

• Peter Fankhauser, Oliver Pfefferkorn

On the Role of Historical Newspapers in Disseminating Foreign Words in German

LREC 2014

Mannheim Corpus of Historical Magazines and Newspapers

- **21 Magazines and Newspapers**
- **652 Individual Volumes**
- **4678 Pages, 4.1 Mio Tokens**
- **Coverage**
 - 18th and 19th century
 - Germany, Austria
- **Metadata:** Genre (Magazine/Newspaper), PubDate, PubPlace, ...
- **Formats:** JPEG, TUSTEP, TEI P5 (& DTA BF), IDS-XCES, XHTML

Mainzer Journal.

N^r. 8. Donnerstag, den 22. Juni. 1848.

 Wegen des hohen Frohnleichnamstages erscheint morgen keine Zeitung. Die zu der heutigen Nummer gehörende Beilage wird Freitag Vormittag um 11 Uhr ausgegeben.

Der erste Artikel der Grundrechte des deutschen Volkes nach dem Entwurf der Verfassungscommission¹⁾.

III Vom Rheine 20. Juni. „Dem deutschen Volke ist gewährt, freier zu sein, als das Volk irgend eines andern Landes. Es ist ihm die Freiheit gegeben, sich nach Belieben zu betheiligen an der Verwaltung der öffentlichen Angelegenheiten, welche die Ausübung dieser Freiheit betreffen, und sich nach Belieben an der Verwaltung der öffentlichen Angelegenheiten zu betheiligen, welche die Ausübung dieser Freiheit betreffen. Eine Anerkennung der Freiheit des deutschen Volkes ist die Freiheit des deutschen Volkes, sich nach Belieben zu betheiligen an der Verwaltung der öffentlichen Angelegenheiten, welche die Ausübung dieser Freiheit betreffen, und sich nach Belieben an der Verwaltung der öffentlichen Angelegenheiten zu betheiligen, welche die Ausübung dieser Freiheit betreffen.“

Wir freuen haben alle wahren und einsichtsvollen Freunde des Vaterlandes und der Religion die Freiheit begrüßt, weil nur auf dem Boden der Freiheit wurde Religionen blühen kann. Sie haben gesagt, daß sie noch endlich die unwürdige Bevormundung der Religion durch den Staatsobersten aufheben, haben auch zuversichtlich gesagt, daß die Freiheit der Gewissensfreiheit den großen und herrlichen Grundbaß der religiösen Freiheit und der Lichtheit der Kirche vom Staat, wie es in Belgien und Nordamerika bezieht, für alle ConfeSSIONen klar, bestimmt und richtig ausprechen werde. Sie haben begreut am so zuversichtlich erwartet, als nur durch die volle und öffentliche Durchführung der Freiheit der Gewissensfreiheit und der Lichtheit der Kirche vom Staat eine solche und innige politische Einheit Deutschlands möglich ist. In allen diesen Erwartungen aber hat sie der angeführte Artikel des Commisfionsentwurfes bitter getäuscht.

[illegible]

freigt, daß in dem vagen Ausdruck „Freiheit des Bekenntnisses“, auch schon die religiöse und kirchliche Freiheit enthalten sey? so bedenkete man, daß dasselbe Wort bisher in allen modernen Bekenntnissen für die Bekenner der drei Hauptbekenntnisse gestanden, aber die schreindsten Eingriffe in die religiöse Freiheit nicht gehindert hat.

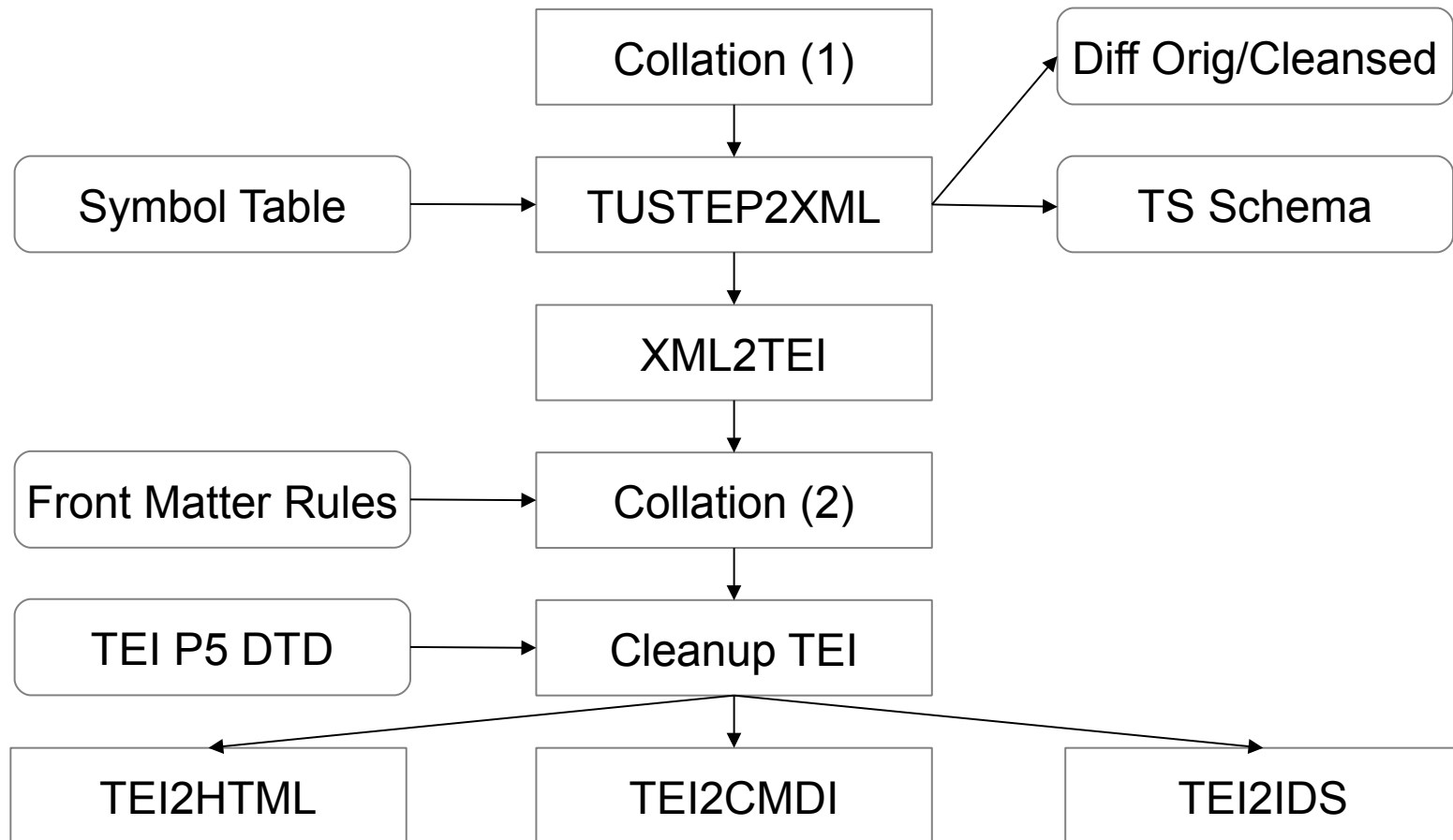
[illegible][illegible]

```

@@@<filename:MainzerJournal(1)0079.jpg/>
$0 @@1@<SE79>
$0079.01 @z <Z><W+1>Mainzer Journal</W+1>.</Z>
$0079.02 #(IMAGE)
$0079.03 <P>_<W-1>Das Mainzer Journal erscheint täglich (mit Ausnahme der höchsten Fes
$0079.04 blättern" schon am Vorabende, die ständige Beilage am Vormittage des betreffer
$0079.05 für Mainz und die nächste Umgebung die Buchhandlung von #S+Kirchheim#S- und #S
$0079.06 vierteljährigen Vorausbezahlungen von 2 fl.; in dem gesammten Gebiete des Fürs
$0079.07 aller Art werden aufgenommen und die dreispaltige Petitzeile oder deren Raum n
$0079.08 #(IMAGE)
$0079.09 <P>{Numero} 43. <Z><W+1>Freitag, den 28. Juli</W+1>.</Z> <R>1848.</R></P>
$0079.10 #(IMAGE)
$0079.11 @z <duolan 2><Z>Der zwölfte Juli und der sechste August.
$0079.12 @Z -----</Z>
$0079.13 <P>_<sj> Wunderbar ist das Walten der Wiedervergeltung in der
$0079.14 Geschichte und es scheint ihr zu gefallen, oft selbst in kleinen,
$0079.15 scheinbar zufälligen Umständen eine Spur und ein Merkzeichen
$0079.16 ihres Waltens auszudrücken.</P>
$0079.17 <P>_Am 12. Juli 1848 hat der Reichsverweser, Erzherzog Jo=
$0079.18 hann von Oesterreich sein Amt feierlich angetreten -- und am
$0079.19 künftigen 6. August wird er vollständig von seiner Gewalt Besitz
$0079.20 ergreifen, indem an diesem Tage die ganze bewaffnete Macht
$0079.21 Deutschlands ihm huldigen und dadurch alle bisherigen Landes=
$0079.22 und Bundestruppen in Reichstruppen verwandelt werden sollen.</P>
$0079.23 <P>_Am 12. Juli des Jahres 1806 aber war es, wo der schmach=
$0079.24 vollste, und am 6. August desselben Jahres, wo der traurigste
$0079.25 Akt der ganzen deutschen Geschichte stattfand. Am 12. Juli 1806
$0079.26 nämlich schlossen die Könige von Bayern und Würtemberg, die
$0079.27 souverainen Fürsten von Regensburg, Baden, Berg, Hessen=
$0079.28 darmstadt, Nassau und andere größere Fürsten des südlichen und
$0079.29 westlichen Deutschlands #S+in Paris den Rheinbund unter
$0079.30 dem Protektorate Napoleons,#S- und sagten sich dadurch
$0079.31 los vom deutschen Reich, und kündeten dem Kaiser, in Ueber=
$0079.32 einstimmung mit dem französischen Gesandten, ihre bisherige
$0079.33 Abhängigkeit auf. Um den eitlen Flitter einer scheinbaren Sou=
$0079.34 verainetät, und den Vortheil eigener Gebietsvergrößerungen,

```

Workflow: TUSTEP to TEI



XHTML

TEI P5

Titel: Mainzer Journal.
Ort: Mainz
Datum: 22-06-1848
Nummer: 8
Seite: 1

Facsimile TEI Tustep

[1] Mainzer Journal.

2) Image

[3] Das Mainzer Journal erscheint täglich (mit Ausnahme der höchsten Festtage) und zwar so, daß das Hauptblatt mit den „Rheinischen Unterhaltungs-
[4] blättern“ schon am Vorabende, die ständige Beilage am Vormittage des betreffenden Tages selbst ausgegeben wird. Bestellungen nehmen alle Postämter an;
[5] für Mainz und die nächste Umgebung die Buchhandlung von Kirchheim, Schott und Thielmann am Leichhofe. Der Preis des Blattes ist hier in Mainz
[6] jährlich 8 fl. in vierteljährigen Vorausbezahlungen von 2 fl.; in dem gesammten Gebiete des Fürstlich Thum- und Taxisschen Postbezirkes jährlich eben=
[7] falls 8 fl. Inserate aller Art werden aufgenommen und die dreispaltige Petitzeile oder deren Raum mit 3 kr. berechnet.

8) Image

9) N^o 8. Donnerstag, den 22. Juni. 1848.

10) Image

[11] Wegen des hohen Frohnleichnamstages erscheint morgen keine Zeitung. Die
[12] zu der heutigen Nummer gehörende Beilage wird Freitag Vormittag um
[13] 11 Uhr ausgegeben.

[14] Der erste Artikel der Grundrechte des deutschen
[15] Volkes nach dem Entwurf der Verfassungs=
[16] commission ¹⁾).

[17] H Vom Rheine 20. Juni. „Dem deutschen Volke ist ge=
[18] währleistet Freiheit des Bekenntnisses, vorbehaltlich der Bestra=
[19] fung von Verbrechen und Vergehen, welche bei Ausübung dieser
[20] Freiheit begangen werden, sowie vorbehaltlich aller staatsbürger=
[21] lichen Pflichten. Einer Anerkennung des Bekenntnisses durch
[22] den Staat bedarf es nicht. Für die Bekenner aller Religionen
[23] Gleichheit vor dem Gesetz. Es ist ausdrücklich die Bildung neuer

(...)leicht, daß in dem vagen Ausdruck „Freiheit des Bekenntnisses,“
[68] auch schon die religiöse und kirchliche Freiheit enthalten sey? so
[69] bedenke man, daß dasselbe Wort bisher in allen modernen Ge=
[70] setzen, wenigstens für die Bekenner der drei Hauptbekenntnisse ge=
[71] standen, aber die schreiendsten Eingriffe in die religiöse Freiheit
[72] nicht gehindert hat.

[73] Allein selbst diesem Ausdruck, „Freiheit des Bekenntnisses,“
[74] hat die Commission noch Clauseln und Bestimmungen angehängt,
[75] die, wenn sich nicht eine geheime, die volle Freiheit des Bekennt=
[76] nisses beeinträchtigende Nebenabsicht dahinter verbirgt, mindestens

```
- <TEI>
+ <teiHeader></teiHeader>
- <text>
- <front>
- <div>
  <pb n="1" facs="http://hdl.handle.net/10932/00-01B8-B175-5D03-A101-2"/>
  <milestone type="tustepsource" facs="http://hdl.handle.net/10932/00-01B8-B175-3573-A001-5" unit="absent"/>
  <p rend="text-center">
    <hi rend="fractlarge">
      <lb n="1"/>
      Mainzer Journal.
    </hi>
  </p>
  <p>
    <lb n="2"/>
    <graphic/>
  </p>
  <p rend="posindent1">
    <hi rend="fractsmall">
      <lb n="3"/>
      Das Mainzer Journal erscheint täglich (mit Ausnahme der höchsten Festtage) und zwar so, daß das Hauptblatt mit den „Rheinischen
    </w>
    Unterhaltungs=
    <lb n="4" break="no"/>
    blättern
  </w>
  “ schon am Vorabende, die ständige Beilage am Vormittage des betreffenden Tages selbst ausgegeben wird. Bestellungen nehmen alle Postämter an;
```

Navigation: based on CMDI

Mannheimer Korpus Historischer Zeitungen und Zeitschriften

[de] Das Mannheimer Korpus Historischer Zeitungen und Zeitschriften besteht aus 21 deutschsprachigen Zeitungen und Zeitschriften des 18. und 19. Jahrhunderts und umfasst etwa 652 einzelne Ausgaben mit insgesamt über 4.1 Mio laufende Wortformen auf 4678 Druckseiten. Dieses Korpus wurde zwischen 2009 und 2011 zusammengestellt und digitalisiert, und 2013 von TUSTEP nach TEI P5 konvertiert.

[en] The Mannheim Corpus of Historical Newspapers and Magazines consists of 21 German newspapers and magazines from the 18th and 19th century. It comprises about 652 individual volumes with over 4.1 Mio word tokens on 4678 pages overall. This corpus has been assembled and digitized from 2009 to 2011, and been converted to TEI P5 in 2013.

Type: collection

PID: <http://hdl.handle.net/10932/00-01B8-AE41-41A4-DC01-5>

Publisher: Institut für Deutsche Sprache

hasPart: Der allerneuesten Europäischen Welt- und Staats=Geschichte II. Theil (1744)
hasPart: Allgemeine Auswanderungs=Zeitung (1846 - 1848)
hasPart: Allgemeine Zeitung (1871)
hasPart: Arbeitgeber. Archiv für die gesamte Volkswirtschaft, Central-Anzeiger für Stellen- und Arbeitergesuche (1870 - 1877)
hasPart: Die Bayerische Presse (1850)
hasPart: Bayreuther Zeitungen (1752)
hasPart: Berlinische Nachrichten von Staats- und gelehrten Sachen (1740 - 1741)
hasPart: Berlinische Privilegirte Zeitung (1737)
hasPart: Conversations=Blatt zur Unterhaltung und Belehrung für alle Stände (1836 - 1838)
hasPart: Das Heiler=Blatt oder Magazin zur Verbreitung gemeinnütziger Kenntnisse (1834)
hasPart: Das Pflennig=Magazin der Gesellschaft zur Verbreitung gemeinnütziger Kenntnisse (1836 - 1855)
hasPart: Das wohlfeilste Panorama des Universums zur erheiternden Belehrung für Jedermann und alle Länder (1834 - 1836)
hasPart: Deutsche Auswanderer=Zeitung (1852)
hasPart: EUROPA Wochenschrift für Kultur und Politik (1905)

Mainzer Journal

Type: collection

PID: <http://hdl.handle.net/10932/00-01B8-B174-96E3-9C01-0>

Publisher: Institut für Deutsche Sprache

isPartOf: Mannheimer Korpus Historischer Zeitschriften und Zeitungen
hasPart: Mainzer Journal, Nr. 8, 22-06-1848
hasPart: Mainzer Journal, Nr. 10, 25-06-1848
hasPart: Mainzer Journal, Nr. 13, 28-06-1848
hasPart: Mainzer Journal, Nr. 15, 30-06-1848
hasPart: Mainzer Journal, Nr. 16, 01-07-1848
hasPart: Mainzer Journal, Nr. 19, 04-07-1848
hasPart: Mainzer Journal, Nr. 22, 07-07-1848
hasPart: Mainzer Journal, Nr. 26, 11-07-1848
hasPart: Mainzer Journal, Nr. 29, 14-07-1848
hasPart: Mainzer Journal, Nr. 30, 15-07-1848
hasPart: Mainzer Journal, Nr. 34, 19-07-1848
hasPart: Mainzer Journal, Nr. 37, 22-07-1848
hasPart: Mainzer Journal, Nr. 39, 24-07-1848
hasPart: Mainzer Journal, Nr. 42, 27-07-1848
hasPart: Mainzer Journal, Nr. 43, 28-07-1848
hasPart: Mainzer Journal, Nr. 45, 30-07-1848
hasPart: Mainzer Journal, Nr. 46, 31-07-1848
hasPart: Mainzer Journal, Nr. 47, 01-08-1848
hasPart: Mainzer Journal, Nr. 48, 02-08-1848
hasPart: Mainzer Journal, Nr. 49, 03-08-1848

Mainzer Journal, Nr. 10

Type: text

PID: <http://hdl.handle.net/10932/00-01B8-B176-54E3-A701-7>

Publisher: Institut für Deutsche Sprache

isPartOf: Mainzer Journal
source: HTML
source: TEI
source: TUSTEP
source: TUSTEP
source: MainzerJournal/MainzerJournal(1)0007.jpg
source: MainzerJournal/MainzerJournal(1)0008.jpg
source: MainzerJournal/MainzerJournal(1)0009.jpg
source: MainzerJournal/MainzerJournal(1)0010.jpg
source: MainzerJournal/MainzerJournal(1)0011.jpg
source: MainzerJournal/MainzerJournal(1)0012.jpg

Metadata Information

Creator: generateCMDI.pl
Creation date: 2013-06-23
Profile: clarin.eu:cr1:p_1366895758244
Collection: Institut für Deutsche Sprache, CLARIN-D Zentrum, Mannheim

Titel: Mainzer Journal.
Ort: Mainz
Datum: 25-06-1848
Nummer: 10
Seite: 7 [Facsimile](#) [TEI](#) [Tustep](#)

[1] Mainzer Journal.

[2] [Image](#)

[3] Das Mainzer Journal erscheint täglich (mit Ausnahme der höchsten Festtage) und zwar so, daß das Hauptblatt mit den „Rheinischen Unterhaltungs=blättern“ schon am Vorabende, die ständige Beilage am Vormittage des betreffenden Tages selbst ausgegeben wird. Bestellungen nehmen alle Postämter an; [4] für Mainz und die nächste Umgebung die Buchhandlung von Kirchheim, Schott und Thielmann am Leichhofe. Der Preis des Blattes ist hier in Mainz [5] jährlich 8 fl. in vierteljährigen Vorausbezahlungen von 2 fl.; in dem gesammten Gebiete des Fürstlich Thurm= und Taxisschen Postbezirkes jährlich eben= [6] falls 8 fl. Inserate aller Art werden aufgenommen und die dreispaltige Petitzelle oder deren Raum mit 3 kr. berechnet.

[8] [Image](#)

[9] N^o 10.

Sonntag, den 25. Juni.

1848

[10] [Image](#)

[11] Die absolute Demokratie und die Geschichte.

[12] J Mainz 23. Juni. Das, worauf eine durch Zahl ansehn= [13] liche und durch ihre Talente und ihre Energie kräftige Partei [14] lossetzt, ist die absolute Demokratie und eine einige und untheil= [15] bare Republik für ganz Deutschland. Hierin allein, glauben [16] jene Männer, könne die Freiheit vollständig verwirklicht und die [17] Größe Deutschlands nach Außen begründet werden. Was das [18] letztere betrifft, so würde in der That vielleicht Deutschland in [19] der gewaltigen Aufregung, in welche es dadurch sich versetzt [20] fände, eine ähnliche Kraft nach Außen entwickeln, wie das re= [21] publikanische Frankreich am Ende des vorigen und am Anfange [22] dieses Jahrhunderts. Was aber das Ende der einigen absolut [23] demokratischen deutschen Republik seyn würde, darüber mag die [24] Geschichte uns belehren. Bei den Männern freilich, die man [25] die Männer des „blinden“ Fortschritts nennen könnte, weil [26] sie gleichsam mit verbundenen Augen voranstritten, ohne rück=

(...)alle Gewalt in der Weise in den Händen des Volkes liegt, daß [72] dasselbe rein nach der Kopfzahl alles entscheidet und regiert oder [73] durch eine Versammlung, die es durch eine ebenso nach der Kopf= [74] zahl berechnete Majorität gewählt hat, beschließen und regieren [75] läßt, so zwar, daß die von der Majorität gefaßten Beschlüsse [76] das einzige und höchste Gesetz sind. Dieses ist jener maj= [77] stätische, unumschränkte, allmächtige Volkswille, der Wille der [78] 40 Millionen, von denen Hecker in seinem Manifest an das [79] Parlament mit einer Art abgöttischer Schwärmerei spricht. Eine [80] solche absolute Demokratie aber, sagen wir, hat nie auf die Dauer [81] bestanden. In allen antiken Republiken war in einem Senat, [82] mochte er heißen wie immer, in alten adeligen Geschlechtern ein [83] höchst zähes, tief in die Geschichte und das Leben verwachsenes [84] conservatives und aristokratisches Element vorhanden, das dem [85] leicht beweglichen und fluthenden Willen der Menge gegenüber [86] die Wage hielt. Außerdem waren die Magistrate, die Consuln, [87] Prätores des alten Roms nichts weniger als bloße Beamten, son=

Foreign Words in Newspapers

- **Hypothesis**

- Historical newspapers popularized foreign words.

- **Corpora & Resources**

- Mannheim Corpus of Historical Newspapers (MKHZ)
 - German Text Archive (DTA, Version Nov 6, 2013)
 - 90 Mio Tokens
 - 1600 – 1920
 - Genres: Factual Writing, Belles Lettres, Learned
 - German Dictionary of Foreign Words (DFWB)
 - ~ 3700 main lemmata for A to Q

Processing

- **Filter by Structure:** Exclude tables/figures/running headers
- **Tokenization:** As given by MKHZ and DTA
- **Lemmatization:** Word Based Lemmatizer (Beliza 1994)
- **Normalization**
 - lower case
 - Expansion by Heuristic Rules
 - $\text{^j} \rightarrow \text{i}$
 - $\text{c}[\text{^h}] \rightarrow \text{k}$
 - $\text{th} \rightarrow \text{t}$
 - $\text{y} \rightarrow \text{i}$
 - ...

Most frequent Foreign Words in MKHZ

	MKHZ	<1800	<1850	>1850	Newspaper	Magazine
ALL	3 913 492	202 654	1 548 983	2 161 855	2 293 010	1 620 482
general/generals/generale/generalen	1 643	321	527	795	1 423	220
personen/person/personen	1 490	133	690	667	1 022	468
minister/ministers/ministern	1 256	223	359	674	1 157	99
partei/parteien/partei/partheien	1 117	10	449	658	922	195
familie/familien	1 097	34	494	569	585	512
präsident/präsidenten	1 084	13	440	631	996	88
interesse/interessen/interessen/interesse	1 051	21	353	677	794	257
fort/forts/forte/forth/fortes	1 004	63	397	544	480	524
millionen/mill/million/mille	965	24	386	555	693	272
armee/armeen	948	163	238	547	851	97
majestät/majestäten	811	296	115	400	707	104
artikel/artikels/artikeln	754	1	292	461	667	87
platz/platze/platzes/platzen	752	33	357	362	363	389
prinzen/prinz/prinze	734	145	215	374	604	130
provinz/provinzen/provinze	727	45	376	306	518	209

Foreign Words by Genre

		MKHZ	DTA
newspapers / factual writing	tokens	2 293 010	7 724 113
	foreign words	59 924	140 250
	percentage	2.61	1.82
magazines / belles lettres	tokens	1 620 482	16 928 838
	foreign words	28 298	216 876
	percentage	1.75	1.28
learned	tokens	–	45 705 849
	foreign words	–	942 084
	percentage	–	2.06
sum	tokens	3 913 492	70 348 800
	foreign words	88 222	1 299 210
	percentage	2.25	1.85

Foreign Words along Time

		MKHZ	DTA
1600 – 1700	tokens	–	8 176 835
	foreign words	–	86 664
	percentage	–	1.06
1700 – 1800	tokens	202 654	14 743 225
	foreign words	5071	211507
	relfreq	2.50	1.43
1800 – 1850	tokens	1 548 983	18565922
	foreign words	31168	348225
	percentage	2.01	1.88
1850 – 1920	tokens	2 161 855	28 872 818
	foreign words	51 965	652 814
	percentage	2.61	2.26

Future Work

- **Corpus**
 - 2500 pages more, filling gaps
 - Integration with DTA: normalization, linguistic processing, structuring
- **Beyond Foreign Words**
 - Analyse diachronic lexical and semantic change
- **Comparison with contemporary News Corpus (Dereko)**
 - Evolution of News as Genre
 - Conventionalization and Diversification
- **Unsupervised, explorative Analysis**
 - HMMs and Topic Modelling