertisiement, musikcn, oalayrae, berlon, komponeradt, boycldieu, jsouard, cimadoro, califen, daleprac, mufiquen

1820 : mnsik, mustk, mnstk, mufik, mnfik, mnflk, muflk, miisik, mnsiken, mufiken, musikoch, östergrensta, musike, stergrensta, mlisik, must, ostergrcnsta, mnstken, stergrenska, östergren, ostergrensta, lstergrensta, musil, bokbandeln, östergrenffa, östergrinffa, ostergrenska, musi, östergrenssa, mnsike

1830 : mnsik, mustk, mufik, mnfik, miisik, muflk, mnflk, mnstk, mnsiken, musil, muiik, musiken, gläser, mustken, nieolo, sång, mnfiken, crusell, jtalienstan, mnflken, miistk, gaveaur, musikeu, musi, trumpan, dalayrae, balletierne, dalanrac, s⟨num⟩na, brcndler

1840 : mufik, mnsik, mufiken, muflk, mustken, mustk, mnsiken, muik, mosik, inusik, musiken, musikeu, mnfik, carufell, masiken, usiken, musi, mnstk, mnflken, balletlerne, muiken, ansöre, nestrov, inusiken, ufik, airangerad, msik, hensler, sång, musike

1850 : mnsik, mufik, nusik, mustk, mnstk, niusik, mustken, mnsiken, inusik, mnfik, musil, mufiken, musikon, insik, musiken, mnfiken, musike, musi, hartzföreningen, alusiken, tengwallfla, muffk, komponetade, muflken, alusik, pilwitz, muflk, blåsinslrumenter, inusiken, exeqneras

1860 : mnsik, mufik, mnsiken, musil, inusik, innsik, mnfik, nusik, niusik, muflk, insik, mustk, musi, sång, nmsik, mnstk, ninsik, ilauptner, niustk, mufiken, mustken, pomposo, masiken, musikeu, aiusik, musikstycke, mnsike, såug, mnsikcn, bleckinstrnmenter

1870 : mnsik, mnsiken, inusik, insik, ninsik, sång, musikon, innsik, mustk, musiken, musikeu, mnstk, leqocq, musil, niusik, parliturspelning, muaik, mufik, nmsik, lererricr, yanloo, partiturspelning, marionettbalett, muik, masik, osenbach, dansmusik, lesterrier, miisik, mufiken

1880 : mnsik, sång, mnsiken, masik, goodtemplarsextetten, nusik, nmsik, mnslk, enstedtska, muik, musikon, elevkapellet, messingsqvartett, versläsning, muaik, musikj, rnnsik, biåsmusik, inusik, deklammation, harmonisällskapets, goodtemplarkören, militärsextett, musikkör, mässingskvintett, knplettsång, fäugkör, hagelqvistska, innsik, mässingssextett

1890 : mnsik, sång, muik, inusik, messingsmusik, mässingskvintett, berättelseläsning, stråkkapell, fiolmusik, sådg, hornmusik, strängqvartett, violmusik, espelundska, messingssextett, såug, muiik, messingssextet, hagelquistska, messingsextett, gitarrmusik, deklamation, messingssoxtett, esplundska, mnslk, mnsiken, brödrasextett, strängorkestern,

sträkorkester, duettsång
1900 : mnsik, sång, hornmusik, sikkår, deklamation, amatörsextetten, messingssextett, muik, musikkår, stränginstrument, duettsång, stråkkapell, musiksextett, musikföreningen, korsång, berättelseläsning, amatörorkester, mnslk, deklammation, orkestermusik, glasmusik, pianoackompagnemang, såne, musikförening, slutmarsch, jtiusik, kvartettsång, kaffeservering, daus, körsång
1910 : sång, hornmusik, deklamation, musik⟨num⟩, mnsik, messingssextett, musikkår, stråkorkester, mässingssextett, musiksextett, musikband, fiolsolo, mnslk, deklammation, strängmusik, slutmarsch, konsertmusik, glasmusik, mnsikkår, hornmusikkår, homorkester, messingsextett, minik, sikkår, korsång, mualk, pianomusik, hornorkester, nationalorkester, duettsång
1920 : sång, musik⟨num⟩, duettsång, stränginstrument, fiolmusik, hornmusik, kvartettsång, musikkapell, stråkkapell, stråkkapellet, hornmusikkår, strängmusik, deklamation, miusik, gluntsång, konsertavdelning, körsång, folklek, violinmusik, kkår, violinsolo, uppläsning, danslek, missionskören, duett, sikkår, hornorkester, bygdemålshistorier, berättelseläsning, orkestermusik

nyhet
1700 : nachdel, päfölgder, wäldsamheter, nästgräntzande, föröfwade, afböya, päfölgd, conjuncturer, owanlige, undersätarne, ansedt, wasendet, swärigheter, pafölgder, undersatarne, wäsende, uprorsmannerne, genueserne, ehwad, märckwärdigt, förtryck, ansiag, fruckta, änstiönt, förmädt, härrils, näpst, oacktad, wicktiga, förtröstan
1750 : päfölgder, förutse, förlägenhet, bedragare, päfölgd, uphäfwa, tilstyndas, aflöpa, affärd, mitznöje, swärigheter, oförrätt, förfarenhet, undersätarne, handelse, oför, swarighet, twingas, pafölgder, nödwandigt, misinöje, namna, blottställa, wigtigare, göromäl, sivärt, sidane, twiflas, änskönt, betänkt
1760 : nyheter, mcdbragte, medbringa, arschau, wekotidning, tidn, inlämnadt, grammatika, norrkjöpings, weko, ntgifwes, larda, inrikestidningar, recension, sammandrag, insördt, utgiftva, continuation, weckoblad, tilsändt, fortfätning, föregäende, brefwet, tillarlskrona, märkwardiga, ifrätt, ifrdn, tyfta, veckoskrift, wcckoblad
1770 : nyheter, onstantinopel, utländfla, utlckndffe, utländfta, allckanda, dagligk, utlandfla, cvgt, agligl, allcanda, allehaiida, ulländffa, allc, nlyheter, dagligl, allcyanda, allchailda, allehailda, agltgt,
ettersburg, gligl, agligk, dagltgl, agligt, allekanda, dagltgt, utllt, allea, utlckndsta
1780 : adrit, nyheter, nlyheter, ettersburg, yheter, utländfla, vtybetev, utländfte, onstantinopel, dräntsen, nvheter, onstantlnopel, utlckndsta, utländfta, sondon, lref, utländfle, oonbon, eapel, lltländste, npheter, heter, utländsie, utlandste, l⟨num⟩ov, utläiidfta, brefwe, ondo, underrättelser, ttyheter
1790 : nyheter, nvheter, utländfla, yeer, utländfta, utländfle, vtybetev, ondo, utlckndsta, onstantinopel, ssndon, ondsn, utlzndsta, utlandska, nvbeter, utlsndsta, adrit, utlandfte, gctober, grter, yheter, ayland, nlybeter, ttybetev, aris, onöon, sndon, arschau, allekanda, vlybetev
1800 : nyheter, oonbon, utländfla, oonöon, utlckndfla, vtybetev, npheter, sondon, ttyheter, nlyheter, utlzndfla, nvheter, oondon, utlzndsta, stifto, utlandfla, yheter, arls, riyheter, utländffa, etev, ukländsta, utlsndsta, utlsndfla, onstantinopel, utlckndsta, yeer, eondon, ondon, ario
1810 : newyorks, chronicle, rikstidning, statesman, nvheter, beklageliga, hoftidningcn, fägnesam, noniteuren, morning, npheter, dagpapper, nyheter, beobachter, dctailler, lnnehälla, evrter, hiladelphia, officicla, underrättelser, nyherer, utrites, däris, dondon, ansöres, monikeuren, monitcuren, lases, iurnaler, iunehälla
1820 : npheter, ärdagens, nyheter, yheter, rekland, krigsteatern, gärdagcns, nvheter, wtrtkes, onstantinopel, tpanien, konstant, wtrikes, kxyeter, särdagens, xheter, rrigsthearern, nyherer, evrter, urkiet, rtzland, torbritannien, ayheter, sonstantinopel, utrikeö, krigsberättelse, mtriles, ortugak, rrigstheatern, rankrike
1830 : nvheter, nyheterne, polkn, inneböllo, underrättelserne, tiduingar, tidringar, jårdagens, notis, hamburgertidningarne, tidoingar, fiankiike, särdagens, underrältelserne, börsrykte, underrattelserne, pariserbladen, lttrikes, hamburgerbladen, gåidagens, nnheter, tidniogarne, nybeter, nylieter, greifswaldsposten, wtriles, nyheterna, krigstheatern, tidmngar, underlättelser
1840 : nyheterne, notis, nybeter, nvheter, söndagsposten, köpenhamnstidningar, slockholmstidningarne, greifswalderpost, nyheler, hamburgertidningar, mtrikes, ankommna, jägtadt, berlinertidningarne, fetstilsrubrik, gårdagspostens, korrespondentunderrättelser, fnällposten, ertrablad, måndagsposten, fredagsposten, stralsundsposterna, stockliolmstidningarne, underrättelferne, hamburgerlidningar, underrättelfen, heratat, köpenhamnstidningarne, nvhet, stockholmslidningarne
1850 : nvheter, nybeter, nylieter, nyheler, postunderrättelser, riksdagsnotiser, postseriptum, stockbolmstidningarne, galanteriartik, hyheter, postunderrättelserna, nvhet, färstaste, nyheter, nyhctcr, telegrafnyheter, notis, ångbåtsposlen, hamburgcrblad, ångbåtsposten, ångbålspost, stockholmsposter, ängbätsposten, telegrafunderrä, nyheterne, postflriptum, ångbåtspost, nvbeter, nyhct, uyheter
1860 : kristianstadsmadet, mristianstadsbladet, nylieter, nyheler, kristianstadsblade, annonstidnmg, hyare, paketfrimärken, aftontiduingarne, annonstidmng, myare, kristianstadsbladel, nybeter, nnheter, hngnesamma, nvhet, pianoorgeln, ännonstidning, aimonstidning, underrättelfen, korrespondensoch, lyare, ånnonstidning, äyare, teaterwcrlden, annonstidniag, utkonimande, yhet, göteborgstidningarna, silkesfabrik
1870 : nynste, nybeter, anstadsbiadet, annonstiduing, anstadsbladef, anstadsbladel, ansladsbladel, annoastidning, ansladsbladet, anstadsmadet, iyaste, anuonstidning, anstadsblade, amionstidning, ånnonstidning, delgifver, krisliansladsbladel, kristianstatisbladet, annonstidaing, aniionstidning, krislianstadsbladet, uyheter, aimonstidnmg, annonstidiiiig, annonstidniiig, annonstidniag, nyhetor, anaonstidning, landsortskorrespondenser, annonstidniug
1880 : kristiansladsbladel, myasle, myasie, åmonstidning, annonstidniiig, krislianstadsbladet, ånnonstidniiig, ånnonstidning, annonstidnmg, ristlanstadsbladet, styaste, lyasle, nyliet, adsbladet, nybeter, jfyaste, htyaste, annonstidnlng, wexiobliadet, aimonstidning, ristianstadsbladet, krislianstadsbladel, wyasle, tfyaste, amonstidning, nyketer, dyheter, myasfe, innonstidning, nylieter
1890 : nybeter, nylieter, jtfyaste, krislianstadsbladel, aimonstidning, annonstidnmg, wexiöbijadet, amonstidning, krislianstadsbladel, armonstidning, atyaste, norrbottenseuriren, wexiobladbt, krislianstadsbladet, uyheter, jlyaste,

jsyaste, wbxiobladet, wexiöbiiadet, annonstidnitig, wexiöbliadet, jstyast, wexiobiiadet, jyaste, jxtyaste, lyaste, jstyat, ånnonstiduing, annonstidmng, ivyaste

1900 : annonstidnin, annonstidn, annonstidning, upplagra, annonstid, nyhetsoch, nyhe, eel, ffftr, garneringsa, nybets, fftf, hyhets, föl⟨num⟩, vidunderkikaren, zonophonskifvor, jänni, spritstryk, melinochemisetten, sotylwta, frarv, nylieter, fapeterier, sotyltato, sotyltata, oeii, mönsteråstidningen, jultidningslitteraturens, spritstrykjern, faconger

1910 : nyhetar, annonstidnin, ånnonstidning, yför, frälsningssånger, nybeter, utrikestelegram, annonstidning, llmånna, herrhattar, musikapparat, århundra, nyheti, senaate, grammophonägare, wonder, aenaste, sammel, fogapparaten, ljudskifvor, nyhetl, lopparbergs, nyhell, damer, nylieter, modenyhet, senaete, gramofonskifvor, biographs, odeonskifvorna

1920 : annonstidning, ånnonstidning, nybets, ktör, avq, chefredaktör, aktör, nyhets, xo, upplagra, redak, lanstidning, hyhets, avo, varannandagstidning, ansvari, kiellman, boktrycker, läos, redakt, diiponent, auo, bikipolitik, aljö, läns, alio, laris, daktör, ansvaris, stidning

politik

1760 : roiga, naini, hafrea, nrom, numa, waraade, handlan, ngo, utivisar, wettenskap, giordr, ycket, älgt, mifibruk, gisivcs, färlia, hästö, välment, lycktvr, riise, ofj, dmgt, nise, urgifwa, ivägnar, hwiltet, besintelige, befintlig, oljer, alrik

1770 : urgifwit, ärsarenheten, clig, sixtus, auctvr, aung, mensklig, päyrka, jorund, tyt, trowärdighet, sincerus, fortsäkning, ttningar, ffrifwet, dryckenskap, werldcnes, laq, skrifwes, folkrätt, qwickare, mannissja, ostyldige, bewandrade, nästföregäende, föredrag, säleves, taeitus, enfkilta, männiffjan

1780 : statsklokhet, politiquen, rädstag, intolerant, ensidighet, fiendskap, catholffe, herrawälde, sedwänjor, aristoteles, carakter, statsklok, partisk, negrernes, cabal, fanatisme, hällsamhet, omänsklighet, despotist, magternas, mysterium, sedolärare, modesti, medborgarne, nationerne, medhällare, tolcrance, romarnes, förderfweliga, philosophicn

1790 : willfarelser, statskonst, gudsdyrkan, despotisme, wilfarelser, sedcrne, moralite, philosophist, falfla, gudalära, uplpsningen, almannare, statsförfattning, theorier, ätstilnaden, politist, iamnlikhet, nplysta, tolerance, philosophi, lhristne, maxim, uplysningcn, definition, konstord, abstracta, werldeu, folkstags, förstand, wetenstapers

1800 : tvstlands, fosterlandskärlek, rädfrägat, pastioner, rätlwisa, jämnwigt, jämnwigten, misskänna, rätthet, humanitet, öfwertygadt, uprätthälla, mängfald, politiffa, civilisation, kritistt, kritist, pasiioner, sträfwande, politissa, assigter, energi, abstrakt, återhällsamhet, framhärdighet, syftemål, policerade, föregångare, earakter, wailan

1810 : samhällsordning, politist, skandinawista, intercste, statskonst, iluropa, politissa, lagstiftning, rligion, systeme, hierarki, kristeliga, äsigter, förhällanden, systemet, polilista, statshushällningen, halsöns, tidchwarf, europ, tidhwarf, system, forswar, grundsattser, utveckling, grundsats, pradr, luropas, moralitet, regeringssätt

1820 : alliancens, grundsats, diplomatik, statsförbund, allianeens, trovpau, poliriken, statsrätt, älliancens, stabilitet, vignon, frisinnad, politifla, menniskoförstånd, statssystem, rvstlands, kabinetternas, statskonst, system, äsigtcr, congrefien, cabinetterna, kombination, abstracta, commerciella, ryfllands, volitista, toryparti, polilik, litik

1830 : polilik, system, styrelsesystem, handelspolitik, angelägenheterne, temporiserande, politisk, svstem, politiker, odelbarhet, kabinetternas, diplomati, förhållandernas, polirista, magtcr, europcista, kabinetternes, politika, perierska, monarkista, grundsats, monarchien, jesuitism, grundsattserna, ministcrens, oltomaniska, regeringsgrundsatser, statssystem, radikalism, noninterventionsprincipen

1840 : polilik, fredspolitik, temporiserande, metternichska, poliiik, frihetsfiendtliga, fredssystem, frihetshatande, politiken, juliregeringens, oförenbart, isoleringspolitiken, thiersta, diplomati, handelspolitik, poliliken, herrssande, partiväsen, julimonarkiens, giftermålsfrågan, ministerutnämningar, utrikespolitik, guizolska, tidningsorgan, nonintervention, historiskhet, urältadt, antiengelsk, framtidsfråga, regeringsgrundsatser

1850 : polilik, utrikespolitik, illfluga, fredspolitik, familjpolitik, eröfringspolitik, poletik, angelägenheterne, diplomati, familjepolitik, berlinerkabinettets, dansttysta, dynastifla, inrikespolitik, olitik, politisk, förwecklingarnc, förwecklingarnes, agressiva, poliliken, krigsfrågan, zarwäldets, antiryska, dynastiförändring, unionsfrågan, oswenska, unionspolitik, iuteressen, antinationel, prcustcns

1860 : utrikespolitik, poliiik, neapoleonska, politiken, nationalitetsidéerna, neutralitetspolitik, polilik, noninterventionspolitiken, agressiva, regeringspolitik, aggressif, enhetspolitik, utwidgningsplaner, öfwerherradöme, diplomati, krigspolitik, ickeintervention, sjelfmotsägande, frankrikeösterrike, ntrikespolitik, agressiv, statskonst, ickeinblandning, hämndpolitik, tuilerierkabinettet, kejsardömet, toryernas, inrikespolitik, absolutistista, polittk

1870 : utrikespolitik, diplomati, beaconsfieldska, noninterventionsprincipen, polilik, donaumonarkien, fredssträfvandena, noninterventionspolitik, fredspolitik, hämndpolitik, toryregeringen, intressepolitik, temporiserande, oriontaliska, sredspolitik, annekteringspolitik, annexionspolitik, öfvergångsministére, ickeinblandning, poiitik, wienpressen, ministéron, orientpolitik, politik⟨num⟩, krigspolitik, torykabinettet, klorikala, systemskifte, manchesterpartiet, iedarne

1880 : utrikespolitik, pauslavistiska, inrikespolitik, framtidspolitik, politisk, bismarcksk, kaluokys, bulgarvänliga, utrikespo, politiked, kyrkopolitik, dagspolitik, conservativa, politik⟨num⟩, balkanpolitik, utrikespolik, tvåkejsarförbundet, vittnessak, rodomontader, småstaternas, lmfvudorgan, baikanhalfön, berlinkabinettets, regeringskretsarnes, förföljelsepolitik, medelhafspolitik, utvidgningsbegär, personalfråga, oppositionspolitik, videlo

1890 : utrikespolitik, politik⟨num⟩, inrikespolitik, politiked, världspolitik, ryskvänlig, vestlandsposteu, skiljedomsartikel, qvadruppelallians, politikeu, konfliktfråga, venstermånnen, panslavistorganet, omstörtningsmännen, stortysk, unionsnpplösande, utrikespo, berlinkabinettets, rikspolitik, antibismarckska, alliansbegär, politisk, bögerpressen, ultraklerikala, minibtéren, fredsmakt, hufvndorgan, intrigpolitik, rabalderpolitiken, unionskonfiikten

1900 : monarkisk, bondeparti, unionsprogram, venstem, monarkistisk, kannstöperi, waldeckrousseaus, rouviers, vensterpressen, imperialistisk, inrikespolitik, unionsstriden, utrikespolitik, chamberlainska, unionsvänlig, stangska, reformpolitik, fredsvänlig, dynastisk, svenskfientlig, franskrysk, nationalism, partiförhållandena, kraterna, frikonservativ, lövlands, rösträttsrörelse, tullunion, oktobristerna, tronkandidatur

1910 : hufvudorgan, hufvudorganet, huvudorgan, regeringsfråga, ententevänlig, rikspolitik, landsfördärfliga, moderatliberal, doktrin, utrikespolitik, nyheterspressen, aggressiv, krigspolitik, stockhtidn, opatriotisk, fredsdiskussionen, konstitutio, nyorientering, regeringspressen, fredspolitik, köpenhamnstidningen, ytterlighetsman, författningsprogram, flottpolitik,

alltysk, inrikespolitik, fredsvänlig, ententepolitik, regeringspolitik, unionsbrott
1920 : klasspolitik, enhetsfront, reformarbete, valupprop, systemskifte, klassparti, radikalism, fredsvilja, förbudsparti, programpunkt, socialdemokrati, folkstämning, kapitalistisk, klassintresse, socialism, omstörtning, framstegspolitik, samlingsregering, valseger, grundåskådning, samhällsutveckling, arbetarmassa, försvarsvän, partiledare, reformpolitik, socia, kapitalism, reaktionär, fackföreningsledare, falang

qvinna

1780 : välkommen, enfliga, beröfva, arete, ålyn, försmägtande, förtviflad, ållkar, vålluft, fmakar, hvilkens, wcrldars, våda, fridiam, vånner, öfverhopa, huflrur, fökte, ewinnerlig, härndén, wräka, hålt, hjettat, wäscnd, sörger, prifar, öfwerger, förvåna, åfyn, bedröfwelser
1790 : poikar, grätcr, wansinnig, ssrik, ffriker, monga, momus, rysiarnc, gräseligt, halfnakna, tilbcdjare, ftukdomar, dumhufwun, räkak, andeliga, ffvgd, stenblind, qwinnors, wistke, torfrige, stngga, qivinna, grähärige, advocaren, behaglighcter, grönländare, bedröfwas, äträr, wanstälde, rilhörde
1800 : mennifkor, ssrister, negerstafwarne, bondaktig, rdfware, kärleken, gillstuga, styddsängel, sullkomlighet, talisman, magta, bäqge, bondpraktika, abcllina, tidebwarf, löfning, barndomen, jndiansta, landprästen, aturens, physiika, jordissa, euripides, mannistjan, undcrjordista, nvdwändiga, wälgöra, romantifl, eklog, romantist
1810 : nakenhet, wekliga, qwäfda, obeständighet, rofgirig, slagtoffer, fjettra, melankoli, sörtwiflan, utmärglad, älstat, klagorop, behagsjuk, qwäfde, halfnakna, förförisk, wansinnige, qminnor, instinkt, uppwackte, uflingen, tillbedjare, willade, swartsjukan, wällusten, självfull, äterger, smälek, sängarn, flyktig
1820 : förtviflan, beslöja, förlwiflan, menskor, känslolös, förkrostad, jämmer, föriagade, uppklättra, wrär, blekhet, återfann, qwidande, hvarann, förtviflad, svärma, lifmet, svärmeri, hwilkak, ohyfsad, gränad, misstänksam, strämda, warclse, nymf, strälarne, skrämsel, flagtoffer, sysslolöshet, glädjas
1830 : menniska, lösaktig, bönpall, männernas, qvinnlig, qvinnau, männerna, kurtisör, följeslagerska, vinnor, varelse, veridens, sammanvexta, demonisk, korsikansk, misantropi, följeslagerskor, martyrdöd, sjelfsvåldig, fullvexta, veild, menniskohjerta, veildens, öfvergifua, meaniskor, libertin, chevaleresk, skyddsande, barnålder, medvetslös
1840 : qviunan, varelse, qvilina, qvirina, menniska, qvionan, qrinnan, qvinuan, flickau, jvinnan, ålderstegna, föiäldrarne, skridskoåkerskan, nervsvag, jvinnans, qvinnsn, fyratioårig, qviiman, qvinnoma, vältänkt, svartögda, ledsagerska, qvinnau, sjde, familjemoder, älskansvärd, audrichons, attigdom, qvalfullt, qvinnorne
1850 : qvilina, menisknr, qvlnnas, qviunan, hvem, giftasvuxen, virman, mertoun, nogarola, tillbedjansvärd, qvinuan, ilon, hjertevarma, qvinnor, qvinnaa, qviunans, qvinno, menniska, nåväl, ehrenstedt, qvlnnan, kurtisan, qvinnaus, henries, gaslons, värdshusfrun, squaw, tvillingsyster, qvinnoma, domitii
1860 : qvin, qvilina, qvinnau, qvimian, qvinno, varelse, qvinlig, qviunan, qvinnor, ilon, qvinuans, qvinnnn, qvinnaus, jvinnor, qvinuan, giftvuxna, afvänjdt, enkoståndet, likväl, tveka, ginter, hvem, qviiiinan, jvinna, qwinnor, lisgards, qvinuau, qvinnorua, ilståndet, förbättringshus
1870 : kvinna, qvinnau, qvinuan, qvilina, qvinnaus, qvimian, qviunans, qviunan, qvinner, qviunans, grindvakterskor, qviunor, qyinnan, qviunors, varelse, qvinnomas, korsdragerska, qvinnnn, litthauisk, qvinnoma, qvinuornas, fyrtiårig, slynåren, fördertvade, qvin, qvinuor, qvinnaa, qvinnan, raan, qvionans
1880 : kvinna, qvinuor, qvinuan, qviunan, qrinnor, qvinnaus, qviunau, qvinner, qvilina, qviunans, qvinuans, qviunor, qviunorna, qvinnau, qvin, qvinnorua, vinnorna, qvinnana, qvinran, hon, man, qviunas, qvinuorna, qvinuas, qviunors, qwinnor, flicka, qvinaor, qvim, qvinuors
1890 : kvinna, qvinuor, qvilina, qrinnor, qviuuor, qviuna, man, qviunor, qvinnau, qvinnorua, qvianor, qvionor, ung, förfäkterska, qviunans, qvinaor, qviunan, qvinuan, stamfordstreet, kvinno, whitechapelmördarens, aubinska, kvin, uppfostringsomsorgerna, qvmnan, qvinnao, qwinnor, borttvinar, qviunors, hon
1900 : kvinna, jemnåriga, logerande, dotterlig, starkenwall, kupékamrat, djelma, spanjorska, lastfull, mannell, slafvinna, förbryterska, hjertevän, mazie, hubling, tillbedt, käraste, béthune, tänksam, holidays, läsarepresten, kärleksdrucken, michselsen, backantinna, fosterfar, henes, gatflicka, vanskapt, kellners, afguda

resa

1600 : anträda, versailles, reefa, reesa, recsa, affreester, begaf, rcesa, hälland, holland, mönstra, mändag, härifrån, mankera, affärda, braband, bcgifwa, reesan, förwäntes, dannemarck, begiswa, recste, windsor, upbryta, engelland, nungen, preusen, grefwcn, afhämta, afskicka
1700 : resi, reson, afresa, återresa, afrest, afreste, hemresa, reia, nesa, asreste, reesa, rosan, återresan, esan, ditresa, begiswa, reesan, aterresa, återkomma, bcgifwa, resek, begisiva, begifwasig, begtfwa, begifwa, degifwa, afrcste, begifwer, hwarisrän, carlsbad
1750 : afresa, reson, esan, risan, resi, rosan, återresa, återresan, återrefan, mistande, anträda, lustresa, reia, afrest, passera, reftin, hwarifrån, begiswa, norrom, fredensborg, avträda, rese, aterresan, jägersborg, derifrån, refan, begifwa, ärnande, arnar, begifwer
1760 : resi, anträda, afresa, återresan, reson, afrest, återrcsan, afreste, återresa, begifwa, begiswa, dannemark, lustslott, anträdt, reftin, rosan, fortsätta, landwagen, incognito, philipsruhe, förwantas, begifwit, afresi, danmark, bortresa, kammerat, laxenburg, achen, zelle, begiftva
1770 : reia, refor, rosan, rcsa, härifrän, resi, refa, resek, rejer, rese, resesällskap, reftin, kammerat, resesällffap, sjöresa, resesällssap, afresa, ditresa, resekamrat, rcfa, återresa, resande, asrest, compiegne, anträda, reson, rcscr, kamerat, orienten, härisrän
1780 : refa, reia, rcsa, relä, refor, resek, reier, refer, rese, återresa, förebud, rcsor, återresa, resi, reson, rrfa, resekamrat, rosan, återresan, rejer, resesällstap, lese, tvskland, rcfa, ditresa, älerresan, postwägen, resesälstap, härifrä, sörnäm
1790 : refa, refor, reia, rcfa, reson, numedien, marilis, madagasear, resi, poiret, japan, maritis, kamtschatka, refer, poirer, thina, resek, rcsa, borchs, cores, återresan, relä, pern, madagascar, postwägen, refan, malra, mears, mariris, rese
1800 : refa, rcsa, reia, refor, resek, reier, nesa, förebud, rcfa, cizico, återresan, rsa, kesa, återrcsan, reson, relä, återrefan, återresa, landtresa, munthers, relor, resr, rescr, resi, krimm, återresa, rcsor, resesälljkap, brydones, resesällffap

1810 : refa, refor, rcsa, förebud, rsor, rese, rsa, reia, rcsor, esor, reftr, reier, resek, återresan, rcser, kesa, grekeland, refer, relor, rejer, rcfa, sjötz, resor, reson, zörebud, sörebud, rcsan, rrsa, luchs, buchs

1820 : refa, refor, reia, rcsa, ressr, rcfa, rescr, relä, resek, retor, rcsor, återresa, reier, reson, förebud, rsor, nesor, refer, reftr, nesa, reior, reler, rejor, moraans, hemresa, restr, äsrvenryr, morgaus, resor, tnorgans

1830 : återresa, färd, hemresa, rosan, resor, reia, refor, kesa, reson, reior, rasan, refa, resesällskap, rcsa, retor, återfärd, återresan, älerresan, landvågen, återresa, konstresa, reftin, exkursion, reja, manfredonia, ätcrresan, rese, resek, rsor, hitresa

1840 : refa, rosan, refor, reftin, rcsa, rese, färd, återresa, refan, återrcsan, reson, fotresa, resau, landtresa, äterrefa, äterrcsa, rcfa, ätcrresan, ätcrrcsan, hitresa, rcsan, begifivit, älerrefan, regiiliera, tesor, reia, återresa, öfverreste, rcfan, besan

1850 : refa, färd, rosan, refor, rcfan, resau, ätcrresan, rcfa, rcsa, återresa, ätcrrcsan, rcsan, reftin, reia, rese, aterrefan, aterresa, hemresa, ätersärden, rasan, resao, derifrön, resatill, reson, ätcrresa, thules, ätcrrcsa, återresa, refau, bemresan

1860 : rosan, refa, rese, resau, hemresa, reson, refan, färd, rasan, refor, återresa, rese, reia, återresau, resai, besan, öfwcrresa, anträda, återrefa, resan, ätcrresan, wistådnet, älerresa, begifwcr, ltesan, återrefan, refau, konstnärsresa, återresan, reian, rejan

1870 : rosan, färd, rasan, resande, hemresa, anträda, afresa, reson, resaus, lesan, färdén, anlrädes, resau, resaa, rese, färdon, resan, auträdes, afreete, fäiden, lcrrdal, forlsälta, pähelsa, qvikkjokk, lauritsala, emigrantfärd, återfärd, refa, bournemouth, pähelsa

1880 : rosan, rasan, resau, färd, hemresa, afresa, resto, resande, anträda, massresor, färdén, vistelse, färdon, rese, vesa, rekreationsfärd, reia, rekreationstur, resan, re⟨num⟩an, amerlka, bournomouth, homresan, resao, eesan, återvända, aftärsresa, återfärd, resaa, afresnn

1890 : rosan, rasan, resaa, tesan, re⟨num⟩an, resau, lesan, färd, hemresa, ressn, rosaus, vapenöfningsorten, resto, kesan, iesan, kurirbåten, resao, auträdes, reean, tågtabell, kontinentalresa, anträda, kerija, färdon, åustralien, emigrantbåt, maristuen, hemresau, resan, pariståget

1900 : rosan, färd, turistfärd, hemresa, besan, anträdt, måndagsångare, enoart, anträda, endart, abazzia, hävre, kalmarsundsbåtarne, lördagsäftnar, srinagar, londonbron, sommarresa, bijett, reisa, nppehåll, götebnewyork, återfärd, rekreationsresa, järnvägsresa, skinnarbol, öfverfärden, hellie, bremerhafen, julbesök, wishys

1910 : hemresa, rosan, hitresa, återfärd, öfverresa, resn, amsrika, rasan, hellenic, kanalväg, amarika, åmerika, wilhelmshaven, rösan, kanalvägen, salongsplats, afgiok, nedresa, vesa, nordjylland, färd, rekreationsresa, tesa, amsrlka, trälieborg, hvetelast, jättefartyg, besan, semesterresa, amarlka

1920 : hemresa, anträdt, överresa, återfärd, bilfärd, landstiga, southampton, affärsresande, sovvagn, rundresa, embarkera, sommarresa, vancouver, turist, amerikaresa, reskassa, färdén, resa⟨num⟩, metropol, besa, sofiero, rekreera, dagsresa, capri, resande, hemfärd, utlandsresa, rivieran, månadslång, röksalong

retorik

1820 : inbill, chrisiendomens, wadjar, öfvermod, tillwiter, ffolen, dynamisk, känslan, hänad, försvarsmedel, riksdagstalare, aerningar, werloens, öfverskåda, känsto, mennilkan, bcrömwärda, totalitet, tillbakahålla, dehwarf, wändberg, försläs, syft, känstolös, grundform, wälbekanta, omdömeskraft, öfvertänka, ulwecklingen, afdessa

1830 : prosodi, etymologisk, rhetoriken, esthetik, rhetorik, poetik, logika, mathcmatiken, konstlära, physiologien, språkstudium, sesthetiska, nationalekonomi, interpunktion, philologer, ästhetiska, skildiing, psychisk, mylhologien, talmud, schellingska, astrologisk, metafysik, religionsfilosofi, encyklopedi, fenomenerna, lärosattser, mytherna, romantikerne, svårfattlig

1840 : rhetorik, heraldik, språkvetenskap, mongoliska, linguistik, aristotelis, mytologisk, skolastisk, lexikografisk, ortografi, kammarlärd, ovetenskaplig, linguistisk, prosodi, verskonst, begelska, flygligt, kommentator, anthologien, hitorien, semitisk, öfverllödar, diktart, förelrädesvis, vokabel, philosopherna, philosofi, rhetorisk, hufvudstudium, philologisk

1850 : floskel, skolastisk, rhetorik, skolastik, obildning, bygnadskonsten, platonism, universalitet, spiritualistisk, skolmässig, rhetoriska, scholastisk, målarekonst, rhetorikens, fouriers, qvinlighetens, folkanden, genomtrångd, förkroppsligande, predominera, ascetiskt, himmelsstormande, epikureisk, exoterisk, grundelementer, historieforskning, absoluti, tånkare, hårklyfveri, nalursanning

1860 : metrik, skolastisk, retorisk, skaldespråk, formfulländadt, formrik, dionysoskulten, skolastiker, skaldeslag, äslhetiska, platonism, sensualism, naturalistisk, skönhetslära, världsåskådning, empirism, förfrns, pindari, romarespråket, aristotelisk, bildprakt, tankereda, bildningsgång, eklektisk, dialogform, sokratisk, tankedjup, hufvudkaraktererna, häfdateckningen, utvecklingsskede

1870 : diktlära, språkvändning, färgglöd, bildrikedom, prosatör, detaljreglor, talarkonst, trimeter, satsbyggnad, formrikedom, didaktisk, estelik, femfotad, renaissansens, rhytm, didaktik, lågtyska, fllosofl, sagoform, literatu, reflexionspoesi, högspråk, klassikerne, jambisk, bildprakt, principio, völuspå, aeneiden, konståskådning, dialektik

1880 : bildprakt, ciceronianska, bildrikedom, svulstighet, rhetoriskt, formrik, lärdomsapparat, schellingska, satsfogning, formbildande, polyfoni, klassiflt, retorisk, ordval, framställningsform, schematism, ordrot, hegelianismen, språkkonst, schematisera, skoldressyren, sofisterne, världsåsikt, rhetoriken, ordprål, oraloriska, romarspråk, dialektik, pretiös, formsäker

1890 : rhetorik, dialektik, utformning, metafor, hyperbolisk, satsbildning, litteraturriktning, skolastisk, kompositionssätt, hegelsk, metafysiker, verskonst, rljud, rabbinsk, högestetiska, lärdomsprål, aristotelisk, skrifart, halfbildning, boileaus, utfrågande, satsbyggnad, anakolut, ordstam, argumenterande, otilllåtligt, pythagoras, världsförklaring, boklärdom, syntetism

1900 : matik, bokhålleriafdelning, läsämne, öfningsämnena, stndenter, kälslära, slöjdafdelning, handafärdighet, realgymnasium, afseenda, afslutningskurs, kunskapsämne, husmoderskur, hufvudämnen, teckningssal, timlärare, undervisningsvana, årsklas, veckotimme, välskrifniug, handelslära, kunskapsgren, skrifprof, uppförandebetyg, liälsolära, öfningsämnen, välskrifnlng, euklides, väiskrifning, militäröfningarne

1910 : kolartro, livsuppfattning, tidsriktning, humani, själfkritik, intolerans, samhällsförbättrare, hypotetisk, altruism, världslighet, kyrkolära, oförblommeradt, jordsocialismen, folklighet, lynnesdrag, ryggradslös, jesuitism,

statsmannamässig, lärobyggnad, återhåll, utlåtelse, vesterländsk, utopist, intuitiv, formuleringskonst, religionslös, erotisk, grovkornig, pjåsk, hedersnamn

telefon

1870 : megafon, mikrofon, fonograf, edison, bellska, taltelegrafen, mikrotasimetern, telefonapparat, elektrist, elektrisk, sonografen, telofon, edisous, elektrifla, telefonbörs, induktionsströmmar, dubbeltelegrafering, edisons, edisonska, elektromotograf, edifons, menlopark, ringapparater, upfinning, elektromagnetisk, tasimetern, snabbskrifvande, apparat, telefonera, sonograsens

1880 : telofon, telef, tolefon, bell, teleton, telefou, allm, teleion, telelon, allni, belltelefon, allin, telof, teiefon, alim, lefon, telefonnr, ailm, telefonor, telefoa, telefonmr, eelltelefon, tolofon, elefon, allm⟨num⟩, telcfonnr, telefonnir, telefun, tolefonor, elofon

1890 : telefou, telelon, telofon, rikstelefon, teleton, biks, teiefon, lefon, iliks, telelou, ltikstelefon, rlkstelefon, tolefon, elefon, kiks, telefon⟨num⟩, kikstelefon, telsfon, llikstelefon, itiks, itikstelefon, statsteleffraf, bikstelefon, rikstelefoa, telcfou, telefoa, riksteleion, liikstelefon, ribstelefon, riksteiefon

1900 : telegraf, nfvetex, nvetex, nketex, nhete, ételegram, nhetex, tvketex, netev, nhetev, ntvetev, senasu, ctelegram, nketev, tvveux, nvete, nvetev, nfvetev, nvvetev, telofon, settasu, ttelegram, tuetex, jyfed, åskledareanläggningar, setvasu, qeflo, celegram, ntvetex, sfelegram

1910 : telofon, kedaktionen, bedaktionen, återförsäljningeställen, örsälj, återförsäljningsstäilen, teleton, telefonar, bedaktoren, återföreäljningsställen, återförsäljningsstftllen, terförsäljningsställen, telefoaer, tolofon, bedaktören, återförsäljningsställen, redaktionsaekr, redaktor, boatad, återförsäljningsstäden, ftterförsäljningsställen, kedaktionssekr, erförsäljningsställen, jnin, teleion, telefnn, återförsftljningsställen, återförsäljningsställon, ningsställen, redaktionseekr

1920 : telofon, namnanrop, ⟨num⟩igv, frnktavd, sige, fraklavd, andréns, passagerarenvd, telef, packhusplalsen, amerikaliniens, packliusplatsen, norlings, packhusplatscn, teleton, passavd, fraktavd, amcrikaliniens, ⟨num⟩ti, walling, tel, vagnverkstad, packhusplatsen, rikstei, drottningtorget, automobiler, itikstel, tolefon, rikslel, motors

telegraf

1800 : telegraph, telegrafdepesch, sheernest, telegra, decres, dovre, expres, telegraphcn, stafette, curiren, basilico, staffett, hoftidninge, dunkirchcn, staffetter, statsbadet, hvftidningen, depesche, statsbäd, ratifikationerne, hitstickade, rönigsberg, signildsstärs, krigsstupen, bculogne, vlicssingen, hirkom, hofkidningen, coimbra, breskens

1810 : louvre, telegraphen, psris, triumfbäge, monit, timor, franksurt, eruz, luiro, illumineradt, prestburg, antigue, merico, undcrrättelserne, porträter, nubien, elyseista, factoriet, statucr, peris, devecher, stafett, portfölj, magdcburg, perfien, riemen, costum, tugall, ditwäntades, trafalgar

1820 : telegra, consols, medförande, eioile, inlupil, bombai, petersburgffa, österr, tamar, parisi, telegraphen, portobello, utspriddc, gekista, vengeur, helleniska, följander, couriern, zirn, franfyst, onkommo, afgången, ögonwittne, bureaucr, porlorico, wallachlet, osservatore, cscadre, telegrapher, vrivatbref

1830 : telegra, estafette, telegrafunderrättelser, telegr, telegrafering, rapporr, telegrafi, telegrafista, telegrafdepcsch, telegrafdepecher, ingäkt, crocodile, depeche, gardagsposten, spahieh, telegrafisk, loulon, telegrafdepeschen, följander, telegrafist, stafett, underlättelser, bayonoe, vayonne, inlnpna, telegrafrapporter, thåre, falmoukh, berchem, telegrafunderrättelse

1840 : rapporter, telegra, telegrafi, elektromagnetisk, signaleradt, afscgling, gosport, bängs, bång, telcgrafrapport, telegraflinea, elektri, slutstation, elektromagnetist, estafetter, telegrafisk, estaffett, alaröy, telegrafunderrättelsen, elektrisk, telegrafering, ropport, telegraflinien, telegraphen, happort, telegrafdepech, telegrafdepescherna, telegraflinje, medföide, telegraflinier

1850 : elektrisk, telegra, submarinska, clektrista, telegras, elektrist, elcktrista, telegrafisk, elcktriffa, legraf, tele, telegrafförbindelse, clcktrista, telegrafera, telegrafe, hamburgcrtidningar, epescher, telegraflinje, telegrafförbindelsen, telegrafkompaniet, elektiista, telegraflinien, lelegraf, telegrafi, elektrifla, nstadstidningen, eleklrist, telcgrammer, underwattens, grafverket

1860 : legrafen, tele, telegra, telegras, legraf, teleg, telegrafisk, telegrafeu, kulapparat, grammerna, telegrafkabelns, ilttb, grafön, buffeloxarna, manftet, kabeln, iltfb, telegramin, danskrysk, telegrafen⟨num⟩, tclegramm, telegi, ystadstidningen, storkliol, gratisbefordring, graferingen, longitudsbestämningar, grafkabeln, telrgrammer, caselliska

1870 : telegratering, telegrat, telegra, telegralering, uphmds, telegrasen, uplnnds, telegraferiug, nosto, uplmds, telegras, telcgrafering, ilgodsmagasiu, uphnds, iplamls, statstelegrafstationerna, telegr, tele, tullwerken, legraf, aprsept, upfonds, stoskholms, skogsstaterna, tullverken, tjenstemeddelande, uplamls, iigodsmagasin, telegrafiinier, ilgodsmagastn

1880 : legraf, telegrat, telegra, statstelegrafstationeu, vägstelegrafen, tele, järnvägstelegraf, ägames, llplumls, telegral, telegrafoch, annorledes⟨num⟩, telegiaf, telefonstationshuset, fonstation, statstelcgrafstationen, llplamls, oscarshaninsbanken, elegraf, telegr, fonnät, telefonkrok, grafstation, uplamls, mjärnväg, jiirnvägstelegrafen, hernösandsbank, ielegraf, iplumls, fonledningar

1890 : järnvägsteiegrafen, telegra, statstelegrafstationen, stelegraf, järnvägstelegraf, statsteleffraf, telefonstation, statsteleg, slatstelegrafstationen, telegral, telefon, grafstation, telegrai, annorledee, vägstelegrafen, statstelegraf, mjärnvägstelegrafen, bankinrä, hemösandsbank, telefontaxa, telefonledningar, telefcnföreningens, tele, brefkorrespondenter, heruösandsbank, statstelegr, töyheter, nyhetei, cgenom, hufvudstaaen

1900 : nfvetex, nketex, nhetev, telefon, nketev, nvetex, ntvetev, tvveux, ttelegram, tvketex, nhete, nkete, nhetex, ételegram, nvetev, nvvetev, ctelegram, sfelegram, jelegram, jyfed, nketet, netev, nfvetev, tfelegram, nvete, ntvetex, celegram, settasu, tuetex, setvasu

1910 : trådlös, telegram, jrted, jtfed, tetefan, jtftd, telefea, jelegram, telefon, legraf, jyled, trådlos, trådläs, jrfed, egraf, jrfea, tetegraf, jfltd, ilgodsexp, jfled, jrtea, tetefen, jfftd, felegram, telegrafering, telefonförbindelse, marconis, telegraflinje, egrom, telefonautomat

1920 : telegrafstation, trådlös, telefonmannaförbundet, telegrafverk, telegra, stationsbyggnad, radiostation, paketpost, trupptransport, vattenfallsverk, sistajiytt, dusseldorf, postkupé, handelsdelegation, montag, moskwa, telefonväsende,

renter, postoch, järnvägsverkstad, osaka, morning, kaukasus, slutstation, calcutta, tändsticksindustri, tyskrysk, basel, linjeskepp, telegrafoch

tidning
1700 : förmala, bekräftelse, tiender, berättelse, tijdender, tidender, ridning, mcdbragt, brefiven, tijdningar, medbrackt, dref, breswen, inlupne, förmäla, tidnin, inlupit, derbent, förmales, inlupen, mcdbrackt, hirkommen, enspannaren, striswes, meddrackt, medbringa, hitkommen, obehagelige, hitbrackt, hitkomme
1750 : inlupen, obehageliga, ridning, bekräftelse, inlupne, enspännare, tidnin, brefwen, berättelse, persien, medbringa, förmäla, brefwcn, förnimma, efterrättelse, förmala, sörnimmes, hitkommen, courier, innehälla, depecher, förnimmes, inlupcn, omständelig, express, behageliga, obchageliga, affardad, ftdnaste, hitkommcn
1760 : tidningar, weko, tidn, inrikestidningen, bihang, inrikes, wcko, larlslrona, nyheter, posttidning, insördt, karlskrona, wekotidning, inrikestidningar, nyhet, tillarlslrona, inrikestidningarne, weckobladet, wekotidningen, artik, deko, posteller, ifrätt, tillarlskrona, ukgifwes, zorrköpings, wrko, nyheterna, posttidningarne, gazett
1770 : tidningar, rivning, tidn, tidninaar, tidni, blader, tidninga, tidnin, ridning, zidningar, blad, utgiswas, utgifwas, utgiiwas, dningar, tidnlngar, idningar, dagli, dagltgk, urgifwas, tidvingar, ulgifwas, blada, dagligl, vtgifwas, tidninzar, allehaiida, tidningame, dagltgl, allehailda
1780 : tidningar, tidntngar, tidnlngar, tidniugar, dageiiga, jnrikestidningarne, allehan, tidni, tiduingar, tldningar, ttdningar, dagcltga, geli, dagltgt, ostsdag, dagltgl, allehand, idn, dagltgk, tidn, dagclig, ottsdag, gligl, ttdntngar, tidninaar, agligt, ztdntngar, agltgl, alleha, rivning
1790 : tidningar, tidnin, allekanda, allchauda, dagligk, allchaiida, posttidning, lidning, daglig, tidningarnc, dagligl, dagltgk, onstantinopel, nyhet, tidni, tidiiingarne, allcöanda, tldningar, agligl, dagltgl, dageltga, dagllgt, agligt, allchanda, alleöanda, alletzanda, dagli, posttidningen, lidningen, dagcl
1800 : tidningar, tidnin, tidni, tidningarnc, lidningar, infördt, tidninga, dagligk, dagligl, lidningarne, tidninaen, innnew, dagltgl, tidnmgarne, dagltgk, dagcltga, tidntngar, dageltga, drter, agligk, dagcllga, tidnlngarne, lidningarnc, ttdnlngar, tidnlngar, idningar, tldningar, allckanda, elzer, allchaiida
1810 : tidningar, tidningsblad, tidnin, dagblad, lidningarne, tidninga, införa, infördt, tidni, inforas, tidniugarne, tidningarne, chronicle, tidnlngarne, dningarne, tldningarne, tidningspapper, noniteuren, newyorks, tidningarnc, morning, statesman, tidntngar, journal, tioningar, monitör, dondon, tidninaar, tidningorne, tidningarn
1820 : tidningar, tidnin, tidn, ärdagens, tidni, tidningarnc, tidningsblad, infördt, lidningar, stockholmsposten, aftonbladet, jnrikestidninge, tiduingarne, granflaren, tidnina, tidningame, tivningarne, torsdsgm, börstidningen, aftonblad, titningarne, tidninqarne, jurnal, tidnmgarne, lidningarne, inforas, tidnlngarne, tidningärne, infird, tidninaen
1830 : tidninga, tiduingarne, lidningen, correspondentcn, stadstidning, tidningar, tidninp, correspondcnten, aftontidning, tidningen, tidni, inrikestidningens, tidningai, aftontidningen, landsortstidning, statstidning, tidningame, granstaren, lidning, tidnin, idningen, provinsblad, stockholmsbladet, provinsbladen, jurnal, statstidnin, morgontidning, kannstöpa, ⟨num⟩lllehanda, gåidagens
1840 : tidningar, tidningarnc, tidiiingen, tidnidgar, lidningarne, lidningen, tidnin, tidningarue, tiduing, tiduingarne, morgonposten, corrcspondenten, tidniugar, tidniogarne, ertrablad, tidningen, tidninge, tidninga, aftonbladet, extrablad, länstidning, norrlandspostcn, idningarne, tidningai, tldningarne, redaetionen, landsortstidningarne, tidningame, iönköpingsbladet, ⟨num⟩lllehanda
1850 : tidninge, lidningen, lidning, tiduingen, tidniugen, tidiiing, tidnidg, tidnin, tidningen, tidn, adressblad, stockholmsbladet, tidniiig, corrcspondenten, tiduing, tidningeu, tidningar, tivningen, artonbladet, tidningcn, tidningarnc, tidningc, tivning, tiduingarne, tidninqen, tidnii, idningen, tiduiug, profnumret, iidningarne
1860 : tidn, tidniugarne, tidninge, tidningeu, lidning, lidningarne, tidningariie, tidniugen, tiduing, tiduingar, tidniiigen, iidningen, tiduingarne, tidniug, tidniiigarne, tidnidgar, prenumeratiousanmälan, tidni, tiduiug, tidningarue, tiduingen, tidningm, iidning, tidnin, tidningarnc, tidningan, lördagsnnmmer, tidninq, lidningarnes, tidiiiiig
1870 : tidni, tidnin, lidningen, tidninge, tidningare, tiduing, tidniugen, östgöten, tidningm, tidningen, tidningan, tidningar, tidningara, tiduingens, tidningeus, tidningeu, tiduingen, linköpingsöstgöten, tidningn, liduingen, tidniugens, tiduiugen, tidniug, tidaiag, iveckoblad, werinlandstidningen, tidniugar, norrboltenskuriren, tidningon, tidningons
1880 : tidniug, tidni, tidning⟨num⟩, tidnidgens, tidoingar, tiduing, tidniugen, tidnngen, tidniugar, tidningea, tidningarb, tidniogar, tidnidgen, tidningar, tidningars, tidnin, tidaingen, tdningen, tidningari, tidningenb, tidningeus, tidningeu, utkommaude, tidningan, halfvecko, oscarshainnsposten, nerikea, tidningför, tidoing, tiduiugar
1890 : tidning⟨num⟩, tidnin, tidningeus, tiddingar, tiduiug, tidni, tidnidgar, tidn, tiduing, tidniugens, tidnirg, tidningen, tidniag, tiduingars, tidningan, tidningarb, tidningenb, iördagsnumret, tidniogen, tidninq, tidniug, tidningt, tidoing, tidningans, bagbladets, correspondenton, tidniugar, tidnrs, tidaing, tidningon
1900 : tidningen, tldn, tidningen⟨num⟩, idningen, tidningeu, tidnirgen, tidnin, tidniagen, tidni, tidningenb, tiduingen, utkomma, tidningi, ntkommer, geflepostens, eration, tidninge, tidaing, _af, återlenmas, tidlingen, ntgifves, tidnmgen, tidningan, tidu, tidningea, tidning⟨num⟩, gftnger, kalmartidningen, tidningenkalmar
1910 : tidninga, tidningan, tidni, kalmartidning, dalp, länstidning, tidningen, annonsöre, tidnin, tidningi, tidaingen, afma, expressupplaga, tidningea, tidning⟨num⟩, ntkommer, tisdagsnummer, pronumoration, måndagstidning, tidnmg, dalademokraten, oskarshamnsredaktion, kalmarkalmar, fidningen, darlgenom, huvudstadstidning, oskarshamns, fidning, tidninge, länstidningen
1920 : dalabladet, dalatidningen, falukurircn, falukuriren, julnummer, faitt, dalademokraten, deum, ienna, tisdagsnummer, länstidnings, dalabladets, aupplaga, itidning, tidningar, insändare, jenna, eratio, daina, folkblad, ntkommer, kurir, dellim, nyckelsida, tuln, östgöte, ortstidning, meration, deritia, annonsavdelning

B. Plotting of RBO vs. absolute frequency