

Toni Bernhart

# „Von Aalschwanzspekulanten bis Abendrotlicht“

Buchstäbliche Materialität und Pathos im *Häufigkeitswörterbuch der deutschen Sprache* von Friedrich Wilhelm Kaeding

Was ist Pathos der Wissenschaften? Eine pathetische Wissenschaft? Eine pathetische Wissenschaftlerin? Wann und wie wird Wissenschaft von Pathos affiziert? Leichter ist es wahrscheinlich zu benennen, was Pathos ist. Größe der Geste, hoher Gültigkeits- und Wahrheitsanspruch, Opulenz und Ornament gehören dazu.

Gegenstand dieses Beitrags ist das *Häufigkeitswörterbuch der deutschen Sprache* von Friedrich Wilhelm Kaeding (1843–1928) aus dem Jahre 1898.<sup>1</sup> Hinsichtlich der Ansprüche, die das Werk erhob, und der Begeisterung, die mit ihm einherging, aber auch gemessen am Aufwand, der für seine Erstellung betrieben wurde, ist das Werk monumental. Es ist Ausdruck großer Geste und Äußerung von Pathos.



Abb. 1: Titelblatt.

<sup>1</sup> [Friedrich] W[ilhelm] Kaeding, Häufigkeitswörterbuch der deutschen Sprache. Festgestellt durch einen Arbeitsausschuß der deutschen Stenographiesysteme, Steglitz bei Berlin 1898.

# 1. Aufbau und Zweck des Wörterbuchs

Das Buch ist 672 Seiten stark. Es gliedert sich in einen längeren ersten und einen kürzeren zweiten Teil.<sup>2</sup> Der erste Teil trägt den Titel *Wort- und Silbenzählungen*. Er enthält Erläuterungen zu Zweck und Ziel des Wörterbuchs, zur angewandten Methode und zum Korpus sowie den eigentlichen Wörterbuchteil, bestehend aus der „Nachweisung der Wortformen“, „Nachweisung der nackten Haupt-(Stamm-)silben“, „Nachweisung der Vorsilben und deren Verbindungen“ und „Nachweisung der Nachsilben“. Der zweite Teil mit dem Titel *Buchstabenzählungen* enthält Abschnitte zu Buchstabenhäufigkeiten nach unterschiedlichen Kriterien und Kombinationsmustern sowie einen „Nachtrag“ mit dem Nachweis der Häufigkeit von Interpunktionszeichen, spezielle Listen zu Silbenhäufigkeiten und einen kurzen Forschungsbericht, in dem Kaeding sein Wörterbuch als vorgänger- und beispielloses Vorhaben charakterisiert und auf William Gambles Häufigkeitsliste chinesischer Schriftzeichen aus dem Jahre 1861<sup>3</sup> als ansatzweise vergleichbaren Versuch hinweist.

Es ist nicht bekannt, in welcher Auflagenhöhe Kaedings *Häufigkeitswörterbuch* erschien. Die kursorische Abfrage einschlägiger Bibliothekskataloge lässt vermuten, dass die Stückzahl der Auflage nicht besonders groß war; auch ist das Buch relativ gering verbreitet und liegt fast ausschließlich in Bibliotheken deutschsprachiger Länder. Eine deutlich weitere Verbreitung lässt die Reprint-Ausgabe (in Auszügen) in den *Grundlagenstudien aus Kybernetik und Geisteswissenschaft* von 1963 erkennen.<sup>4</sup> Geschuldet ist die relativ geringe Verbreitung der Originalausgabe wohl auch dem Umstand, dass das Buch im Selbstverlag des Herausgebers erschien.

---

<sup>2</sup> Der längere erste (S. 1–502) und der kürzere zweite Teil (S. 503–672) sind fortlaufend paginiert, haben aber jeweils ein eigenes Titelblatt mit dem Erscheinungsjahr 1897. Möglicherweise waren Exemplare der beiden einzelnen Teile als Vorab-Ausgaben im Umlauf. Die heute noch erhaltenen und im Umlauf befindlichen Buchausgaben bestehen aus der Zusammenführung der beiden Teile, ergänzt um einen Vorspann (S. I–VI), aus einem zusätzlichen Titelblatt, welches das *Häufigkeitswörterbuch* auf das Jahr 1898 datiert, und einem Inhaltsverzeichnis, das sich über den ersten und den zweiten Teil erstreckt. Dieser Umstand hat zur Folge, dass das *Häufigkeitswörterbuch* bibliographisch mit unterschiedlichen Erscheinungsjahren erfasst wird (1897 oder 1898).

<sup>3</sup> William Gamble, Two lists of selected characters containing all in the Bible and twenty-seven other books, Shanghai 1861.

<sup>4</sup> Friedrich Wilhelm Kaeding, Häufigkeitswörterbuch der deutschen Sprache, 1. Teil: Wort- und Silbenzählungen (auszugsweise Reproduktion), 2. Teil: Buchstabenzählungen (Auszug aus dem Nachtrag), Quickborn bei Hamburg 1963 (Grundlagenstudien aus Kybernetik und Geisteswissenschaft 4).

Eine Besonderheit des Buches sind die Corrigenda. Im Abschnitt „Fehlerberichtigung“<sup>5</sup> sind sie in gedruckter Form aufgelistet. Gleichzeitig sind sie in wohl allen Exemplaren auch handschriftlich umgesetzt. Der Vergleich dreier Exemplare aus unterschiedlichen Aufstellungsorten<sup>6</sup> lässt erkennen, dass die handschriftlichen Korrekturen von derselben Hand ausgeführt wurden. Dies lässt die Annahme zu, dass sie entweder von Kaeding selbst oder von einer von ihm beauftragten Person vor der Auslieferung in allen gedruckten Exemplaren vorgenommen wurden. Auch darin drückt sich Kaedings akribische Arbeitsweise aus.

Man kann im Kaeding nachschlagen, wie oft ein Wort, ein Buchstabe, eine Silbe oder ein Satzzeichen in der deutschen Sprache vorkommen. So erfahren wir, dass „die“, „der“ und „und“ die drei häufigsten Wörter, der Vokal e und die Konsonanten n und r die drei häufigsten Buchstaben sowie Komma und Punkt mit deutlichem Abstand vor allen anderen die zwei häufigsten Interpunktionszeichen der deutschen Sprache sind.<sup>7</sup> Kaedings Zählungen und Berechnungen geben Auskunft darüber, dass ein deutsches Wort aus durchschnittlich 1,83 Silben<sup>8</sup> und eine deutsche Silbe aus durchschnittlich 3,03 Buchstaben besteht<sup>9</sup> und dass eine überaus geringe Zahl von Wörtern (0,13%) eine überaus große Zahl der Belege liefert und umgekehrt: 49,14% aller Wörter sind nur ein einziges Mal belegt.<sup>10</sup> Die Beobachtung dieser Verteilung, die für alle Sprachen gleichermaßen gilt, ist eine Vorwegnahme dessen, was später nach George K. Zipf (1902–1950) als Zipf'sches Gesetz bezeichnete wurde, welches besagt, dass das Produkt aus Rang und Häufigkeit eines Wortes konstant ist.

Kaedings Anspruch aber ist weit spezifischer und differenzierter, als dass es ihm lediglich um die positivistische Dokumentation quantitativer Daten aus Sprache ginge. Ziel seiner Häufigkeitsuntersuchungen war die Optimierung und Standardisierung der Stenographie, die bis zur Erfindung von Geräten für akustische Sprachaufzeichnung eine konkurrenzlose Kulturtechnik war. Im Besonderen ging es Kaeding also um

---

5 Kaeding, Häufigkeitwörterbuch, S. 670f.

6 Miteinander verglichen wurden die Exemplare der Staatsbibliothek zu Berlin – Preußischer Kulturbesitz mit der Signatur HA 5 Rx 2911 und der Universitäts- und Landesbibliothek Tirol, Innsbruck, mit der Signatur GB 1571 K11 sowie das kassationierte Exemplar der ehemaligen Pädagogischen Hochschule Rheinland, Abteilung Bonn, mit der Signatur IIIc 1301, das der Verfasser dieses Beitrags in einem Antiquariat erworben hat.

7 Kaeding, Häufigkeitwörterbuch, S. 53, 643, 648f.

8 Ebd., S. 32.

9 Ebd., S. 513.

10 Ebd., S. 44.

1. die Häufigkeit der einzelnen Wörter, 2. die Häufigkeit der einzelnen Silben, 3. die Häufigkeit der einzelnen Laute, 4. die Häufigkeit der einzelnen Buchstaben und zwar a) der *einzelnen* Buchstaben, b) der Zusammensetzungen, c) der Verschmelzungen.<sup>11</sup>

Gegen Ende des 19. Jahrhunderts gab es, allein für die deutsche Sprache, viele hundert miteinander konkurrierende stenographische Schriftsysteme. Namhafte Kurzschriftpioniere waren der bayerische Ministerialbeamte Franz Xaver Gabelsberger (1789–1849), der Berliner Versicherungsbeamte und Privatgelehrte Heinrich August Wilhelm Stolze (1798–1867) und der Berliner Knopffabrikant und Schreibmaschinenhändler Ferdinand Schrey (1850–1938). Im Jahr 1897 wurde die so genannte „Systemurkunde“ unterzeichnet, welche die Schreibsysteme von Stolze und Schrey vereinheitlichte. Dieses Stolze-Schrey-System ist in der Schweiz bis heute gültig. Deutschland und Österreich einigten sich 1924 auf die „Deutsche Einheitskurzschrift“ (DEK), die nach einer weiteren Reformierung 1968 in der auch heute noch gültigen „Wiener Urkunde“ festgeschrieben wurde. Die DDR trat 1970 diesem Regelwerk bei.

Kaeding wurde am 18. September 1843 in Rathenow geboren, besuchte dort das Gymnasium bis zur Sekunda und war später beim dortigen Kreisgericht beschäftigt. Im Jahre 1868 zog er nach Berlin, wo er ab 1873 bei der Reichsbank tätig war (1882 Kalkulator, 1895 Oberkalkulator, 1899 Rechnungsrat, 1910 Geheimer Rechnungsrat). Zusammen mit dem Stenographen Adolf Dreinhöfer (1852–1896) gründete er 1874 den Verband Stolze'scher Stenographenvereine. Als Verbandsfunktionär und Autor zahlreicher Fachbeiträge zur Stenographie war er maßgeblich an der Vereinheitlichung der Schriftsysteme beteiligt.<sup>12</sup>

Stenographische Schreibsysteme sind für jeweils nur eine einzelne Sprache verwendbar. Sie operieren nach den Prinzipien der Kürzung und Zusammenziehung wiederkehrender Buchstaben-, Silben- und Wortfolgen. Stenographie zielt darauf ab, häufige Wörter und häufige Buchstabenfolgen durch möglichst einfache Schriftzeichen abzubilden und dadurch möglichst schnell schreiben zu können. Kurzschrift in ihrer effizientesten Form, der so genannten Eilschrift, ist in der Lage, deutlich schneller zu schreiben, als jemand spricht.

Untersuchungen zur Feststellung der Häufigkeit deutscher Wörter, Silben, Laute und Lautverbindungen sind bisher in größerem Umfange noch nicht ausgeführt worden [...]. Diese Untersuchungen sind aber unbedingt erforderlich, wenn man ein wirklich zweckmäßiges stenographisches System aufstellen oder ein vorhandenes ausbilden will [...]. Die

<sup>11</sup> Ebd., S. 10, Hervorhebung im Original.

<sup>12</sup> Biographische Notiz nach Karl-Heinz Best, Friedrich Wilhelm Kaeding (1843–1928), in: *Glottometrics* 18 (2009), S. 81–87, hier S. 81. Enthält auch eine Übersicht von Kaedings Publikationen zu Stenographie.

Frage nach der Häufigkeit eines Wortes, einer Silbe oder eines Buchstaben, sowie irgend welcher Verbindung dieser Sprachbestandteile untereinander müssen sich die Vertreter der stenographischen Systeme stets vorlegen, wenn sie Verbesserungsanträge in erschöpfender Weise prüfen wollen.<sup>13</sup>

Kaedings Innovation besteht darin, die Entwicklung und Optimierung stenographischer Schriftzeichen erstmals auf eine systematische und empirische Basis zu stellen: Wenn empirisch gesichert ist, welche Wörter, Buchstabenfolgen und Laute in einer Sprache die häufigsten sind, lassen sich umso gezielter effiziente Kurzschriften entwickeln. Die Abb. 2 zeigt einen Auszug aus der Seite 53 mit den am häufigsten belegten Wörtern.

Auffallend ist auch Kaedings begeisterte Zuversicht, dass nicht nur die „stenographische Wissenschaft“<sup>14</sup> Nutznießerin seiner Forschung sei, sondern dass auch eine Reform der Schreibschrift insgesamt daraus abgeleitet werden könnte:

Erst jetzt bekommen wir einen klaren Einblick in die Zusammensetzung der Sprache, erst jetzt gewinnen wir den Überblick über die Wertverhältnisse der Buchstaben zu einander, welcher uns eine genaue Abschätzung ermöglicht. Wir werden nun auch mit Leichtigkeit berechnen und nachweisen können, in welcher unverantwortlicher Weise die deutsche Schrift die Kraft des Schreibers vergeudet, wir werden die Wege zur Abhilfe finden lernen und in absehbarer Zeit den Schriftverkehr vereinfachen können, nicht durch die Abschaffung der Formen, sondern, was wichtiger ist, durch die Abschaffung des überflüssigen Ballasts an Zeichen. Dass neben der Kurrentschrift auch die stenographische Schrift von dieser Möglichkeit den ausgiebigsten Vorteil ziehen wird, ist selbstverständlich und bedarf hier keiner weiteren Erläuterung.<sup>15</sup>

---

13 Kaeding, Häufigkeitwörterbuch, S. 5.

14 Ebd., S. 505.

15 Ebd.

## Abteilung A I.

Nachweisung der Wörter aus Abteilung A nach der Häufigkeit geordnet.

**Vorbemerkung:** Die Häufigkeitszahlen dieser Nachweisung enthalten auch diejenigen der „trennbaren Vorsilben“ (z. B. zu, in, auf, an u. s. w.), sowie die Ergebnisse aus den Zerlegungen zusammengesetzter Wörter (als erstes Wort, Mittelwort, Schlußwort).

die	358 054	durch	59 999	haben	29 270	wohl	17 608
der	354 526	als	59 279	einer	28 396	sehr	17 352
und	320 985	um	56 968	hin	27 958	dann	16 950
zu	258 584	bei	55 553	mir	26 804	schon	16 748
in	214 308	wie	54 988	doch	25 562	meine	16 199
ein	153 095	für	54 406	ihm	25 214	habe	16 156
an	145 968	ab	52 261	diese	24 706	ihre	16 098
den	141 542	im	50 909	einem	24 247	alle	15 511
auf	127 349	nehmen	45 139	ihr	23 790	seiner	15 311
das	127 137	aber	44 929	uns	23 765	ihnen	15 254
von	118 088	man	44 302	mehr	23 490	kommen	15 168
nicht	115 342	unter	41 432	zum	23 279	selben	15 096
mit	109 958	her	41 358	zur	22 796	Ordnung	14 416
dem	103 691	noch	41 275	hier	22 601	lassen	14 398
das	102 171	mann	40 311	mann	22 117	haben	14 344

**Abb. 2:** „Nachweisung der Wörter aus Abteilung A nach der Häufigkeit geordnet“ (S. 53) mit handschriftlichen Korrekturen, vermutlich von Kaedings Hand.

## 2. Vorlauf und Arbeit am Wörterbuch

Den Hintergrund und das Entstehen seines *Häufigkeitwörterbuchs* dokumentiert Kaeding ausführlich. Die Ende des 19. Jahrhunderts etablierten Kurzschriftsysteme entbehren laut Kaeding einer empirisch gesicherten Grundlage: Sie basieren auf unzureichend großen Korpora und stellen einen „Notbehelf“ dar. Kaeding errechnet, dass erst ein Korpus (er spricht von „Zählstoff“) mit 10 Millionen Wörtern eine tolerierbare „Abweichung[] von der Durchschnittshäufigkeit bei den in Rechnung gestellten Wörtern“ von 3% aufweist und hinreichend groß ist, um sprachspezifisch hoch frequente Wörter, Buchstaben und Zeichenfolgen zu bestimmen und auf dieser Grundlage ein stenographisches System zu optimieren.<sup>16</sup> Das Korpus, das Kaeding dem *Häufigkeitwörterbuch* zugrunde legt, umfasst genau 10.906.235 Wörter<sup>17</sup> bzw. 60.558.018 Buchstaben.<sup>18</sup>

<sup>16</sup> Ebd., S. 6. Damit erfüllt Kaeding hinsichtlich des Korpusumfangs auch heute noch empfohlene Größenordnungen: Willy Martin, *The Frequency Dictionary*, in: F. J. Hausmann u.a. (Hrsg.), *Wörterbücher. Ein internationales Handbuch zur Lexikographie*, Berlin 1990 (Handbücher zur Sprach- und Kommunikationswissenschaft 5), Bd. 2, S. 1314–1322, hier S. 1317, empfiehlt für ein Frequenzwörterbuch ein Korpus aus mindestens 500.000 Wörtern.

<sup>17</sup> Kaeding, *Häufigkeitwörterbuch*, S. 25. ‚Wörter‘ hier im Sinne der Häufigkeit insgesamt: 258.173 Wörter (S. 43) (im Sinne von Tokens) kommen insgesamt 10.906.235-mal vor.

<sup>18</sup> Ebd., S. 513.

Auf Vorschlag des Verbandes Stolze'scher Stenographenvereine, nicht unmaßgeblich in der Person von Kaeding selbst, beschloss der so genannte Stolzetag, die jährliche Hauptversammlung des Verbandes, im September 1891 in Berlin die „baldige Inangriffnahme“ „ausgedehnter Häufigkeitsuntersuchungen“, übertrug Kaeding die Planung und Leitung der „Untersuchungen zur Feststellung der Häufigkeit der Wörter, Silben und Laute in der deutschen Sprache“ und setzte dazu einen Arbeitsausschuss ein. Als große Schwierigkeit erwies sich die Aufbringung der erforderlichen Mittel. Die Stolze'sche Schule steuerte „an Geldmitteln mehr als die Hälfte des ganzen Bedarfes“ bei; woher der andere Teil der Mittel kam und wie hoch die Projektkosten insgesamt waren, dazu macht Kaeding keine Angaben.<sup>19</sup> Der größte Teil der Arbeitsleistung der schätzungsweise über 1.000 Mitarbeiter, die zwischen 1891 und 1897 mit dem Zählen und „Buchen“ der Wörter, Silben und Buchstaben und den zahlreichen Berechnungen beschäftigt waren, erfolgte ehrenamtlich. Unterstützt wurde das Vorhaben auch durch zahlreiche Aufrufe in stenographischen Fachzeitschriften und der Tagespresse.<sup>20</sup> Die größte Herausforderung bestand mangels entsprechend umfassender und systematischer Vorgängerprojekte in der Entwicklung der Systematik und Methode für eine derartige Untersuchung.

Ein erster Schritt war die Korpusbildung.<sup>21</sup> Kaeding war es wichtig, dass der „verwendete Zählstoff“ „möglichst alle Wissensgebiete“ berücksichtigte; er wurde in 16 „Abteilungen“ untergliedert:

1. juristisch, 2. kaufmännisch (aus Abhandlungen allgemeiner Art), 3. desgl. (aus Briefen),
4. theologisch, 5. medizinisch, 6. geschichtlich, 7. gemischt (aus Zeitungen und Büchern verschiedener Zweige), 8. militärisch (allgemeinen Inhalts), 9. militärisch (Briefe), ferner entnommen
10. aus Privatbriefen, 11. dem Buch der Erfindungen von Reuleaux,<sup>22</sup> 12. Klassikern und Novellisten, 13. der deutschen Rundschau, 14. der Bibel (hiervon nur 100.000 Worte), außerdem
15. parlamentarisch, 16. volkswirtschaftlich.<sup>23</sup>

---

**19** Ebd., S. 7.

**20** Ebd., S. 8.

**21** Nachweis aller Korpustexte: ebd., S. 12–20.

**22** Gemeint ist Franz Reuleaux, Das Buch der Erfindungen, Gewerbe und Industrien. Rundschau auf allen Gebieten der gewerblichen Arbeit, 8 Bde., 8. Aufl., Leipzig 1884–1888.

**23** Kaeding, Häufigkeitswörterbuch, S. 11. Unter den „Klassikern und Novellisten“ befinden sich Ludwig Börne, Ernst Eckstein, Goethe, Wilhelm Hauff, Johann Peter Hebel, E.T.A. Hoffmann, Jean Paul, Kleist, Theodor Körner, August von Kotzebue, Lessing, Martin Luther, Jean Baptist Racine (in deutscher Übersetzung), Ernst Raupach, Jean-Jacques Rousseau (in deutscher Übersetzung), George Sand (in deutscher Übersetzung), Schiller, Shakespeare (in deutscher Übersetzung) u.a.m.

Vorbereitend erstellte der Arbeitsausschuss unter Kaedings Leitung einen Arbeitsplan mit „Arbeitsanweisungen“; hierbei war „der Leiter des Königlich Preußischen Statistischen Büreaus, Herr Geheimer Oberregierungsrat Blenck in Berlin“<sup>24</sup> beratend tätig. In der „Abteilung 1“, welche dezentral die „Anfangsarbeit“ leistete, waren 665 Mitarbeiter tätig. Sie hatten

die Aufgabe, die sämtlichen zur Zählung bestimmten Wörter auf einzelne Zählzettel auszuschreiben und dadurch die Grundlage für die weitere Verarbeitung zu bilden. Von den beiden Möglichkeiten „Strichelung“ und „Zählzettel“ wählte der Arbeitsausschuß den sichersten Weg, das Ausschreiben *jedes Wortes* der zu untersuchenden Druckbogen auf einen besonderen Zählzettel von 3 cm Höhe und 7 cm Länge.<sup>25</sup>

Die „Abteilung 2“, in der 167 Mitarbeiter tätig waren, bestand aus den so genannten „Sammelstellen“: „Je 100.000 Wörter bildeten einen größeren Arbeitsteil, ‚Sammelstelle‘ genannt, und solcher Arbeitsteile waren 100 nötig, um die als zweckmäßig erkannte Zahl von 10 Millionen Wörtern unterzubringen.“<sup>26</sup> In den Sammelstellen wurden die Zählungen aus den „Anfangsarbeiten“ zu Teilkorpora von jeweils 100.000 Wörtern zusammengeführt: „Die Inhaber der Sammelstellen hatten die Pflicht, die ihnen von den ‚Anfangsstellen‘ zugehenden alphabetisch vorgeordneten Zettel in eine einzige alphabetische Ordnung zu bringen, [...] das Ganze doppelt zu prüfen und Einheitszettel für jedes Wort ihrer Stelle zu schreiben.“<sup>27</sup>

In der „Abteilung 3“ führten 106 Mitarbeiter die „Buchungen“ durch: Anhand der „Einheitszettel“ aus der Abteilung 2 und der „Zählzettel“ aus der Abteilung 1 wurden hier „Buchungspäckchen“ gebildet und für jedes einzelne Wort ein „Buchungsblatt“ ausgefüllt. Darin wurden die Häufigkeiten nach ihrer Verteilung über die „Wissensgebiete“ (juristisch, kaufmännisch, theologisch etc.) untergliedert. Es gab 84 Buchungsstellen.<sup>28</sup>

In der „Abteilung 4“ erfolgte die „Anlegung der alphabetischen Nachweisung“. 94 Mitarbeiter wirkten darin mit. Es wurde „eine einzige alphabetische Liste des ganzen Stoffes aufgestellt“; sie „umfaßt 817 Hefte von je 5 Bogen, also 4.085 Bogen“. „Jedes Heft wurde für sich doppelt aufgerechnet, nachdem die Eintragungen doppelt geprüft worden waren.“ Die dabei ermittelte Fehlerquote

24 Ebd., S. 22.

25 Ebd., Hervorhebung im Original.

26 Ebd., S. 23.

27 Ebd., S. 24.

28 Ebd., S. 25–31.

beträgt „0,00019% der Wörter“: „eine Zahl, die für die peinliche Sorgfalt bei der Arbeit spricht“, wie Kaeding kommentiert.<sup>29</sup>

Die „Abteilung 5“ (148 Mitarbeiter) zerlegte zusammengesetzte Wörter in ihre alphabetisch geordneten Glieder, die „Abteilung 6“ (72 Mitarbeiter) war für die „Zerlegung der einfachen Wörter und der nach der Abtrennung der Vorsilben übrig gebliebenen Wörter und Wortstümpfe“ zuständig, in der „Abteilung 7“ zerlegten 68 Mitarbeiter die „Wörter in die Unterbestandteile: Konsonanten und Vokale“, in den Abteilungen 8 bis 11 schließlich wurden Überblickslisten aller Vor-, End- und Nebensilben sowie aller Vokale und Konsonanten erstellt. Die „Abteilung 12a und b“ erbrachte „die Schlußarbeitsanweisungen für die Nachprüfung der alphabetischen Liste und der Liste der nackten Stämme“, 50 Mitarbeiter und „Herr Pastor Koch in Töchtelborn“ waren damit betraut.<sup>30</sup> Kaeding verfasste am Ende die „Urschrift“ des *Häufigkeitwörterbuchs* und besorgte die Drucklegung. Nicht ohne Stolz schreibt er im ersten Absatz seines Vorworts:

Mit dem vorliegenden Werke übergebe ich dem deutschen Volke die Ergebnisse einer mehr als fünfjährigen angestrengten Arbeit vieler Personen mit dem Wunsche, daß die erreichten Feststellungen brauchbare Unterlagen für weitere wissenschaftliche Forschungen bieten mögen. Ich bin überzeugt, daß trotz der aufgewendeten Mühe, nicht alle von den verschiedenen Kreisen zu stellenden Fragen ihre Beantwortung finden werden, man erhält aber überall die Grundlage für die zu bestimmten Zwecken erforderlichen Anschlußarbeiten und Ergänzungen, da der Stoff bis in die kleinsten Einzelheiten zergliedert und übersichtlich zusammengestellt worden ist.<sup>31</sup>

---

**29** Ebd., S. 31f., Hervorhebung im Original.

**30** Ebd., S. 32–37.

**31** Ebd., S. 1.

### 3. Das Archiv des Wörterbuchs

Welch gewaltige Anstrengung die Realisierung eines solchen Projekts bedeutete, lässt sich unschwer erahnen. Mit welchem Eifer und welcher Begeisterung Kaeding – neben seiner Arbeit als Kalkulator der Reichsbank – sich für das Vorhaben engagierte, geht deutlich zwischen den Zeilen (genauer: zwischen den Buchstaben und Zahlen) seines opus magnum hervor. Kaeding scheut auch nicht davor zurück, die ablehnende Haltung oder nicht, zu spät oder nur widerwillig eingebrachte Arbeits- und Verbesserungsvorschläge seitens einiger Stenographenkollegen zu rügen. Auch bedauert er, dass es seitens der zeitgenössischen Sprachwissenschaft kein Interesse und auch keine Unterstützung gab:

Was zunächst die Sprachforschung anlangt, so hätte nach dieser Seite hin noch vieles erreicht werden können, wenn die deutsche Gelehrtenwelt dem Werke von Anfang an mehr Beachtung geschenkt, dem Arbeitsausschusse ihre Wünsche dargelegt und die nötigen Arbeitskräfte und Geldmittel zugeführt hätte. Die ersten Arbeitsanweisungen hätten dementsprechend abgefaßt werden und die Arbeiten ohne wesentliche Vermehrung der Arbeitslast den weitergehenden Wünschen Rechnung tragen können. Da aber keinerlei Beteiligung von Seiten der Gelehrten sich zeigte, so hatte der lediglich für stenographische Zwecke eingesetzte Arbeitsausschuß keine Veranlassung, sich mit Dingen zu befassen, die seinem eigentlichen Wirkungskreise fern standen. Trotzdem bin ich überzeugt, daß auch für die Sprachwissenschaft sehr viel Neues und Wissenswertes aus unserem Werke sich ergeben wird, wenn die weitere Verarbeitung des Stoffes von kundiger Hand geschieht.<sup>32</sup>

Kaeding ahnte und wusste sehr wohl, dass seine Arbeit und jene seiner aberhundert Mitarbeiter unterschiedliche Anschlussstellen für weiterführende Forschungen böte:

Die Buchungsblätter bieten einen ganz vorzüglichen Stoff für denjenigen, dem es auf das Vorkommen bestimmter Wörter in bestimmten Stoffgattungen ankommt. Der Linguist z.B. kann sich Auszüge fertigen lassen aus den ihn besonders anziehenden Abschnitten: Klassiker und Novellisten sowie dem privaten Briefstil. Ebenso kann getrennt festgestellt werden, wie oft die Wörter vorkommen: im juristischen Stoff, im militärischen, im gemischten u.s.w. Dieser Nachweis ist für jedes Wort bis ins kleinste möglich.<sup>33</sup>

Die Daten, die für das *Häufigkeitswörterbuch* erfasst wurden, sind weit umfangreicher, als sie dort veröffentlicht werden konnten. Kaeding war es ein Anlie-

---

<sup>32</sup> Ebd., S. 10.

<sup>33</sup> Ebd., S. 31.

gen, die Wissensbestände, die sich während der Arbeit akkumulierten, für Anschlussarbeiten aufzubewahren und der Nachwelt zu überliefern:

Eine Veröffentlichung des Inhaltes der sämtlichen Buchungsblätter ist natürlich unmöglich. Die Buchungsblätter und die Zählzettel der zusammengesetzten Hauptwörter, die Umschrift der alphabetischen Liste, die sämtlichen Arbeitsanweisungen und eine Sammlung von auf die Untersuchungen bezüglichen Schriftstücken sollen der Königlichen Bibliothek in Berlin übergeben werden.<sup>34</sup>

Das vollständige Archiv des *Häufigkeitswörterbuchs* liegt heute unter der Signatur „Nachlass 394 (Friedrich Wilhelm Kaeding)“ in der Handschriftenabteilung der Staatsbibliothek zu Berlin – Preussischer Kulturbesitz.<sup>35</sup> Kaedings konsequente Arbeitsweise berechtigt zur Annahme, dass er bald nach Abschluss der Arbeit am *Häufigkeitswörterbuch* und nach Erscheinen desselben, also bald nach 1898, die oben genannten Bestände der ehemaligen Königlichen Bibliothek in Berlin übergab. Daraufhin aber musste das Wörterbuch-Archiv über sehr viele Jahrzehnte in Vergessenheit geraten sein, denn erst im Jahre 2001 und im Zuge der Zusammenführung der Handschriftenabteilungen der ehemaligen Deutschen Staatsbibliothek in Berlin (Ost) und der ehemaligen Staatsbibliothek Preussischer Kulturbesitz in Berlin (West) wurde der Bestand als „Nachlass 394 (Friedrich Wilhelm Kaeding)“ formell von der Staatsbibliothek akzessioniert.<sup>36</sup> Weder die Benutzungsgeschichte des Nachlasses noch die ohnedies spärliche Forschung zu Kaeding und seinem Wörterbuchprojekt lassen erkennen, dass jemals irgendjemand – weder Zeitgenossen Kaedings noch Nachgeborene – Einsicht in den Nachlass genommen hätte.

Kaedings Nachlass ist überaus umfangreich und umfasst 153 Kästen.<sup>37</sup> Darin enthalten sind die Archivalien, deren Übergabe an die Königliche Bibliothek Kaeding angekündigt hatte: die Arbeitsanweisungen, die Zählzettel der zusammengesetzten Hauptwörter, die Buchungsblätter, die Nachweise der Korpustexte und weitere Arbeitsunterlagen (darunter auch Korrespondenz) in Zusammenhang mit dem *Häufigkeitswörterbuch*. Anhand der Dokumentation des

---

<sup>34</sup> Ebd.

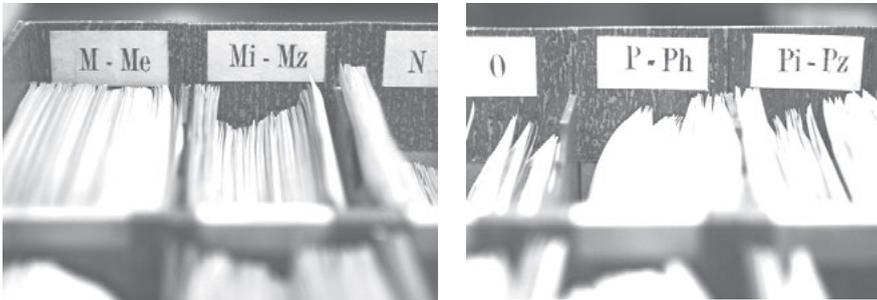
<sup>35</sup> Der Verfasser dankt der Handschriftenabteilung der Staatsbibliothek zu Berlin – Preussischer Kulturbesitz für die Benutzung des umfangreichen Nachlasses von Friedrich Wilhelm Kaeding und für die Erlaubnis, Fotos daraus anzufertigen und diese hier zu veröffentlichen.

<sup>36</sup> Persönliche Mitteilung von Dr. Jutta Weber, stv. Leiterin der Handschriftenabteilung der Staatsbibliothek zu Berlin – Preussischer Kulturbesitz, an den Verfasser am 22.6.2013.

<sup>37</sup> Eine grobe Übersicht über den Bestand vermittelt die Archivdatenbank Kalliope über „Sucheinstieg Bestände“ unter <http://kalliope.staatsbibliothek-berlin.de/> [letzter Zugriff am 16. Januar 2015].

Arbeitsprozesses im *Häufigkeitswörterbuch* (siehe oben) lassen sich der Funktionszusammenhang der Archivalien und der gesamte Arbeits- und Entstehungsprozess des Wörterbuchs rekonstruieren. Der außerordentlich umfangreiche und lückenlose Materialbestand macht die Arbeit am Wörterbuch erfahrbar und erlebbar. Darüber hinaus sind die Materialien in ihrer Objekthaftigkeit auch von besonderem ästhetischem Reiz. Exemplarisch werden einige davon im Folgenden vorgestellt.

Im gedruckten *Häufigkeitswörterbuch* sind die Nachweise der Korpustexte, wohl aus Platzgründen, relativ knapp gehalten.<sup>38</sup> Das Archiv jedoch enthält eine Kartei, bestehend aus mehreren Kästen, in der sämtliche Quellen nachgewiesen sind. Die Abbildungen 3 und 4 zeigen einen Blick in die Kartei über die Buchstabenstrecke M bis Pz.



**Abb. 3 und 4:** Blick in die Kartei der Korpustexte.

Die in ihrer Objekthaftigkeit und Materialität wohl faszinierendsten Archivalien sind die „Zählzettel der zusammengesetzten Hauptwörter“. Jedes einzelne Wort (im Sinne eines ‚Token‘) aus dem Korpus wurde „auf einen besonderen Zählzettel von 3 cm Höhe und 7 cm Länge“ „ausgeschrieben“.<sup>39</sup> Die Zählzettel der einfachen Wörter wurden offensichtlich nicht aufgehoben, wohl aber jene der Komposita. In fünf Kästen mit jeweils acht Laden sind diese Zettel, zu Päckchen geschnürt, aufbewahrt. Öffnet man einen solchen Kasten, gelangt man wie durch Schichten zum einzelnen Wort.

<sup>38</sup> Kaeding, *Häufigkeitswörterbuch*, S. 12–20.

<sup>39</sup> Ebd., S. 22.



Abb. 5: Beschriftung des 1. Kastens der „Zählzettel der zusammengesetzten Hauptwörter“.

Hebt man den Deckel ab, wird das Deckblatt sichtbar, das auf der ersten Lade mit den Wörtern „Von Aalschwanzspekulanten bis Azurblau“ liegt:

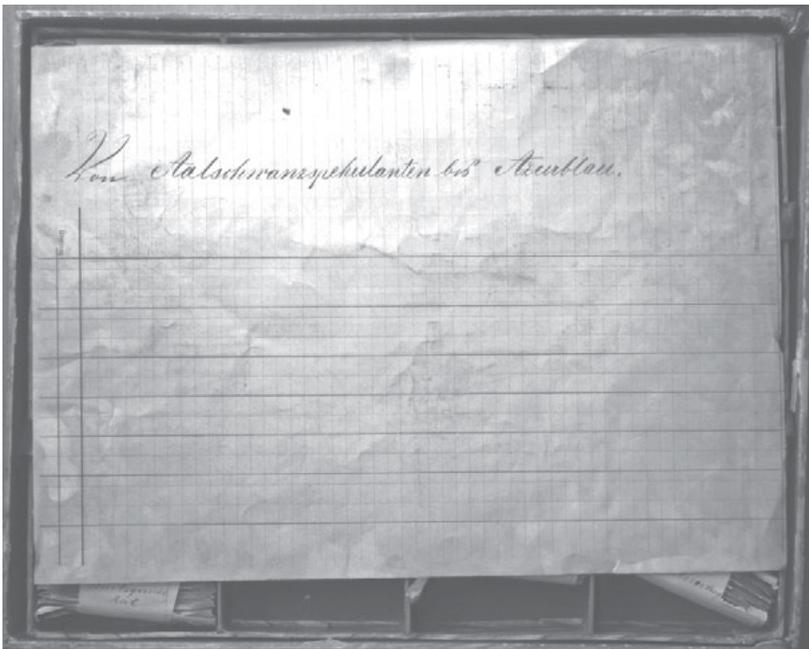
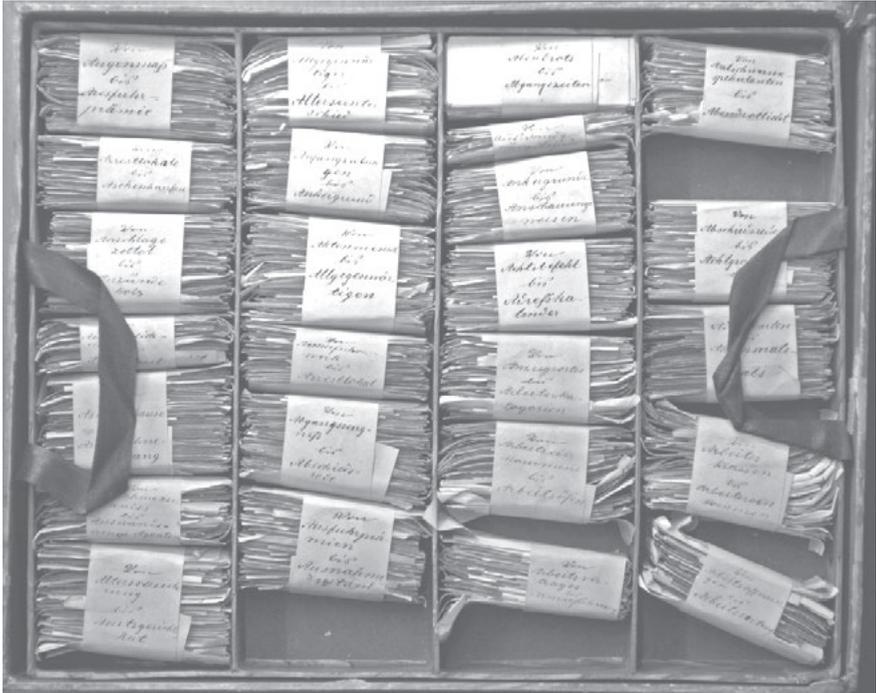


Abb. 6: Deckblatt der ersten Lade: „Von Aalschwanzspekulanten bis Azurblau“.

Die Lade unter dem Deckblatt (Abb. 7) ist in vier Fächer unterteilt, welche die Päckchen mit den 3 mal 7 cm großen Zetteln enthalten. Schwarze Stoffflaschen erleichtern das Herausheben der Lade aus dem Kasten. Acht solcher Laden, die aus steifem Karton gefertigt sind, sind übereinander in jedem der fünf Kästen enthalten. Jede Lade enthält rund 25 bis 35 solcher Päckchen von unterschiedlichem Umfang.



**Abb. 7:** Blick auf die erste Lade.

Die einzelnen Päckchen werden von verklebten Papierlaschen zusammengehalten. Alle Laschen sind beschriftet und benennen den jeweiligen Inhalt. Abb. 8 zeigt das Päckchen „Von Aalschwanzspekulanten bis Abendrotlicht“. Es ist in alphabetischer Reihenfolge das erste Päckchen der ersten Lade, dort (Abb. 7) befindet es sich in der rechten oberen Ecke.



Abb. 8: Das Päckchen der Zählzettel „Von Aalschwanzspekulanten bis Abendrotlicht“.

Der oberste Zettel enthält das Wort „Aalschwanzspekulanten“ (Abb. 9). Die Ziffer „1“ (in roter Tinte) bedeutet, dass ein Beleg für dieses Wort in der „Abteilung 1“, die für das „Ausschreiben“ der Wörter aus dem „Zählstoff“ zuständig war, gefunden wurde. Der Stempel (in blauer Farbe) weist nach, dass dieses Wort im Rahmen der „Buchung“, die in „Abteilung 3“ anhand der „Zählzettel“ und „Einheitszettel“ erfolgte, im „Buch No. 042“ erfasst wurde.

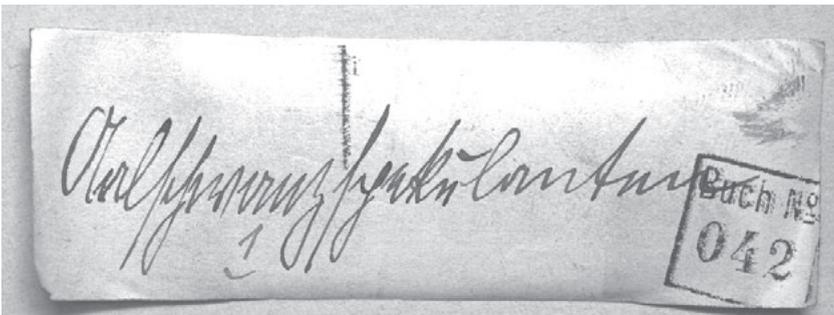


Abb. 9: Zählzettel „Aalschwanzspekulanten“.

Mittels der „Buchungsblätter“ erfolgte in „Abteilung 3“ gewissermaßen die Lemmatisierung der Tokens. Kaedings Nachlass enthält sämtliche Buchungsblätter aller 258.173 Wörter, die im Projektverlauf erfasst wurden, also auch die Buchungsblätter der Wörter mit einer Häufigkeit von weniger als 4, die aus Platz- und Kostengründen im *Häufigkeitswörterbuch* nicht nachgewiesen werden konnten.<sup>40</sup> Abb. 10 zeigt exemplarisch das Buchungsblatt für das Wort

<sup>40</sup> Ebd., S. 43: „Wollte man den ganzen Stoff abdrucken, [...] so würde [...] das Ganze [...] 298 Druckbogen erfordern.“ „Die ganz geringen Häufigkeiten 1, 2 und 3 sind weggelassen, weil eine

„Kaiser“ mit der Gliederung der Fundstellen nach den „Wissensgebieten“ (juristisch, kaufmännisch, theologisch etc.).

*Kaiser*

Gesamt  
2474

Häufigkeit auf

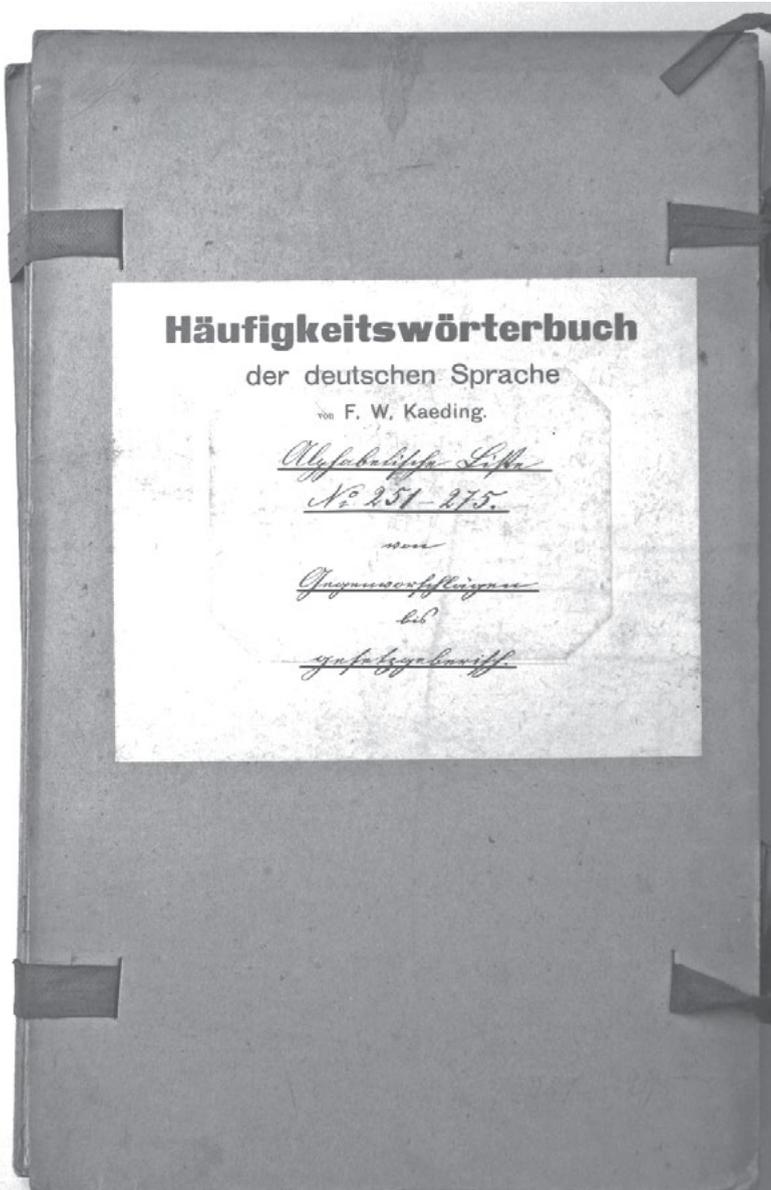
Nr.	Stoff	100 000 Wörter	Nr.	Stoff	100 000 Wörter	Nr.	Stoff	200 000 Wörter	Nr.	Stoff	2319 Wörter	
1.	juristisch	1	Uebertrog juristisch	591	1	Uebertrog juristisch	1164	1	Uebertrog	2319	2319	
2.	kaufmännisch	2	Uebertrog kaufmännisch	591	1	Uebertrog kaufmännisch	1164	1	Uebertrog	2319	2319	
3.	theologisch	3	Uebertrog theologisch	591	1	Uebertrog theologisch	1164	1	Uebertrog	2319	2319	
4.	sonstige	4	Uebertrog sonstige	591	1	Uebertrog sonstige	1164	1	Uebertrog	2319	2319	
5.	sonstige	5	Uebertrog sonstige	591	1	Uebertrog sonstige	1164	1	Uebertrog	2319	2319	
6.	sonstige	6	Uebertrog sonstige	591	1	Uebertrog sonstige	1164	1	Uebertrog	2319	2319	
7.	sonstige	7	Uebertrog sonstige	591	1	Uebertrog sonstige	1164	1	Uebertrog	2319	2319	
8.	sonstige	8	Uebertrog sonstige	591	1	Uebertrog sonstige	1164	1	Uebertrog	2319	2319	
9.	sonstige	9	Uebertrog sonstige	591	1	Uebertrog sonstige	1164	1	Uebertrog	2319	2319	
10.	sonstige	10	Uebertrog sonstige	591	1	Uebertrog sonstige	1164	1	Uebertrog	2319	2319	
11.	sonstige	11	Uebertrog sonstige	591	1	Uebertrog sonstige	1164	1	Uebertrog	2319	2319	
12.	sonstige	12	Uebertrog sonstige	591	1	Uebertrog sonstige	1164	1	Uebertrog	2319	2319	
13.	sonstige	13	Uebertrog sonstige	591	1	Uebertrog sonstige	1164	1	Uebertrog	2319	2319	
14.	sonstige	14	Uebertrog sonstige	591	1	Uebertrog sonstige	1164	1	Uebertrog	2319	2319	
15.	sonstige	15	Uebertrog sonstige	591	1	Uebertrog sonstige	1164	1	Uebertrog	2319	2319	
16.	sonstige	16	Uebertrog sonstige	591	1	Uebertrog sonstige	1164	1	Uebertrog	2319	2319	
17.	sonstige	17	Uebertrog sonstige	591	1	Uebertrog sonstige	1164	1	Uebertrog	2319	2319	
18.	sonstige	18	Uebertrog sonstige	591	1	Uebertrog sonstige	1164	1	Uebertrog	2319	2319	
19.	sonstige	19	Uebertrog sonstige	591	1	Uebertrog sonstige	1164	1	Uebertrog	2319	2319	
20.	sonstige	20	Uebertrog sonstige	591	1	Uebertrog sonstige	1164	1	Uebertrog	2319	2319	
21.	sonstige	21	Uebertrog sonstige	591	1	Uebertrog sonstige	1164	1	Uebertrog	2319	2319	
22.	sonstige	22	Uebertrog sonstige	591	1	Uebertrog sonstige	1164	1	Uebertrog	2319	2319	
23.	sonstige	23	Uebertrog sonstige	591	1	Uebertrog sonstige	1164	1	Uebertrog	2319	2319	
24.	sonstige	24	Uebertrog sonstige	591	1	Uebertrog sonstige	1164	1	Uebertrog	2319	2319	
25.	sonstige	25	Uebertrog sonstige	591	1	Uebertrog sonstige	1164	1	Uebertrog	2319	2319	
Uebertrog		591	Uebertrog		1164	Uebertrog		2319	Summe			2474

Abb. 10: Buchungsblatt für das Wort „Kaiser“.

Einer der letzten Schritte vor der Fertigstellung des *Häufigkeitswörterbuchs* war die Erstellung der „Alphabetische[n] Liste“ über den gesamten „Zählstoff“ in der „Abteilung 4“. Es handelt sich hierbei um jene 817 Hefte, die alle Wörter in

so seltene Anwendung eines Wortes doch nur zufällig ist und für die weiteren Schlußfolgerungen der stenographischen Wissenschaft kaum in Betracht kommt. Für die Sprachwissenschaft haben allerdings gerade die seltener vorkommenden Wörter unter Umständen einen besonders hohen Wert; da dem Arbeitsausschuß aber aus diesen Kreisen keine Geldmittel zur Bestreitung der erhöhten Druckkosten zugeführt wurden, so konnten die über das stenographische Bedürfnis hinausgehenden Wünsche keine Berücksichtigung finden.“ Kaeding nennt hier (S. 43) auch die zahlreichen Institutionen und Personen, bei denen Druckkostenzuschüsse beantragt wurden; die Bescheide waren bis auf sehr wenige Ausnahmen negativ.

alphabetischer Reihenfolge verzeichnen. Die „Alphabetische Liste“ besteht aus 35 Mappen, die in der Regel jeweils 24 fortlaufend nummerierte Hefte enthalten. Abb. 11 zeigt die Mappe der „Alphabetischen Liste“ mit den Heften der Nummern 251–275 über die Strecke der Wörter von „Gegenvorschlägen“ bis „gesetzgeberisch“. Abb. 12 zeigt als Beispiel die Seite der „Alphabetischen Liste“ zu den Wörtern „Geheimensanitätsrat“ bis „Geheimhaltungen“.



**Abb. 11:** Mappe der „Alphabetischen Liste“ mit den Heften 251 bis 175 von „Gegenvorschlägen“ bis „gesetzgeberisch“.

Bezeichnung des Wortes	Häufigkeit auf 10 000 235 Wörter oder 20 Millionen Silben	Gesamt- Silbenzahl (Ergebnis aus Spalte 1 + 2)	Verteilungen in Abteilungen und Zusammenfassungen als			Gesamt- Häufigkeit (Summe der Spalten 2, 4, 5, 6)	Häufigkeit auf 100 000 Silben (Spalte 7 geteilt durch 200)
			Erstes Wort	Mittel- Wort	Schluss- Wort		
1.	2.	3.	4.	5.	6.	7.	8.
Übertrag:	994	3057	1114		46	2107	
<u>Gefahrenparitätswal</u>	1	7				1	
<u>Gefahren</u> (58)	1	162	1			58	} 158
<u>Gefahren</u>	157	389				103	
<u>Gefahrenwale</u>	2	10				2	
<u>Gefahrenrot</u>	227	1048				227	
<u>Gefahren</u>	1	4				1	
<u>Gefahrenkommunizierend</u>	1	8				1	
<u>Gefahren</u>	40	120				40	
<u>Gefahrensüß</u>	1	3				1	
<u>Gefahrensfallener</u>	2	12				2	
<u>Gefahrensöld</u>	1	3				1	
<u>Gefahrenfallbar</u>	1	4				1	
<u>Gefahrensfallen</u>	2	8				2	
<u>Gefahrenfallbar</u>	1	4				1	
<u>Gefahrenfallung</u>	28	112				28	
<u>Gefahrenfallungen</u>	1	5				1	
Übertrag:	1520	4976	1755		116	2720	

Abb. 12: Beispiel einer Seite aus der alphabetischen Liste.

## 4. Rezeption und wissenschaftsgeschichtlicher Zusammenhang

Wenngleich Kaeding mit großer Geste „dem deutschen Volke die Ergebnisse einer mehr als fünfjährigen angestrengten Arbeit vieler Personen mit dem Wunsche“ übergab, „daß die erreichten Feststellungen brauchbare Unterlagen für weitere wissenschaftliche Forschungen bieten mögen“,<sup>41</sup> blieb die Rezeption seines *Häufigkeitswörterbuchs* über die ‚stenographische Wissenschaft‘ hinaus<sup>42</sup> und abgesehen von jener frühen (1928) durch den späteren Stanford-Germanisten Bayard Quincy Morgan,<sup>43</sup> sehr lange sehr verhalten.<sup>44</sup> Zwei Aspekte waren vermutlich dafür ausschlaggebend. Kaeding war Dilettant, und zwar im besten und wahrsten Sinne des Wortes. Zum einen erschwerte also wohl ein akademisches Selbstverständnis es, dass „die deutsche Gelehrtenwelt“ mit dem Nicht-Akademiker Kaeding auf Augenhöhe kooperieren konnte. Zum anderen ist auch eine tatsächliche mangelhafte Anschlussfähigkeit Kaedings an die „Sprachforschung“ zu diagnostizieren: Zu fern lag Kaedings Ansatz vom zeitgenössischen sprachhistorischen oder phonologischen Wissensstand und Forschungsinteresse, zu stark konzentrierte er sich auf die Belange der Stenographie, die zudem nie ein akademisches Fach war.

Erst in der zweiten Hälfte des 20. Jahrhunderts, auch im Zuge der Historisierung seiner Person und seines Wörterbuchprojekts, wird die Linguistik auf Kaedings Arbeit aufmerksam und seine Leistung anerkannt. Prominente Rezeptionsbeispiele sind der Nachdruck (in Auszügen) als Band 4 (1963) der *Grundla-*

---

41 Ebd., S. 1.

42 Eine Übersicht über die durch Kaeding angeregte Weiterentwicklung der Stenographie sowie über linguistische Forschungen, die an Kaeding anschließen, liefert die Bibliographie in Best, Friedrich Wilhelm Kaeding, S. 84–86.

43 Bayard Quincy Morgan, German frequency word book. Based on Kaeding’s Häufigkeitswörterbuch der deutschen Sprache, New York 1928.

44 Darauf wies zuletzt auch Best, Friedrich Wilhelm Kaeding, S. 82, hin. Als Symptom dafür, dass sich bis heute auch eine kulturwissenschaftlich orientierte Linguistik nicht mit Kaeding beschäftigt, darf der Umstand gesehen werden, dass etwa Ulrike Haß-Zumkehr weder in ihrer Monographie *Deutsche Wörterbücher – Brennpunkt von Sprach- und Kulturgeschichte*, Berlin und New York 2001, noch in ihrem Sammelwerk *Große Lexika und Wörterbücher Europas. Europäische Enzyklopädien und Wörterbücher in historischen Portraits*, Berlin und Boston 2012, auf Häufigkeitswörterbücher im Allgemeinen und auf Kaeding im Besonderen eingeht. Auch Michael Schläefer, *Lexikologie und Lexikographie. Eine Einführung am Beispiel deutscher Wörterbücher*, Berlin 2002 (Grundlagen der Germanistik 40), geht auf Häufigkeitswörterbücher nicht ein und erwähnt Kaeding lediglich sehr knapp.

genstudien aus Kybernetik und Geisteswissenschaft<sup>45</sup> und, ausgehend vom Verständnis, dass Frequenz das „Maß der Nützlichkeit“ ist,<sup>46</sup> die Verwendung im Bereich der Sprachdidaktik z.B. durch Wolf Dieter Ortmann.<sup>47</sup> In der gegenwärtigen Quantitativen Linguistik, einem jungen Teilgebiet der Linguistik, wird regelmäßig auf Kaeding Bezug genommen. Hervorgehoben wird u.a. der Umstand, dass erst die fortgeschrittene computergestützte Linguistik in der Lage ist, für Häufigkeitsuntersuchungen größere Korpora als das Kaeding-Korpus zu berücksichtigen: „In actual practice however one can observe that up till now few of the published frequency dictionaries exceed Kaeding’s corpus of 11 million wordtokens“, wie Willy Martin in seiner grundlegenden Darstellung *The Frequency Dictionary* festhält.<sup>48</sup> Erst das *Frequenzwörterbuch der deutschen Zeitungssprache* (1972–1977) von Inger Rosengren,<sup>49</sup> das mithilfe von Hochleistungstechnik am Rechenzentrum der Universität Hamburg erarbeitet wurde, übertraf mit ca. 12,4 Millionen Wörtern<sup>50</sup> geringfügig Kaedings 11 Millionen Wörter großes Korpus.<sup>51</sup>

---

45 Kaeding, Häufigkeitswörterbuch 1963.

46 Inger Rosengren, Ein Frequenzwörterbuch der deutschen Zeitungssprache. Die Welt, Süddeutsche Zeitung, 2 Bde., Lund 1972–1977, hier Bd. 1, S. XXV.

47 Wolf Dieter Ortmann, Hochfrequente deutsche Wortformen, 4 Teile, München 1975–1978; ders., Minimalpaare im Deutschen. Typen, Häufigkeiten, Übungsbeispiele, rechnerortiert anhand von 7.995 hochfrequenten Wortformen der Kaeding-Zählung, mit einem Anhang: Reimlexikon zur Kaeding-Wortliste, München 1981.

48 Martin, *The Frequency Dictionary*, S. 1317.

49 Rosengren, *Frequenzwörterbuch*.

50 Ebd., Bd. 1, S. XXIII. Das Korpus besteht aus repräsentativen Stichproben aus *Die Welt* und *Süddeutsche Zeitung*, Zeitraum: 1.11.1966 bis 30.10.1967.

51 Weitere Frequenzwörterbücher für das Deutsche sind Hans-Heinrich Wängler, *Rangwörterbuch hochdeutscher Umgangssprache*, Marburg 1963; Arno Ruoff, *Häufigkeitswörterbuch gesprochener Sprache, gesondert nach Wortarten, alphabetisch, rückläufig alphabetisch und nach Häufigkeit geordnet*, 2. Aufl., Tübingen 1990 und das *Wortschatz-Portal* der Universität Leipzig, das auch Angaben zu Wörterhäufigkeiten bereitstellt: <http://wortschatz.uni-leipzig.de/> [letzter Zugriff am 16. Januar 2015]. Erratisch und singulär ist die *Deutsche Sprachstatistik* (1964) des Lehrers und Privatgelehrten Helmut Meier, der vier Jahrzehnte Sammelarbeit vorausgegangen waren. Am 19. Dezember 1964 verlieh die Philosophische Fakultät der Universität Hamburg Meier für seine „Lebensarbeit“ die Ehrendoktorwürde. Aus Freude und zum Dank widmete Meier der Universität Hamburg einen zweiten Band, der 1967 zusammen mit dem ersten Band als Manuskript-Reprint bei Olms erschien: Helmut Meier, *Deutsche Sprachstatistik*, Hildesheim 1967. Lutz Mackensen dazu in seinem Geleitwort, S. V: „Dieses Buch mutet wie ein Heldenlied an, Beispiel nicht nur eines unbeirrbar geistigen Ringens, sondern auch einer einzigartigen Kraft des Herzens. [...] Bücher wie dieses werden alle Jahrhunderte einmal geschrieben.“

Nach gegenwärtigem Kenntnis- und Verständnisstand ist Kaedings *Häufigkeitswörterbuch der deutschen Sprache* weltweit das zweite (oder dritte) Häufigkeitswörterbuch für eine Sprache. Als „starting point“ in der Geschichte der Häufigkeitswörterbücher gilt William Gambles Häufigkeitsliste chinesischer Schriftzeichen,<sup>52</sup> die 1861 unter dem Titel *Two lists of selected characters containing all in the Bible and twenty-seven other books* in Shanghai erschien.<sup>53</sup> Das Buch, das als Handreichung für Setzer chinesischer Schriftzeichen gedacht war, damit diese besser abschätzen konnten, welche Zeichen in welcher Menge vorzuhalten sind, erschien wahrscheinlich ist sehr geringer Stückzahl und erhob auch nicht den Anspruch eines Referenzwerks. So ist es heute ein Rarum, das weltweit in sehr wenigen Bibliotheken verfügbar ist. Das hat bis heute zur Folge, dass die Forschung Gamble vorbehaltlos würdigt, sein Buch aber nie in Händen hält. Kaeding war einer der wenigen, der nachweislich Gambles Werk in der Hand hatte und auch daraus zitiert.<sup>54</sup>

Emmerich Kelih weist in seiner Dissertation *Geschichte der Anwendung quantitativer Verfahren in der russischen Sprach- und Literaturwissenschaft* (2008) darauf hin, dass 1894, also vier Jahre vor Kaeding, „das erste russische Häufigkeitswörterbuch“ erschienen sei.<sup>55</sup> Es handelt sich dabei um die Arbeit des ukrainischen Pädagogen Viktor Nikolajewitsch Kunickij,<sup>56</sup> dem als Korpus ein literarischer Text, die Komödie *Gore ot uma / Verstand schafft Leiden* (1824) des russischen Schriftstellers Alexander Sergejewitsch Griboedov, zugrunde liegt. „Das erklärte Ziel von Kunickij ist es, eine möglichst exakte und umfassende lexikographische Bestandsaufnahme des in der Komödie verwendeten Wortschatzes zu leisten.“<sup>57</sup> Legt man die zwei wesentlichen Kriterien für ein Häufigkeitswörterbuch, Unabhängigkeit von individuellem Sprachgebrauch und ausreichend großes und über unterschiedliche Textsorten gestreutes Kor-

---

52 Martin, *The Frequency Dictionary*, S. 1319.

53 Gamble, *Two lists*.

54 Kaeding, *Häufigkeitswörterbuch*, S. 669: „Gamble hat dann in einem, soweit mir bekannt, sehr selten gewordenen Büchlein, das mir einmal ein glücklicher Zufall in Paris in die Hände spielte, als ich dort danach fahndete, diese rund 5.000 Schriftzeichen mit Häufigkeitszahlen versehen und nach der Häufigkeit geordnet. Einige Hauptergebnisse veranschaulicht das Folgende [...]“. Auf S. 670 folgt sodann eine Tabelle mit Daten aus Gambles *Two lists*.

55 Emmerich Kelih, *Geschichte der Anwendung quantitativer Verfahren in der russischen Sprach- und Literaturwissenschaft*, zugl.: Diss., Univ. Graz 2007, Hamburg 2008 (Studien zur Slawistik 19), S. 42.

56 V[iktor] N[ikolajewitsch] Kunickij, *Jazyk i slog komedii Gore ot uma. K stolétiju dnja roždenija A.S. Griboėdova 4 Janvarja 1795 g. – 4 Janvarja 1895 g.*, Kiev 1894.

57 Kelih, *Geschichte der Anwendung quantitativer Verfahren*, S. 42.

pus,<sup>58</sup> zugrunde, ist Kunickijs Arbeit nicht als Häufigkeitwörterbuch in eigentlichem Sinn, also einer Sprache, sondern als Häufigkeitwörterbuch zu einem einzelnen Autor, ähnlich einer Konkordanz, anzusehen.

Überhaupt stehen Konkordanzen in funktioneller Verwandtschaft zu Häufigkeitwörterbüchern, und sie gehen diesen, historisch betrachtet, voraus, worauf Manlio Cortelazzo und Arjuna Tuzzi in ihrem Buch *Metodi statistici applicati all'italiano* (2008), insbesondere im Abschnitt zur Geschichte der Häufigkeitsforschung zur italienischen Sprache und Literatur, aufmerksam machen.<sup>59</sup> Konkordanzen gab es lange vor den Häufigkeitwörterbüchern (bereits aus dem 13. Jahrhundert sind Bibel-Konkordanzen bekannt), sie beziehen sich in der Regel auf hoch kanonische Werke und Autoren wie die Bibel, Dante, Shakespeare oder Goethe, sie sind heuristische Hilfsmittel für die hermeneutische Arbeit und dienen als Heuristiken zur Entwicklung und Bearbeitung philologischer Fragestellungen. Beispiele für Konkordanzen aus der germanistischen Literaturwissenschaft sind die *Indices zur deutschen Literatur*, die ab 1968 im Athenäum-Verlag, ab 1982 bei Niemeyer erschienen. Der erste Band in dieser Reihe ist der Index zu Novalis' *Heinrich von Ofterdingen*.<sup>60</sup> Während Häufigkeitwörterbücher als linguistische Hilfsmittel gelten, sind Konkordanzen Nachschlagewerke von literaturwissenschaftlichem Interesse.

In der Linguistik waren Frequenzwörterbücher mehrfach Gegenstand der Forschung.<sup>61</sup> Besonders aus den Jahren um 1970, und zwar v.a. aus dem Kontext der Bestrebungen zu maschineller Übersetzung, sind hoch spezialisierte, auch dreisprachige Häufigkeitwörterbücher überliefert.<sup>62</sup> Martin stellt fünf For-

---

**58** Martin, *The Frequency Dictionary*, S. 1314; Hadumod Bußmann, *Lexikon der Sprachwissenschaft*, 3. Aufl., Stuttgart 2002, S. 270.

**59** Manlio Cortelazzo und Arjuna Tuzzi, *Metodi statistici applicati all'italiano*, Bologna 2008, S. 80–124, hier S. 83.

**60** Helmut Schanze, *Index zu Novalis' Heinrich von Ofterdingen*, Frankfurt a.M. 1968 (*Indices zur deutschen Literatur* 1) [Reprint Nendeln 1979].

**61** Duncan Harkin, *The history of word counts*, in: *Babel* 3 (1957), S. 113–124; Rolf-Dietrich Keil, *Einheitliche Methoden in der Lexikometrie*, in: *International Review of Applied Linguistics in Language Teaching* 3 (1965), S. 95–122; P. M. Alekseev, *Statistische Lexikographie. Zur Typologie, Erstellung und Anwendung von Frequenzwörterbüchern*. Lehrbuch, Übersetzt von Werner Lehfeldt, Bochum 1984 (*Quantitative linguistics* 22); Martin, *The Frequency Dictionary*; mit dem Thema „Frequenzeffekte“ beschäftigt sich Heft 169 (2013) der Zeitschrift für Literaturwissenschaft und Linguistik (LiLi).

**62** In der Reihe *Häufigkeitwörterbuch Russisch, Englisch, Französisch* erschienen im Leipziger Verlag Enzyklopädie ab 1970 die Bände *Fachwortschatz Medizin* (1970, 1986 in 7., unveränderter Aufl.), *Fachwortschatz Mathematik*, *Fachwortschatz Chemie*, *Fachwortschatz Physik*, *Fachwortschatz Tierproduktion*, *Veterinärmedizin* und *Fachwortschatz Bauwesen*. Eine Übersicht über Häufigkeitwörterbücher zu technischen und ingenieurwissenschaftlichen

schungs- bzw. Anwendungszusammenhänge für Häufigkeitswörterbücher fest: pragmatisch (als Information über Verwendungshäufigkeiten von Wörtern), sprachdidaktisch (um einen Basiswortschatz aus hoch frequenten Wörtern zu erstellen), psycholinguistisch (zur Vorbereitung von Experimenten, z.B. Worterkennungstests), stilometrisch („as a kind of norm or yardstick against which possible deviations can be evaluated“) und als Daten-Input für künstliche Sprachgenerierung.<sup>63</sup> Die Geschichte der Häufigkeitswörter gliedert er in drei Phasen:

- eine erste Phase von den Anfängen bis 1920, die dadurch gekennzeichnet ist, dass Häufigkeitswörterbücher oft nicht für linguistische oder literaturwissenschaftliche, sondern für andere Zwecke erstellt wurden (z.B. als Handreichung für Setzer chinesischer Schriftzeichen bei Gamble, zur Entwicklung der Kurzschrift bei Kaeding);

- eine zweite Phase von 1920 bis 1960, in der Häufigkeitswörterbücher vor allem für Belange der Sprachendidaktik erstellt wurden (für das Spanische gab es als erstes Listen dieser Art);<sup>64</sup>

- eine dritte Phase von 1960 bis zur Gegenwart, die vor allem an die Entwicklung von Korpuslinguistik und Computertechnik gekoppelt ist.<sup>65</sup>

Eine literaturwissenschaftliche, kultur- und theoriegeschichtliche Verortung von Häufigkeitswörterbüchern (auch von Konkordanzen) gibt es bislang nicht. Eine solche wäre eng an die Geschichte der Lexikographie, aber auch die Geschichte der Anwendung quantitativer Verfahren in Literatur- und Sprachwissenschaft geknüpft und würde auf Computational Stylistics in Digital Humanities verweisen.

Pathos ist wahrscheinlich auch etwas, das sich in seiner Wirkung konfiguriert: Je dichter der Verständniszusammenhang um einen Gegenstand wird, umso weniger wirkt dieser pathetisch und umso mehr nähert er sich der Verständlichkeit (und auch der Selbstverständlichkeit). Kaedings *Häufigkeitwörterbuch der deutschen Sprache* ist nicht nur ein wissenschaftsgeschichtlich interessierendes Beispiel für ein außerordentliches Wörterbuchvorhaben, des-

---

Fachsprachen unterschiedlicher Sprachen bietet Alekseev, *Statistische Lexikographie*, S. 110–117.

<sup>63</sup> Martin, *The Frequency Dictionary*, S. 1319.

<sup>64</sup> Hayward Keniston, *Common Words in Spanish*, in: *Hispania* 3 (1920), S. 85–96; ders., *Spanish idiom list. Selected on the basis of range and frequency of occurrence*, New York 1929 (Publications of the American and Canadian Committees on modern languages 11).

<sup>65</sup> Martin, *The Frequency Dictionary*, S. 1319. Vgl. auch Cortelazzo, Tuzzi, *Metodi statistici*, S. 80.

sen Ansatz in seiner Zeit in hohem Maße innovativ war, das aber gleichzeitig ob seiner eher geringen Anschlussfähigkeit an die zeitgenössische Philologie bereits etwas isoliert dastand. Kaedings monumentales *Häufigkeitswörterbuch* ist auch Ausdruck von Pathos insofern, als es mit großer Geste hohen Gültigkeitsanspruch behauptet, der im Laufe des vergangenen Jahrhunderts auch aufgehoben wurde in der weiteren Entwicklung von Verständnisweisen korpusbasierter Frequenzeffekte.

