МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

**НЕМЕЦКИЙ ЯЗЫК
В СОВРЕМЕННОМ МИРЕ:
ИССЛЕДОВАНИЯ СТАТУСА И КОРПУСА
И ВОПРОСЫ МЕТОДИКИ
ПРЕПОДАВАНИЯ**

Материалы II Международного научного форума

18–19 сентября 2019 г.

Томск
Издательский Дом Томского государственного университета
2019

A.-M. Puolokainen

Petrosawodsker Staatliche Universität

DAS BILD VON TÜBINGEN IM DEUTSCHSPRACHIGEN RAUM: KORPUSUNTERSTÜTZTE ANALYSE

Zusammenfassung. Dieser Artikel enthält Statistiken zur Verwendung des Wortes „Tübingen“ im deutschen Referenzkorpus DeReKo COSMAS II sowie eine Analyse der Verwendung des Wortes „Tübingen“ zu verschiedenen Zeitpunkten und aus verschiedenen Quellen.

Schlüsselwörter: Korpuslinguistik, Korpus Cosmas II, Tübingen.

Die Stadt Tübingen hat einen weltweiten Ruf als eine Universitätsstadt, denn an Eberhard-Karls-Universität zu Tübingen haben viele bekannte Wissenschaftler, Dichter und Theologen studiert, und heute ist jeder vierte Bürger irgendwie mit der Universität verbunden. Die meisten Menschen, die etwas von Tübingen wissen, verbinden diesen Stadtnamen also mit der Universität.

Das Ziel meiner Untersuchung war, diese theoretische Hypothese am Beispiel von konkretem sprachlichem Material aus deutschen Sprachkorpora empirisch zu überprüfen. In dieser Arbeit wird der Name der Stadt *Tübingen* mittels spezifischer Software im riesigen Sprachkorpus quantitativ beziehungsweise qualitativ analysiert, um die Annahme zu bestätigen, dass man das Wort *Tübingen* im Kontext der Universität und des Studiums am meistens verwendet.

Zur Hilfe wurde das Deutsche Referenzkorpus DeReKo des Instituts der Deutschen Sprache genommen, das über COSMAS II kostenlos abfragbar ist. Mit 43 Milliarden Wörtern (Stand 18.03.2019) ist das die weltweit größte Sammlung elektronischer Korpora mit geschriebenen deutschsprachigen Texten aus der Gegenwart und der neueren Vergangenheit.

Dieses Korpus gibt uns die Möglichkeit, nicht nur die Verwendung von Wörtern im Kontext zu sehen, mit diesem Korpus kann man aber auch automatische quantitative Analyse durchführen.

Ich möchte die Ergebnisse der Recherche im Archiv der geschriebenen Sprache vorstellen. Dieses aus 790 Korpora und über 8,5 Milliarden Wörtern bestehende Archiv erwies fast 46000 Erwähnungen des Wortes *Tübingen*. Werkzeuge von COSMAS II ermöglichen automatisierte Analyse nach

verschiedenen Kriterien, wie Quellen, Themen, Schreibweise usw. Es entsteht das folgende Bild von *Tübingen* anhand von diesen Ergebnissen.

1. Quellenansicht

Treffer	Texte	von	bis	Quelle
6.615	4.786	1985	2018	Mannheimer Morgen
6.478	4.774	1992	2018	Süddeutsche Zeitung
3.959	3.096	1996	2018	Rhein-Zeitung
2.941	2.193	1947	2018	Der Spiegel
2.809	1.935	1953	1998	Die Zeit
2.667	2.023	2000	2018	die tageszeitung
2.294	1.678	1998	2018	Berliner Morgenpost
1.958	1.365	1999	2018	Spiegel-Online
1.953	1.543	1990	2018	Nürnberger Nachrichten
1.376	971	2002	2018	Nürnberger Zeitung
1.197	990	2000	2018	Neue Zürcher Zeitung
1.153	754	1999	2014	Die Zeit (Online-Ausgabe)
986	764	2009	2018	Berliner Zeitung
975	705	2005	2013	Braunschweiger Zeitung
879	699	2000	2018	Nordkurier
725	218	1996	2012	Plenarprotokolle des Parlaments Landtag von Baden-Württemberg
623	508	1997	2018	St. Galler Tagblatt
619	481	1997	1999	Frankfurter Rundschau
610	548	1991	2018	Die Presse
556	422	2000	2018	FOCUS
45.913	33.904	1808	2018	92 Quellen

Abb. 1. Ausschnitt aus der Ergebnistabelle „Quellenansicht“ mit den höchsten Werten

Wenn wir die Ergebnistabelle Quellenansicht analysieren, können wir sehen, dass das Wort *Tübingen* in verschiedenen Arten der Quellen vorkommt – in Handbüchern, Texten von Goethe, Belletristik des 20. und 21. Jahrhunderts, in Plenarprotokollen des Parlaments usw. Es ist interessant, dass dieses Wort uns in den Zeitungen und Zeitschriften öfter als in anderen Quellen begegnet. Wir können daraus schließen, dass diese Stadt bei Journalisten sehr beliebt ist. Merkwürdig ist dabei, dass nicht nur regionale Ausgaben wie Mannheimer Morgen und Rhein-Zeitung hohe Trefferwerte zeigen, sondern auch alle überregionalen größten Zeitungen und Zeitschriften wie Süddeutsche Zeitung, Der Spiegel, Die Zeit, die Tageszeitung usw. Tausende von Erwähnungen zeugen davon, dass die Stadt Tübingen einen wichtigen Platz im heutigen Leben von Deutschland einnimmt.

2. Ansicht nach Ländern.

Treffer	Texte	von	bis	Land
41.263	30.039	1808	2018	D
2.986	2.449	1996	2018	CH
1.599	1.359	1991	2018	A
65	57	2008	2018	L
45.913	33.904	1808	2018	4 Länder

Abb. 2. Ergebnistabelle „Ansicht nach Ländern“

Gemessen an der Trefferanzahl des Wortes *Tübingen* in den Quellen aus verschiedenen Ländern, ist der Ort in Deutschland bekannter als in anderen Ländern. Das ist aus der Lage der Stadt vorhersehbar. Aber es gibt auch Tausende von Treffern in den ausländischen Quellen, in erster Linie in den schweizerischen. In österreichischen Texten des Archivs wurde *Tübingen* rund 1600 mal zwischen 1991 und 2018 erwähnt, was fast zweimal so wenig ist, wie in den Texten aus der Schweiz. Auch in Luxemburg kennt man *Tübingen*, auch wenn die Zahl von Erwähnungen in den letzten 20 Jahren unter 70 ist.

3. Ansicht nach Textsorten

Ansicht nach Textsorten belegt wieder, dass *Tübingen* am meisten als Nachrichtengrund gebraucht wird. In der Ergebnistabelle sieht man in der ersten Zeile, dass die meisten Textsorten undefiniert sind. Volltexte zeigen aber, dass es meist Berichte und Nachrichten sind. Dass sie als solche nicht definiert wurden, kann wohl damit verbunden sein, dass die entsprechenden Korpora im untersuchten Archiv nach Textsorten nicht annotiert wurden.

Die häufigsten Treffer des Wortes *Tübingen* finden sich in den Texten der Berichte und in den Nachrichten, insbesondere im 21. Jahrhundert. Die seltensten Treffer sind in Serien und Interviews. Der früheste Treffer in den nach Textsorten annotierten Quellen ist Feuilleton aus dem Jahr 1949.

4. Ansicht nach Themen

Die Ansicht nach Themen lässt uns auch manches erschließen. Die Artikel über *Tübingen* beziehen sich auf ganz verschiedene Themen: Das Archiv definiert 48 Themen. Merkwürdig ist, dass das Thema Wissenschaft erst die dritte Zeile belegt, folgend nach Politik und Kultur.

Seit 1947 kommt *Tübingen* am häufigsten in den Texten über Innenpolitik vor.

Den zweiten Platz belegen die Texte über Literatur, wie z.B. Nachrichten über Literaten und ihre Biographien.

Erst dann kommt das Thema Wissenschaft. Der Blick in die Volltexte zeigt, dass es da meistens um Forschungen und Meinungen von Wissenschaftlern aus der Universität Tübingen geht.

Bemerkenswert ist, dass zwischen den Themen Wissenschaft und Bildung, die in Bezug auf die Universitätsstadt Tübingen vorsehbar sind, auf der vierten Zeile Thema Sport / Fußball platziert ist. Das entdeckt für uns, wie wichtig Fußball für Tübingen ist.

Am seltensten kommt das Wort *Tübingen* in Texten über Freizeit und Unterhaltung vor.

Treffer	Texte	von	bis	Textsorte
35.395	26.378	1808	2018	undefiniert
5.004	3.760	1993	2014	Bericht
1.200	839	2014	2018	Bericht ; Nachrichten
917	360	1996	2012	Protokoll; Plenarprotokoll
737	531	1949	2018	Feuilleton
506	369	2000	2018	Meldung; Agenturmeldung
403	281	1992	2018	Porträt
357	331	1969	2018	Leserbrief
273	217	1994	2017	Rezension; Buchrezension
262	199	1994	2014	Kommentar
193	137	1992	2014	Interview
87	61	1993	2013	Serie
72	67	1995	1999	Bericht; Veranstaltungshinweise
62	52	2014	2018	Bericht ; Reportage
56	48	1992	2013	Personalie
55	43	1956	2011	Meldung
49	22	2007	2013	Reportage
38	27	2014	2018	Bericht ; Feuilleton
32	12	1995	2003	Dossier
22	15	2001	2012	Dokumentation
45.913	33.904	1808	2018	52 Textsorten

Abb. 3. Ausschnitt aus der Ergebnistabelle „Ansicht nach Textsorten“ mit den höchsten Werten

Treffer	Texte	von	bis	Thema (>= 50%)	Zuverlässigkeitswe
4.939	3.369	1947	2018	Politik; Inland	
4.477	3.497	1947	2018	Kultur; Literatur	
4.304	3.324	1950	2018	Wissenschaft; Populaerwissenschaft	
4.080	2.608	1967	2018	Sport; Fussball	
3.641	2.477	1947	2018	Staat Gesellschaft; Bildung	
2.362	1.694	1948	2018	Freizeit Unterhaltung; Reisen	
2.349	1.800	1954	2018	Gesundheit Ernährung; Gesundheit	
2.337	1.886	1947	2018	Staat Gesellschaft; Kirche	
2.143	1.493	1948	2018	Staat Gesellschaft; Biographien Interviews	
1.810	1.454	1948	2018	Staat Gesellschaft; Familie Geschlecht	
1.756	1.325	1947	2018	Kultur; Bildende Kunst	
1.284	841	1948	2018	Staat Gesellschaft; Recht	
1.090	859	1961	2018	Sport; Vermischtes	
933	659	1947	2018	Politik; Kommunalpolitik	
820	698	1949	2018	Politik; Ausland	
813	545	1955	2018	Staat Gesellschaft; Verbrechen	
800	680	1948	2018	Kultur; Musik	
644	588	1960	2018	Freizeit Unterhaltung; Vereine Veranstaltungen	
591	487	1947	2018	Kultur; Darstellende Kunst	
532	341	1992	2013	Rest; Ligatabellen	
45.913	33.904	1808	2018	48 Themen (>= 50%)	

Abb. 4 Ausschnitt aus der Ergebnistabelle „Ansicht nach Themen“ mit den höchsten Werten

5. Ansicht nach Jahrzehnt.

Treffer	Texte	Jahrzehnt
18.549	13.835	2000-2009
15.908	11.588	2010-2019
7.346	5.522	1990-1999
1.378	943	1980-1989
1.191	871	1960-1969
1.010	742	1970-1979
448	337	1950-1959
82	65	1940-1949
1	1	1800-1809
45.913	33.904	9 Jahrzehnte

Abb. 5. Ausschnitt aus der Ergebnistabelle „Ansicht nach Jahrzehnt“ mit den höchsten Werten

Wenn man die Anzahl der Treffer von *Tübingen* in verschiedenen Jahrzehnten vergleicht, so stellt man fest, dass sich ihre Anzahl im Laufe der Zeit stets erhöht hat. Im Vergleich zur Mitte des 20. Jahrhunderts hat sich die Anzahl der Erwähnungen im 21. Jahrhundert um etwa vierzigmal erhöht. Das kann man als einen der Beweise der ständigen Entwicklung der Stadt betrachten.

6. Ansicht nach Jahr

Wenn man sich die Trefferstatistik des 21. Jahrhunderts genau erforscht, sieht man, dass das Interesse an *Tübingen* relativ einheitlich und stabil war – von etwa 1.700 bis zu 2.000 Treffern. Die größte Anzahl von Treffern war im Jahr 2006, die kleinste – im Jahr 1998, aber der Unterschied dazwischen beträgt nur 667 Treffer, wenn man die Anzahl nur des 21. Jahrhunderts vergleicht, die kleinste war 2002, und der Unterschied ist – 404.

7. Ansicht nach Tag

Die Funktion von COSMAS II, die Erwähnungen nach Tag automatisiert zu zählen, erschließt uns das Datum, an dem Tübingen das größte öffentliche Interesse rund in Deutschland erweckt hat: Die meisten Erwähnungen von *Tübingen* fallen auf den 22. April 2017. An diesem Tag, als man den Tag der Erde traditionell gefeiert hat, fand in Tübingen der Wissenschaftsmarsch statt. Die Teilnehmer des Wissenschaftsmarsches förderten die Wissenschaft als eine Säule der menschlichen Freiheit und des menschlichen Wohlstands. Diese Tatsache betont nun wieder die erwartete Verbindung Tübingens mit der Wissenschaft und der Universität.

Treffer ▾	Texte	Jahr
⊞	2.135	1.548 2006
⊞	2.110	1.521 2012
⊞	2.085	1.595 2009
⊞	1.994	1.493 2001
⊞	1.967	1.447 2011
⊞	1.918	1.399 2010
⊞	1.896	1.481 2000
⊞	1.893	1.367 2007
⊞	1.789	1.298 2017
⊞	1.765	1.301 2003
⊞	1.759	1.356 2008
⊞	1.757	1.311 2013
⊞	1.731	1.263 2002
⊞	1.700	1.158 2018
⊞	1.673	1.208 2015
⊞	1.654	1.244 2005
⊞	1.637	1.187 2004
⊞	1.610	1.172 2014
⊞	1.502	1.122 1999
⊞	1.468	1.126 1998
45.913	33.904	73 Jahrgänge

Abb. 6 Ausschnitt aus der Ergebnistabelle „Ansicht nach Jahr“ mit den höchsten Werten

Treffer ▾	Texte	Jahr, Tag, Monat
⊞	58	39 2017, 22. April
⊞	48	22 2013, 11. Juni
⊞	47	12 2010, 06. Oktober
⊞	45	22 1999, 11. November
⊞	44	15 2006, 24. Oktober
⊞	38	16 2018, 05. Dezember
⊞	37	25 2009, 14. Mai
⊞	33	22 2000, 18. Dezember
⊞	32	9 2006, 12. Oktober
⊞	32	21 2010, 15. Oktober
⊞	31	8 2004, 15. April
⊞	31	22 2007, 22. Dezember
⊞	31	19 2010, 29. November
⊞	31	16 2018, 30. Juni
⊞	30	1 1957, 12. Dezember
⊞	30	11 2005, 11. November
⊞	30	20 2010, 03. November
⊞	29	4 2002, 17. August
⊞	29	4 2008, ??, ??
⊞	29	17 2008, 19. März
45.913	33.904	10359 Tage

Abb. 7 Ausschnitt aus der Ergebnistabelle „Ansicht nach Tag“ mit den höchsten Werten

8. Ansicht nach Wort-Types

Die Recherche der Ergebnisse der automatisierten Wort-Types-Analyse entdeckt die Verbindung der Stadt zu anderen Orten. Statistik offenbart nicht, in welchem Aspekt die Orte verbunden sind, aber auch rein statistische Daten kann man als Beleg für Gemeinsamkeiten von Tübingen und relevanten Orten betrachten.

Am meisten steht *Tübingen* mit Strich mit Städten Reutlingen, Berlin, Basel, Stuttgart und Heidelberg. Und wenn man die Verbindung mit Reutlingen, Stuttgart und Heidelberg durch Zugehörigkeit zu den größten Städten von Baden-Württemberg erklären kann und Berlin als Hauptstadt in der Wortverbindung auch zu erklären ist, so ist Basel in dieser Statistik eher unerwartet. Das wäre ein interessanter Punkt für kommende Forschungen.

Treffer	Texte	rel [%]	Wort-Types (Eb+Rb+Db+Si)
43.417	32.719	94.564	Tübingen
1.028	1.009	2.239	TÜBINGEN
852	755	1.856	Tübingens
28	27	0.061	Tübingen/Reutlingen
23	23	0.050	Tübingen/Berlin
22	18	0.048	tübingen
18	17	0.039	Tübingen/Basel
18	18	0.039	Tübingen/Stuttgart
16	6	0.035	Metz
16	13	0.035	Tübingen-Stuttgart
13	10	0.028	Tubingen
11	11	0.024	Tübingen-Reutlingen
10	9	0.022	Tübingen/Heidelberg
9	7	0.020	Tübingen-Hirschau
9	9	0.020	Tübingen-Spiel
8	8	0.017	Tübingen/München
7	5	0.015	
7	7	0.015	Tübingen/Mannheim
7	7	0.015	Tübingen-Derendingen
6	6	0.013	Tübingen/Hechingen
45.913	--	100.000	289 Wort-Types

Abb. 8. Ausschnitt aus der Ergebnistabelle „Ansicht nach Wort-Types“ mit den höchsten Werten

Im Großen und Ganzen lässt diese korpusbezogene Untersuchung feststellen, dass Tübingen nicht nur in der wissenschaftlichen Hinsicht einen wichtigen Platz im Leben von Deutschland und anderen deutschsprachigen Ländern einnimmt, sondern auch als politisch und kulturell engagierter Ort gilt.

Quellenverzeichnis

1. Bickel H. Das Internet als linguistisches Korpus // Linguistik online. 2006. S. 71–83.
2. Belica C. Statistische Kollokationsanalyse und Clustering. Korpusanalysemodul. Institut für Deutsche Sprache. Mannheim, 1995. URL: <http://corpora.ids-mannheim.de> (Stand 10.09.2019).
3. Das Deutsche Referenzkorpus DeReKo des Instituts der Deutschen Sprache COSMAS II]. URL: <https://www.ids-mannheim.de/cosmas2/> (Stand 10.09.2019).
4. Lüdeling A. Das Zusammenspiel von qualitativen und quantitativen Methoden in der Korpuslinguistik // Kallmeyer W. & Zifonun G. (Hrsg.): Sprachkorpora – Datenmengen und Erkenntnisfortschritt. Berlin : Walter de Gruyter, 2006. S. 28–48.
5. Offizielle Seite der Stadt Tübingen. URL: <https://www.tuebingen.de/> (Stand 10.09.2019).
6. Reznicek M. et al. Das Falco-Handbuch. Korpusaufbau und Annotationen Version 2.01. Humboldt-Universität zu Berlin, 2012. URL: <https://www.linguistik.hu-berlin.de/de/institut/professuren/korpuslinguistik/forschung/falko/standardseite> (Stand 10.09.2019).
7. Denktionary. Die Arbeit mit COSMAS II. URL: https://wiki.uni-mannheim.de/denktonary/index.php?title=Die_Arbeit_mit_COSMAS_II (Stand 10.09.2019).
8. Wikipedia. Deutsches Referenzkorpus. URL: https://de.wikipedia.org/wiki/Deutsches_Referenzkorpus (Stand 10.09. 2019).