

# Analyseblatt

## AcI-ähnliche WIE-Sätze in Abhängigkeit von sprachexternen Faktoren

Folgende Schritte werden durchgeführt:

- absolute Häufigkeiten des AcI-ähnlichen *Wie*-Satzes anzeigen
- Korpusgrößen (absolute Häufigkeit aller Vorkommen des Lemmas *sehen*) anzeigen
- normierte Zahlen anzeigen
- Diagramm der normierten Zahlen (hierzu werden die absoluten Zahlen zu den Korpusgrößen in Bezug gesetzt: absolute Zahl / Korpusgröße \* 1 000 000)
- Chi-Quadrat-Test (für den Vergleich der erwarteten Werte mit den absoluten Zahlen werden für die Berechnung der erwarteten Werte die Korpusgrößen herangezogen)
- Residuen anzeigen (signifikante Abweichungen werden gelb schattiert; die Residuen sind in der Reihenfolge dargestellt, wie die Ausprägungen im Diagramm abgebildet sind)

## Medium

```
> show.table(TABLE) # absolute Häufigkeit des sehen-wie
```

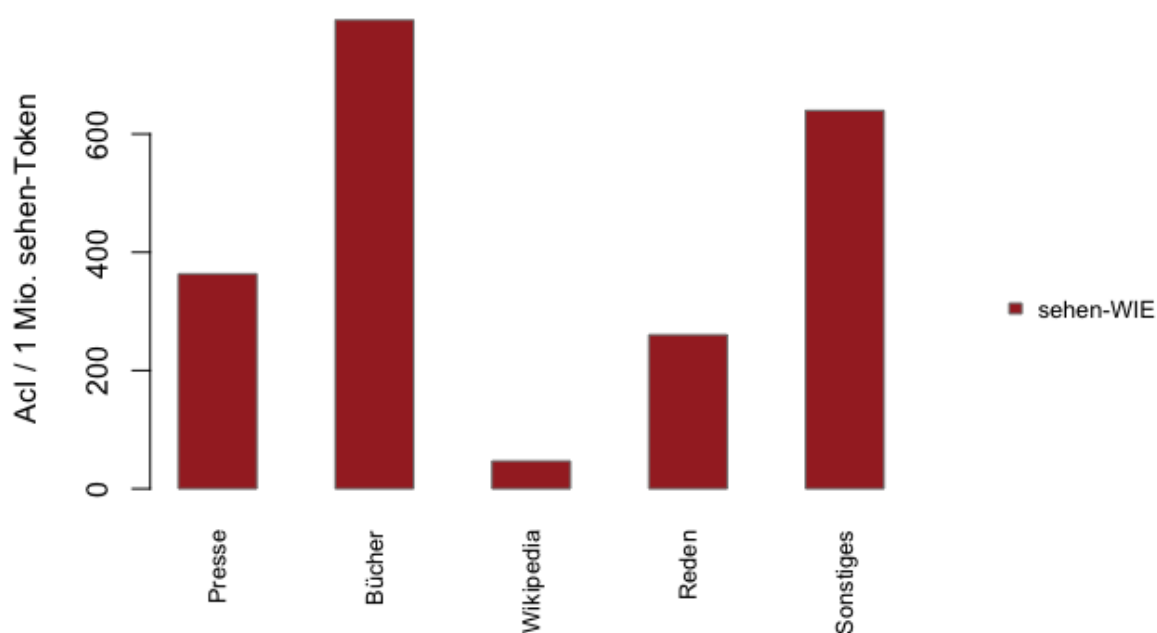
	sehen-WIE
Presse	1821
Bücher	31
wikipedia	35
Reden	78
Sonstiges	7

```
> show.table(corpus.TABLE) # alle vorkommen des Lemmas sehen
```

	sehen-WIE	Korpus
Presse	5009403	
Bücher	39111	
wikipedia	745273	
Reden	299464	
Sonstiges	10941	

```
> show.norm.table(TABLE, corpus.TABLE) # normierte Zahlen
```

	sehen-WIE
Presse	363.51637
Bücher	792.61589
wikipedia	46.96266
Reden	260.46536
Sonstiges	639.79527



```
> chi2
X-squared = 234.96, df = NA, p-value = 0.0004998 (simulated, based on 2000 replicates)
```

```
> residuals(chi2)
5.04 5.17 -13.26 -1.91 1.84
```

## Jahrzehnt

```
> show.table(TABLE) # absolute Häufigkeit des sehen-wie
```

Jahrzehnt	sehen-WIE
-69er	8
70er	5
80er	5
90er	549
2000er	898
2010er	507

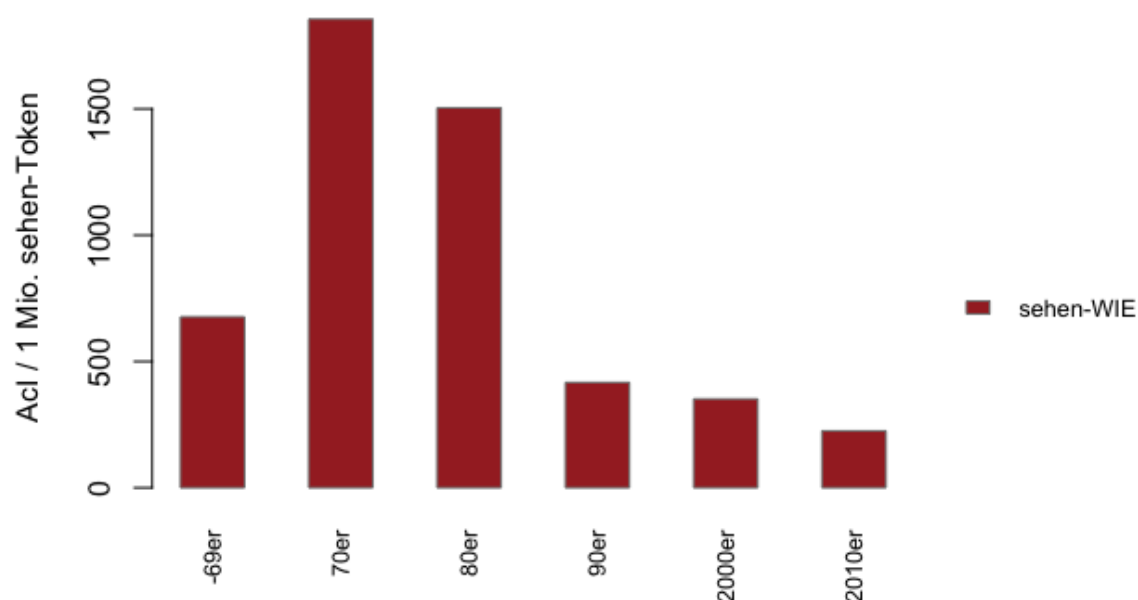
```
> show.table(corpus.TABLE) # alle Vorkommen des Lemmas sehen
```

Jahrzehnt	sehen-WIE	Korpus
-69er		11823
70er		2695
80er		3324
90er		1316805
2000er		2552673
2010er		2251129

```
> show.norm.table(TABLE, corpus.TABLE) # normierte Zahlen
```

Jahrzehnt	sehen-WIE
-69er	676.6472
70er	1855.2876
80er	1504.2118

90er	416.9182
2000er	351.7881
2010er	225.220



```
> chi2
```

```
X-squared = 148.41, df = NA, p-value = 0.0004998 (simulated, based on 2000 replicates)
```

```
> residuals(chi2)
```

```
2.16  4.44  3.81  6.12  2.72 -8.04
```

## Register

```
> show.table(TABLE) # absolute Häufigkeit des sehen-wie
```

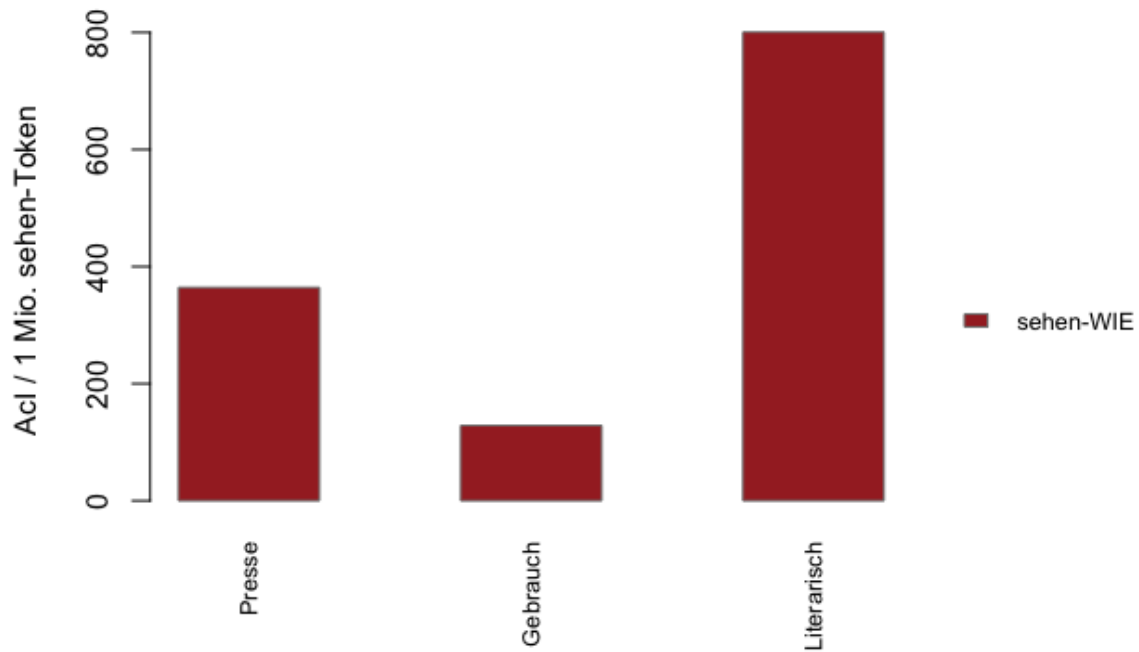
	sehen-WIE
Presse	1790
Gebrauch	153
Literarisch	29

```
> show.table(corpus.TABLE) # alle Vorkommen des Lemmas sehen
```

	sehen-WIE	Korpus
Presse	4911363	
Gebrauch	1190860	
Literarisch	36226	

```
> show.norm.table(TABLE, corpus.TABLE) # normierte Zahlen
```

	sehen-WIE
Presse	364.4609
Gebrauch	128.4786
Literarisch	800.5300



```
> chi2
```

```
X-squared= 192.2, df = NA, = 0.0004998 (simulated, based on 2000 replicates)
```

```
> residuals(chi2)
```

```
5.34 -11.73 5.1
```

## Domäne

```
> show.table(TABLE) # absolute Häufigkeit des sehen-wie
```

	sehen-WIE
unkl	87
Fiktion	32
Kultur	546
Mensch	33
Politik	1135
Technik	139

```
> show.table(corpus.TABLE) # alle vorkommen des Lemmas sehen
```

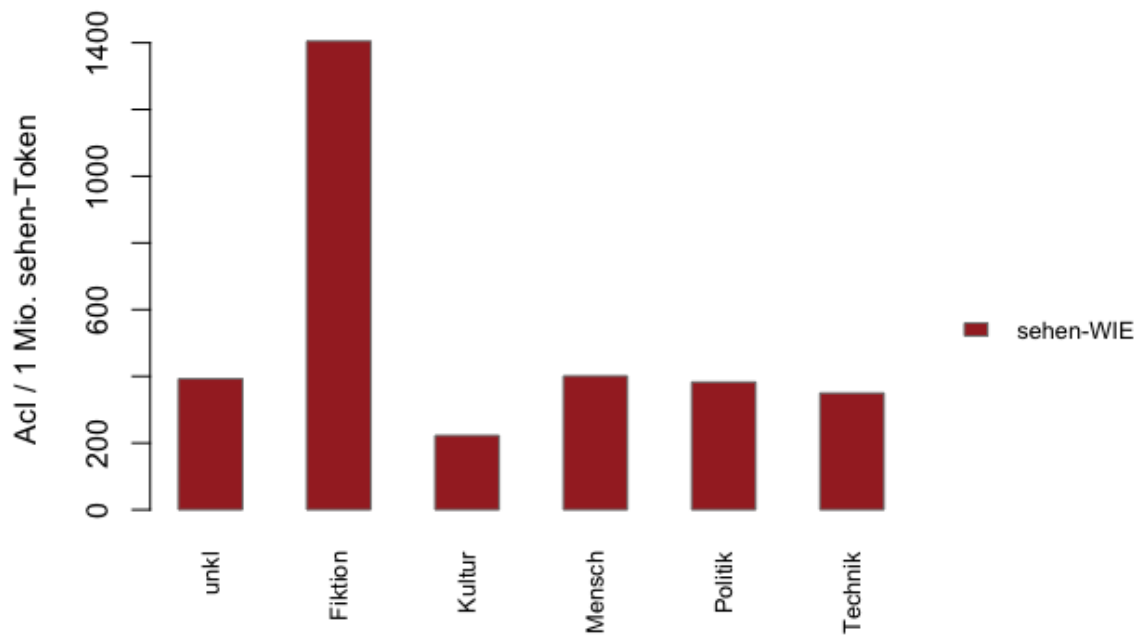
	sehen-WIE	Korpus
unkl	221643	
Fiktion	22778	
Kultur	2449922	
Mensch	82261	
Politik	2964557	
Technik	397288	

```
> show.norm.table(TABLE, corpus.TABLE) # normierte Zahlen
```

```

sehen-WIE
unkl  392.5231
Fiktion 1404.8643
Kultur  222.8642
Mensch  401.1622
Politik 382.8565
Technik 349.8721

```



```
> chi2
```

```
X-squared = X-squared = 198.25, df = NA, p-value = 0.0004998 (simulated, based on 2000 replicates)
```

```
> residuals(chi2)
```

```
1.87  9.13 -8.59  1.28  5.92  1.01
```

## Land

```
> show.table(TABLE) # absolute Häufigkeit des sehen-wie
```

```

sehen-WIE
D          1620
Dost (bis 1990)  5
DWest (bis 1990) 29
A          199
CH         112
LU           7

```

```
> show.table(corpus.TABLE) # alle Vorkommen des Lemmas sehen
```

```

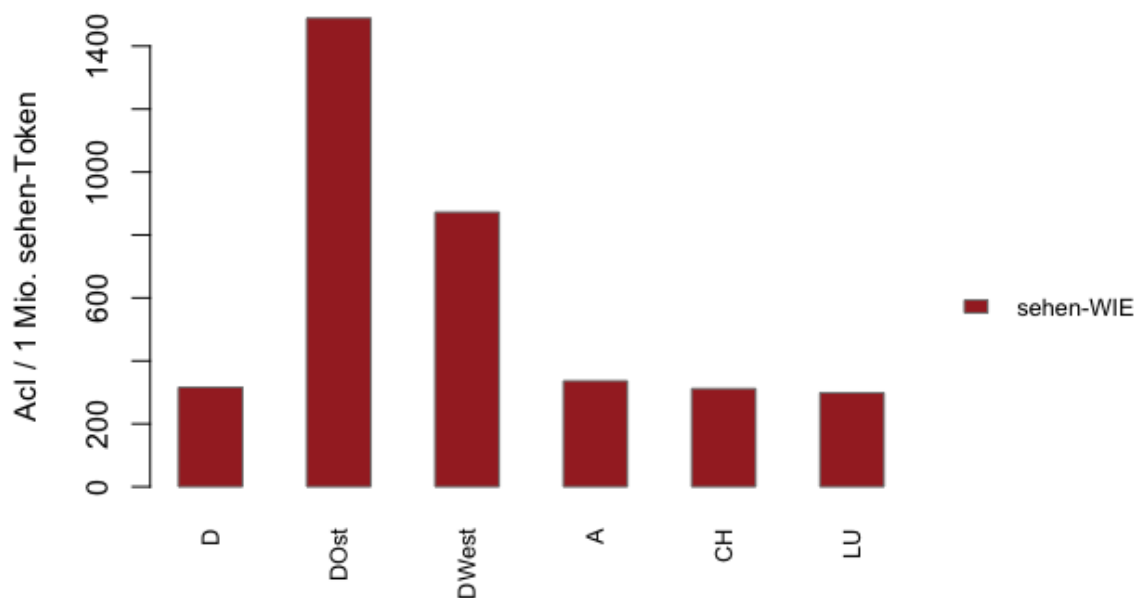
sehen-WIE Korpus
D          5127779
Dost (bis 1990)  3358

```

```
DWest (bis 1990) 33230
A                591503
CH              359128
LU              23451
```

```
> show.norm.table(TABLE, corpus.TABLE) # normierte Zahlen
```

```
sehen-WIE
D                315.9263
DOst (bis 1990) 1488.9815
DWest (bis 1990) 872.7054
A                336.4311
CH              311.8665
LU              298.4947
```



```
> chi2
```

```
X-squared = 46.722, df = NA, p-value = 0.0004998 (simulated, based on 2000 replicates)
```

```
> residuals(chi2)
```

```
-0.67  3.78  5.61  0.65 -0.31 -0.19
```

## Region

```
> show.table(TABLE) # absolute Häufigkeit des sehen-wie
```

```
sehen-WIE
überreg    425
unbek       7
nicht zuzuord 35
Südwest   134
Mittelost   7
```

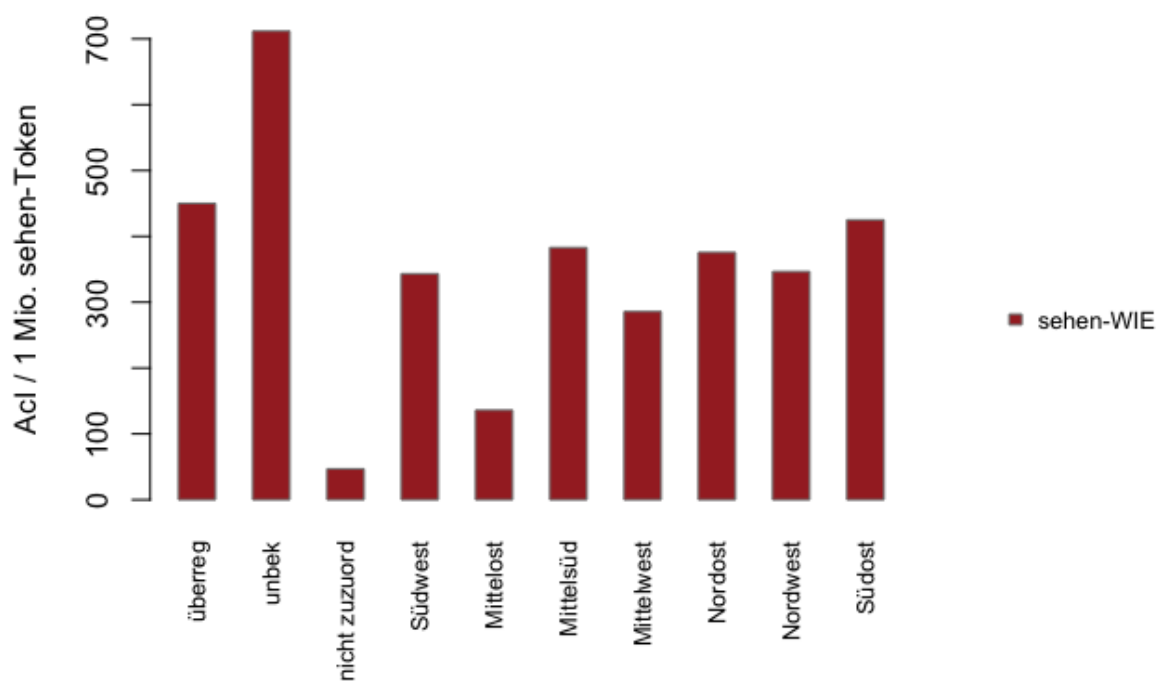
Mittelsüd	128
Mittelwest	345
Nordost	285
Nordwest	154
Südost	452

> show.table(corpus.TABLE) # alle Vorkommen des Lemmas *sehen*

	sehen-WIE	Korpus
überreg		943984
unbek		9836
nicht zuzuord		745601
Südwest		390341
Mittelost		51464
Mittelsüd		334153
Mittelwest		1207271
Nordost		758778
Nordwest		444227
Südost		1063187

> show.norm.table(TABLE, corpus.TABLE) # normierte Zahlen

	sehen-WIE
überreg	450.2195
unbek	711.6714
nicht zuzuord	46.9420
Südwest	343.2896
Mittelost	136.0174
Mittelsüd	383.0581
Mittelwest	285.7685
Nordost	375.6039
Nordwest	346.6696
Südost	425.1369



> chi2

```
X-squared = 275.83, df = NA, p-value 0.0004998 (simulated, based on 2000 replicates)
```

```
> residuals(chi2)
```

```
6.34  2.07 -13.5  0.4  -2.44  1.64  -2.76  2.11  0.56  5.30
```