

Guido Mensching, Jean-Yves Lalande,  
Jürgen Hermes und Claes Neufeind (Hrsg.)

# Sprache – Mensch – Maschine

Beiträge zu  
Sprache und Sprachwissenschaft,  
Computerlinguistik und Informationstechnologie

für Jürgen Rolshoven aus Anlass seines  
sechsendsechzigsten Geburtstages

**Köln 2018**

Guido Mensching,<sup>1,2</sup> Jean-Yves Lalande,<sup>2</sup> Jürgen Hermes<sup>3</sup> und Claes Neuefeind<sup>3</sup> (Hrsg.). 2018. *Sprache – Mensch – Maschine. Beiträge zu Sprache und Sprachwissenschaft, Computerlinguistik und Informationstechnologie für Jürgen Rolsboven aus Anlass seines sechsundsechzigsten Geburtstages*. Köln.

<sup>1</sup> Georg-August-Universität Göttingen, Seminar für Romanische Philologie

<sup>2</sup> ehemals Universität zu Köln, Sprachliche Informationsverarbeitung

<sup>3</sup> Universität zu Köln, Sprachliche Informationsverarbeitung

Bibliografische Information der Deutschen Nationalbibliothek  
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der  
Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im  
Internet über <http://dnb.dnb.de> abrufbar.

ISBN: 978-3-00-061795-9 (Digital)

KUPS - Kölner UniversitätsPublikationsServer  
<https://kups.ub.uni-koeln.de/>

Köln 2018

# Inhaltsverzeichnis

Einleitung GUIDO MENSCHING, JEAN-YVES LALANDE, JÜRGEN HERMES, CLAES NEUEFEIND	7
---	---

## ROMANISCHE PHILOLOGIE UND LINGUISTIK

---

Zur Vor- und Frühgeschichte des romanischen Infinitivs mit explizitem Subjekt HANS DIETER BORK	23
--	----

Das Aromunische – ein Dialekt des Rumänischen? WOLFGANG DAHMEN	43
---	----

Gebundene Rede und fiktiver Diskurs: Was macht die Dichtung zur Dichtung? ANDREAS KABLITZ	57
---	----

Instrumente der Bedeutungsintensivierung im Bündnerromanischen FLORENTIN LUTZ	105
---	-----

Die ‘Fauna des Theaters’ von Calderón de la Barca im Kontext ihrer Zeit MANFRED TIETZ	121
---	-----

## ALLGEMEINE SPRACHWISSENSCHAFT

---

Die (sub-)kategoriale Distribution von Anglizismen in deutschen Tweets HEIKE BAESKOW	159
--	-----

Para una morfosintaxis del infinitivo en la “Indogermanische Grammatik”: Sincronía, comparación y reconstrucción JOSÉ LUÍS GARCÍA-RAMÓN	203
---	-----

προθέλυμος, θαύσια und ein griechisches u-Präsens  
DANIEL KÖLLIGAN 219

„Same same but different“: Scrambling-Effekte  
in einigen europäischen Sprachen  
JÜRGEN LENERZ 231

## COMPUTERLINGUISTIK UND INFORMATIONSTECHNOLOGIE

---

Permutiert, klassifiziert, kommasepariert: Computerlinguistik  
im Umfeld medizinischer Klassifikationen  
CHRISTOPH BENDEN 247

Bild, Schrift, Unicode  
CHRISTA DÜRSCHIED 269

Zu den zwei Grundbausteinen der Salzburger  
Dialektometrie: Arbeitskarten und Taxat(areal)e  
HANS GOEBL 287

Simulation of a Functional Grammar in Prolog  
PAUL OTTO SAMUELSDORFF 311

„Strings and Structures“ – eine Begegnung von  
Computerlinguistik und Bioinformatik  
THOMAS WIEHE 327

# Zu den zwei Grundbausteinen der Salzburger Dialektometrie: *Arbeitskarten* und *Taxat(areal)e*

Hans Goebel

*Lieber Jürgen!*

*Unser erstes Zusammentreffen fand vor 40 Jahren (= 2017 –1977) in Romanischbünden statt, und zwar im Zeichen des Surselvischen. Schon damals wurden große Parallelen zwischen unser beider Interessenslagen deutlich, die sich von der „guten alten Romanistik“ – auf ewig unvergesslich werden mir Deine aromunischen Abenteuer sein! – bis zu den damals brandneuen sowie hinsichtlich Funktion und Effizienz noch sehr wackelig dastehenden „Computern“ alias „Rechnern“ erstreckten. Die vorliegenden Zeilen beziehen sich auf mit diesen „Teufelsgeräten“ durchgeführte Forschungen, die aber etwas jünger sind als unsere Freundschaft.*

*Ad multos faustos felicesque annos!*

## 1. Vorbemerkung

Die zwei im Titel erwähnten „Grundbausteine“ der in Salzburg praktizierten Dialektometrie (DM) – *Arbeitskarten* und *Taxate* bzw. *Taxatareale* – beziehen sich auf jene nominal skalierten Vektoren bzw. Elemente, aus denen die in Salzburg verwendeten Datenmatrizen bestehen.<sup>1</sup> Allesamt wurden sie aus diversen – nicht nur romanischen – Sprachatlanten im Wege einer meist von Hand durchgeführten Mess- und Klassifizierungsprozedur abgeleitet, die vor vielen Jahren<sup>2</sup> den Namen *Taxierung* erhalten hat. Werfen wir zur Klarstellung gleich einen Blick auf die bunte Karte 1, die – als Teil der in den Jahren 1995–1999 durchgeführten Dialektometrisierung des französischen Sprachatlantes ALF<sup>3</sup> – ein diesbezüglich sehr brauchbares Muster darstellt:

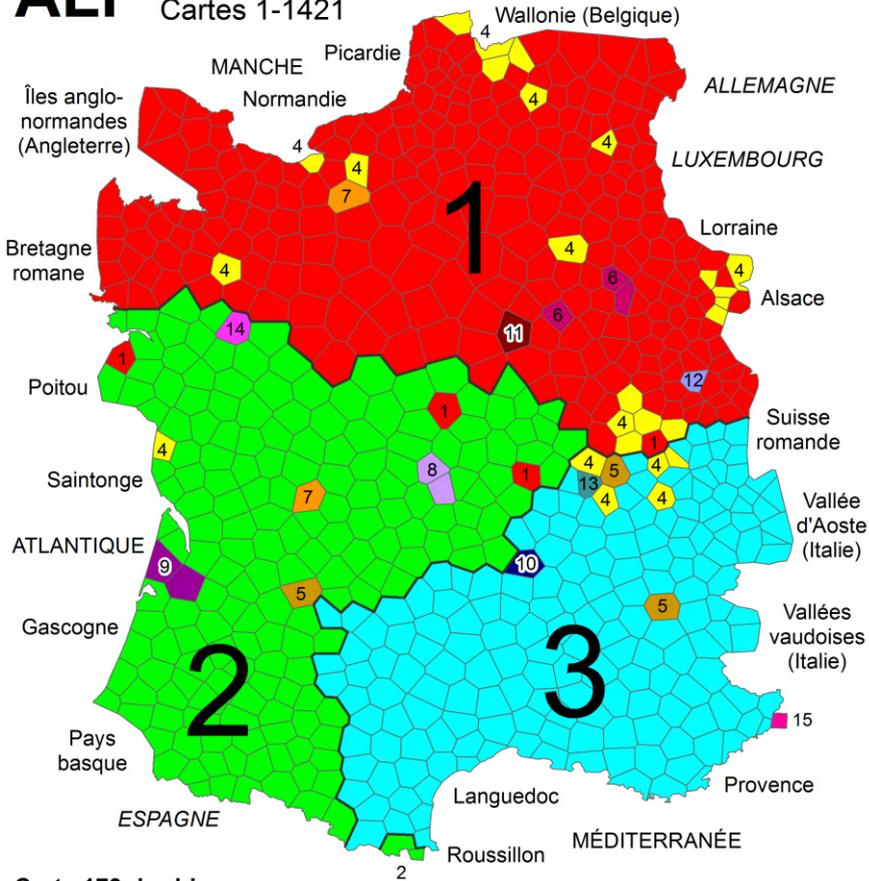
---

<sup>1</sup> Zu den vorliegenden Darstellungen existiert eine argumentativ leicht anders getönte englische Vorgänger-Version: Goebel/Smečka (2016).

<sup>2</sup> Dies geschah schon in den 1970er Jahren: siehe dazu Goebel (1984:Bd. I, passim).

<sup>3</sup> Siehe dazu unsere Beiträge Goebel (2002a, 2003, 2005, 2006, 2007b, 2009, 2010 und 2012).

# ALF Série A: Cartes 1-1421



## Carte 173: brebis

1 brebis [245]	5 agnello [3]	9 tyoc [2]	13 marote [1]
2 ouaille [177]	6 goerlette [3]	10 boebina [1]	14 mère mouton [1]
3 fedo [175]	7 berc [2]	11 canette [1]	15 pecora [1]
4 mouton [26]	8 gode [2]	12 foueyotte [1]	

Karte 1: Zu dialektometrischen Zwecken erstellte Arbeitskarte. Datenbasis: Karte 173 (la brebis 'das Mutterschap') des französischen Sprachatlases ALF. Es wurden 15 lexikalisch relevante Taxate (bzw. Geo-Synonyme) ermittelt. Zahlenwerte zwischen eckigen Klammern: Anzahl der von dem betreffenden Taxat besetzten Messpunkte bzw. Polygone.

Es handelt sich um eine nach lexikalischen Kriterien erstellte ‘Typisierung’ – so hieß das in der klassischen romanischen Sprachgeographie – der Karte 173 des ALF, die 639 galloromanische Bezeichnungen für das Muttertier des Schafes (*la brebis*) enthält. Eine an der Schnittstelle zwischen romanistisch-etymologischen und statistisch-klassifikatorischen Reflexionen stehende Analyse dieser Originalkarte des ALF hat 15 Typen zu Tage gefördert, die in Salzburg *Taxate* heißen.

Bekanntlich lag und liegt der eigentliche Sucus solcher Typisierungen in der Entdeckung der erstaunlichen Variabilität der betreffenden Diffusions- oder Verteilungsareale, die im vorliegenden Fall zwischen 245 Messpunkten bzw. Polygonen (für das *Taxat 1 brebis*) und nur einem Messpunkt-Polygon (für die *Taxate 11–15*) liegt. Aufgabe nicht nur der romani(sti)schen Sprachgeographie bzw. Onomasiologie war es denn auch, möglichst viele ‘erklärende’ Informationen zur Größe und Verteilung dieser *Taxata*areale beizubringen.

Im vorliegenden Fall hat sich darum kein Geringerer als Walther von Wartburg (1888–1971) bemüht. Die seiner im Jahr 1918 von der Berliner Akademie publizierten Habilitations-Schrift beigegebene Karte ist allerdings nicht mit jener empirischen Akribie erstellt worden, die für die hier sichtbare Karte in Salzburg entfaltet wurde. Bei von Wartburg scheinen nur die *Taxate 1–3*<sup>4</sup> auf, während für den verbleibenden Rest auf der betreffenden Kartierung weiße Flächen vermerkt sind. Für einen auf formale Genauigkeit und Exhaustivität erpichten Sprach-Statistiker ist eine derartige ‘Großzügigkeit’ nicht nur an sich unzulässig, sondern – *pis encore* – im wahrsten Wortsinn sogar Erkenntnis-verschleiernd.

## 2. Zu Struktur und Inhalt der Datenmatrizen der Salzburger DM: das Fallbeispiel Frankreich (ALF)

Bekanntlich bestehen die ‘großen’ romanischen Sprachatlanten wie ALF<sup>5</sup> oder AIS<sup>6</sup> aus mehreren Hundertschaften an Messpunkten und einer deutlich noch viel größeren Anzahl an Atlaskarten. Für die Serie A des ALF sind das 1.421 und für den Gesamtumfang des AIS 1.705 großformatige Atlas-Karten<sup>7</sup>. Eine der Haupt-

<sup>4</sup> Hier die lateinischen Etyma dieser drei *Taxate*: *Taxat 1: brebis* (< VERVÍCE ‘Hammel’), *Taxat 2: onaille* (< OVÍCULA ‘[kleines] Schaf’), *Taxat 3: fedo* (< fêta ‘Tier, das geworfen hat’).

<sup>5</sup> Siehe dazu die groß angelegte und mit vielen Farbgraphiken ausgestattete ALF-Exegese von Brun-Trigaud/Le Berre/Le Dù (2005).

<sup>6</sup> Siehe dazu den nach wie vor als Meisterwerk einzustufenden Einführungsband zum AIS (Jaberg/Jud 1928).

<sup>7</sup> Ein für die romanische Sprachgeographie ganz entscheidender Umstand ist die außerhalb der Romanistik kaum bekannte (und schon gar nicht entsprechend gewürdigte) Tatsache, dass ‘unsere’ Atlaskarten ‘Volltext-Karten’ sind, die somit die im Feld gesammelten Daten in ihrer originalen und nicht in kodierter Form enthalten. Jede seri-

motivationen von Jules Gilliéron (1854–1926), dem genialen Autor des ALF, war ja die Beibringung einer möglichst großen Menge an empirisch optimal (und zudem in vergleichbarer Form) erhobenen ‘Bilanzen’ zur geographischen Verteilung möglichst vieler ‘Wörter, Laute und Formen’, um das einmal ganz unfachmännisch auszudrücken. Als dann diese Materialien ab 1902 in gedruckter Form auf dem Tisch lagen und ausgewertet werden konnten, war das Erstaunen der Fachwelt über die allen Vorerwartungen zuwiderlaufende räumliche Variabilität all dieser ‘Wörter, Laute und Formen’ grenzenlos. Das betraf natürlich die Sprachgeographien aller Philologien gleichermaßen, wiewohl sich angesichts der alt- und allbekannten interdisziplinären Blindheit auf Gegenseitigkeit die Kunde davon nur sehr langsam in allen Philologien bzw. in ‘ganz Europa’ verbreitet hat.

Die im Rahmen der Romanistik seit etwa 1905<sup>8</sup> anhand der Daten des ALF getätigten ‘Typisierungen’ (im Deutschen auch ‘Kartendiskussionen’ genannt) waren bzw. sind zum einen sehr zahlreich und zum anderen in vielerlei Hinsicht sehr wertvoll. Leider existiert dazu bis heute keine exhaustive Bibliographie<sup>9</sup>. Wiewohl ja seit dem ALF alle Sprachatlanten über die formale Struktur einer zweidimensionalen Matrix (N Messpunkte mal P [originale] Atlaskarten) verfügten, ist diese Tatsache den allein *qualitativ* denkenden und vorgehenden Linguisten eigentlich bis zum Auftreten der Dialektometrie (DM) verborgen geblieben. Das hat auch damit zu tun, dass vor der DM von den traditionellen Sprachgeographien (aller Philologien und Länder) eine an sich *quantitativ* ausgerichtete Tugend nur in minimalen Ansätzen geübt wurde, die in anderen, mit analogen Massendaten befassten empirischen Disziplinen schon viel früher praktiziert worden war: nämlich die systematische Synthese einer größeren Menge von Sprachatlas-Daten. Das erwähnte Minimum ereignete sich meistens bei der Erstellung von Isoglossen-Synthesen, wovon man schon in der meisterhaften Erst-Vorstellung des ALF durch Karl Jaberg (1908) einige sehr sprechende Kartenbeispiele (alle in Farbe) findet.<sup>10</sup>

Im Falle der schon erwähnten Gesamt-Dialektometrisierung des ALF ging es darum, nach Maßgabe der in Salzburg vorhandenen ‘Ressourcen’ (hinsichtlich Mitarbeitern und Forschungsmöglichkeiten) zum einen das Gesamtnetz des ALF einzubeziehen und zum anderen eine maximale Anzahl der an sich vorhandenen 1.421 Originalkarten (der Serie A) zu berücksichtigen bzw. zu taxieren. Tatsäch-

---

öse ‘Lektüre’ einer solchen Volltext-Karte gerät in der Romanistik zu einem genuin sprachklassifikatorischen Akt mit nachfolgender Visualisierung unter Einsatz ‘stummer Karten’: siehe dazu Jaberg (1906, 1908) sowie Goebel (2018).

<sup>8</sup> Das Datum bezieht sich auf die Publikation der bekannten Pionier-Arbeit von Gilliéron/Mongin (1905).

<sup>9</sup> Eine solche existiert jedoch für die in der Anglistik getätigten Kartendiskussionen: Fischer/Ammann (1991).

<sup>10</sup> Eine kartentechnisch sehr gut gelungene Isoglossen-Synthese liegt beispielsweise in Rosenqvist (1919) vor. Die Karte wurde in leicht verbesserter Form erneut publiziert in Berschin/Felixberger/Goebel (2008<sup>2</sup>:261).

lich konnten davon 626 Stück (~ 44%) berücksichtigt und zu 1.681 ‘Arbeitskarten’ (AK) verarbeitet werden. Die damit generierte Datenmatrix (N mal p) hatte also die Dimensionen: N = 641 Messpunkte mal p = 1 681 AK. Dazu zwei Kommentare:

a) Zur Anzahl der Messpunkte:

Die 638 originalen Messpunkt-Vektoren des ALF<sup>11</sup> wurden durch die Hinzufügung dreier künstlicher Vektoren auf 641 erweitert. Diese bezogen sich auf die Hochsprachen *Französisch*, *Italienisch* und *Katalanisch*, wobei unter *Französisch* die von Gilliéron gewählten Kartentitel und bei den beiden anderen Sprachen deren lexikographisch ermittelbare Äquivalente zu verstehen sind. Die Hinzufügung dieser Kunstpunkte ist für die Aufzeichnung des flächigen Impakts dieser Hochsprachen auf die Dialektlandschaften der Galloromania sehr nützlich.

b) Zur Anzahl der AK:

Der Fachausdruck *Arbeits-Karte* (AK)<sup>12</sup> weist darauf hin, dass hier etwas prinzipiell Anderes als eine *Original-Karte* des ALF vorliegt. Dieser Unterschied beruht vor allem auf der genau definierten kategoriellen Wertigkeit der aus den ALF-Originalkarten abgeleiteten AK. Die hier in Karte 1 abgebildete Taxierung der Karte 173 (*brebis*) des ALF ist *lexikalisch* ausgerichtet. In aller Regel kann einer *lexikalisch* relevanten Originalkarte des ALF nur eine *lexikalisch* relevante AK entnommen werden. Anders ist die Situation bei einer ALF-Karte, die nur aus *phonetisch* relevanten Varianten ein und desselben Basis-Etymons besteht, wie dies etwa bei der ALF-Karte 233 *chanter* (< lat. CANTÁRE) der Fall ist. In einem solchen Fall können die einzelnen etymologischen Komponenten des betreffenden Etymons (wie z. B. CA-, -À]-, -NT-, Á[, -R, -E) jeweils zum Gegenstand einer gesonderten Taxierung samt nachfolgender Kartierung werden, so dass aus einer einzigen *Original-Karte* mehrere *Arbeits-Karten* (hier: sechs Stück) entstehen können.

Die taxierten AK werden vor ihrer Einfügung in die dialektometrische Datenmatrix hinsichtlich ihrer kategoriellen Zugehörigkeit entsprechend indiziert.

---

<sup>11</sup> Das ALF-Netz enthält Informationen zu 638 *Ortschaften*, wobei vom ALF-Explorator Edmond Edmont (1849–1926) zum Messpunkt 284 (St-Pol-sur Ternoise, Pikardie), seinem Geburtsort, zwei *Enquêtes* gemacht worden sind, deren Resultate ebenso auf den ALF-Karten vermerkt sind.

<sup>12</sup> Dafür existieren seit geraumer Zeit die folgenden internationalen Entsprechungen: frz. *carte de travail*, engl. *working map*, ital. *carta di lavoro*, span. *mapa de trabajo*, katal. *mapa de treball*.

Schon der frühen Sprachgeographie ist – sicher auch als Folge der schweißtreibenden Einfärbung der für die Romanistik kennzeichnenden ‘stummen Karten’ – die große Variabilität der sich bei dieser Tätigkeit einstellenden Kartenbilder aufgefallen. Leider ist dieser Umstand vor der DM nie zum Anlass irgendwelcher Zählungen genommen worden; jedoch hat man, durchaus in Fortführung älterer Vorahnungen aus dem letzten Viertel des 19. Jahrhunderts, im Rahmen dieser Arbeiten immer wieder mit ebenso großem Bedauern wie Erstaunen festgestellt, dass sogar bei kategoriell eng verwandten Kartierungen (wie z. B. für lat. CA- in phonetisch analogen Karten wie ALF 229 *chandelle* [< lat. \*CANDÍLLA], ALF 225 *champ* [< lat. CAMPU] oder ALF 262 *chemin* [< kelt. CAMMÍNU]) die herauspräparierten Verteilungsareale (hier zu CA- > k-) ausnahmslos mehr oder weniger deutlich voneinander differierten und somit auch die entsprechenden Umgrenzungslinien (‘Isoglossen’) nicht – wie man naiverweise erwartete – ‘exakt’ zusammenfielen.

Immerhin lautete aber eine der daraus gezogenen Lehren: „Jedes Wort hat seine eigene Geschichte“<sup>13</sup>, deren wortwörtliche Befolgung letztendlich zum phänomenalen Aufschwung der romanistischen Lexikologie (à la FEW) beigetragen hat.

Eine andere – wissenschaftlich nur indirekt ertragreich gewordene – Lehre war das vor allem in Frankreich zirkulierende Verdikt von der ‘Nicht-Existenz von Dialekten’,<sup>14</sup> das aber glücklicherweise immer mit dem Nachsatz verbunden war, dass nur das Studium der räumlichen Verteilung vieler linguistischer Einzel-tatsachen im wahrsten Wortsinn ‘wissenschaftlich’ und daher mit aller Kraft anzustreben sei.<sup>15</sup> Und just darauf beruht ja letztendlich der *furor empiricus* von Jules Gilliéron, dem wir den wunderbaren ALF und die ganze Sprachgeographie verdanken.

### 3. Zur *Polynymie* der Arbeitskarten

Unter *Polynymie* versteht man die taxat-spezifische ‘Zerreißung’ bzw. Granulation einer AK. Die Gliederung der in Abschnitt 1 abgedruckten Karte 1 beruht auf der räumlichen Ko-Präsenz von 15 Taxaten: die ganze AK ist daher 15-nym. Theoretisch kann die Polynymie einer AK zwischen 2 und N variieren: da eine *mono-*

<sup>13</sup> Siehe dazu die Überblicke bei Christmann (1971) und Malkiel (1967). Erstaunlicherweise ist dieses auf das frühe 19. Jahrhundert zurückgehende Diktum im Licht der Sprachgeographie nicht in der folgenden Weise umgemünzt worden: ‘Jedes Wort hat seine eigene *Geographie*’.

<sup>14</sup> Die dafür hauptverantwortlichen Propugnatoren waren Paul Meyer (1840–1917) und Gaston Paris (1839–1903).

<sup>15</sup> Zum Postulat der Nicht-Existenz von Dialekten siehe vor allem Gaston Paris (1888), zu jenem der Nicht-Klassifizierbarkeit von Dialekten in erster Linie Hugo Schuchardt (1900 [1870]).

nyme AK keinerlei Variation zeigt, kommen mononyme (und daher variationslose) AK eo ipso für unseren variationsorientierten Forschungskontext nicht in Frage. Umgekehrt ist es theoretisch denkbar, dass jeder der N Messpunkte einer AK über ein eigenes Taxat verfügt.

Bei der erwähnten Dialektometrisierung des ALF oszillierte die Polynymie der 1.681 erstellten AK zwischen 2 und 90.<sup>16</sup> Wenn man nun die bei der Taxierung erstellten AK nach ihrer Polynymie sortiert und die sich dabei ergebende Ordnung in passender Weise visualisiert, entsteht eine überaus prägnant strukturierte Kurve: siehe dazu die Figur 1 (auf der nächsten Seite) und auch die auf den ALF bezogenen Informationen der Tabelle 1 in Abschnitt 6. Kurven dieser Art haben sich bei allen unseren dialektometrischen Analysen ergeben, so dass schon sehr früh die Vermutung aufgekommen ist, dass es sich hier um eine dem ‘System Sprachatlas’ prinzipiell immanente Gesetzmäßigkeit handeln müsse: siehe dazu den Beitrag von G. Altmann aus dem Jahr 1985, worin die Genese (= Geburt) und das Verschwinden (= Tod) der auf den Karten eines Sprachatlases beobachtbaren Taxatareale mit mathematisch modellierbaren Geburts- und Todes-Prozessen (allesamt als ‘Goebel-Gesetz’ bezeichnet) in Verbindung gebracht werden.<sup>17</sup>

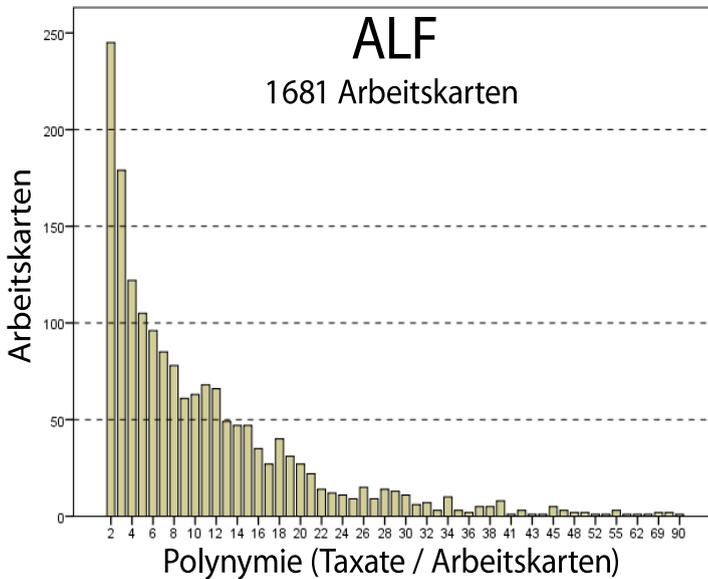
Allgemein gesprochen zeigt die Figur 1 eine Exponentialverteilung, wie sie auch in der allgemeinen Sprachstatistik bei den von George Kingsley Zipf (1902–1950) entdeckten Rang-Frequenz-Korrelationen auftritt. Aus graphischen bzw. arbeitspraktischen Gründen hat sich die traditionelle Sprachgeographie stets nur mit Arbeitskarten kleiner und mittlerer Polynymie befasst. Allzu reichhaltig gegliederte Atlaskarten wurden kaum je ‘diskutiert’ oder sonstwie hinsichtlich ihrer Struktur reflektiert. Damit blieb natürlich der Blick auf diese Gesetzmäßigkeit verstellt.

Mit der in Figur 1 sichtbaren gesetzesgesteuerten Verteilung ist noch ein Kuriosum verbunden, das aber erst in den 1990er-Jahren entdeckt worden ist: man kann sich ja auch fragen, was bei der ‘normalen’ dialektometrischen Synthese einer größeren Menge von AK vergleichbarer Polynymie ‘herauskommt’. Immerhin war ja schon seit den 1970er-Jahren bekannt, dass bei Vorliegen einer hinsichtlich des Kennwerts  $p$  (= Gesamtmenge der AK) hinreichend großen Datenmatrix im Falle

---

<sup>16</sup> Das Polynymie-Maximum von 90 (Taxaten / AK) ergab sich bei der Taxierung der *lexikalisch* relevanten Karte ALF 545 *fauvette* (‘Grasmücke’: ein kleiner Vogel). Bei der vorliegenden Netzgröße des ALF (= 641 Messpunkte) könnte theoretisch die Polynymie bis zu diesem Wert gehen. Dass beim ALF der mögliche Spielraum nur zu 14% (= 90 : 641) ausgenutzt wurde, hat mit der im Vergleich zu anderen romanischen Atlanten relativ geringen Binnenvariabilität der betreffenden Daten zu tun. Diese Tatsache wiederum geht weitgehend auf die in Frankreich seit vielen Jahrhunderten laufende ‘Zentralisierung’ (= Französisierung) fast aller Bereiche des bäuerlich-bürgerlichen Lebens zurück.

<sup>17</sup> Weitere Literatur zu diesem „Diversifikations-Problem“ findet man bei Best (2014).



Figur 1: Histogramm zur Veranschaulichung der Relation zwischen dem Grad der Polynymie ( $x$ -Achse) und der Anzahl der entsprechenden Arbeitskarten ( $y$ -Achse). Datenbasis: 1.681 aus 626 Originalkarten des ALF abgeleitete Arbeitskarten (AK); alle linguistischen Kategorien.<sup>18</sup>

einer aleatorisch vorgenommenen *Halbierung*, *Drittellung*, *Viertelung* etc. des  $p$ -Potentials dieser Datenmatrix praktisch dieselben dialektometrischen Analyse-Resultate erzielt werden.<sup>19</sup>

Der erste der sich daraus ergebenden Schlüsse war, dass geolinguistische Netze bzw. die in ihnen angelegten Ordnungsstrukturen über eine sehr große innere Redundanz verfügen, wie man das seit den 1940er-Jahren aus der Kommunikationstheorie (‘Sender-Empfänger-Modell’ von Claude E. Shannon [1916–2001] und Warren Weaver [1894–1978]) und den dazu passenden physikalischen Verhältnissen kennt.

Eine weitere, direkt die Polynymie einbeziehende Versuchsanordnung bestand darin, aus einer vorliegenden umfangreicheren Datenmatrix drei in polynymischer Hinsicht verschieden strukturierte Unter-Korpora zu bilden, die aus *größer* (= ‘*oligo-nym*’), *mittel* (= ‘*meso-nym*’) und sehr *fein* bzw. *bunt* (= ‘*poikilo-nym*’) geglie-

<sup>18</sup> Verständnishilfe:  $x$ -Achse: die Polynymie schwankt zwischen 2 und 90 Taxaten pro AK.  $y$ -Achse: die Zahl der AK variiert zwischen 245 (gilt für 2-nyme AK) und 1 (gilt für 90-nyme AK).

<sup>19</sup> Siehe dazu Goebel (1984:Bd. I, 206–219).

derten AK bestehen. So kann man die 1.681 AK des vorliegenden ALF-Gesamtkorpus wie folgt in drei mengenmäßig vergleichbar große Untermengen zerlegen:

Polynymie 1–5: 536 AK  
 Polynymie 6–12: 557 AK  
 Polynymie 13–90: 588 AK

Zu unserem großen Erstaunen haben sich für alle drei Korpora praktisch dieselben dialektometrischen Analyseresultate ergeben<sup>20</sup>, woran auch anders vorgenommene polynymische Stückelungen nichts geändert haben. Exakt dieselben Erfahrungen konnten auch bei der analogen Verarbeitung der zu anderen Sprachatlanten erstellten Datenmatrizen gemacht werden.

Erneut verweist dieser Befund auf eine systemimmanente Redundanz innerhalb der globalen Tiefenstrukturen geolinguistischer Netze. Seit geraumer Zeit sehen wir darin das Resultat einer ‘raumbewirtschaftenden’ Tätigkeit, die von der Gesamtheit aller basilektalen Sprecher eines gegebenen Großraumes ausgeübt wird und als kommunikative Leistung natürlich gewissen Strukturbedingungen unterworfen ist.

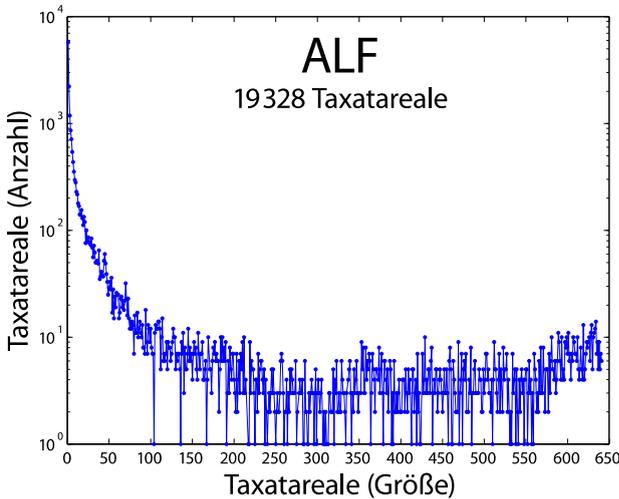
#### 4. Zur Größe der Taxatareale

Im Fall der hier besprochenen Arbeitskarte (Karte 1) entsprechen den 15 von unserem Taxierungsteam herauspräparierten Taxaten ebensoviele Taxat-*Areale*, deren Größe hinsichtlich der Zahl ihrer Messpunkte zwischen 245 und 1 schwankt. Klarerweise kann bzw. sollte man sich fragen, wieviele Taxate die zum ALF erstellte Datenmatrix in toto enthält und wie sich die Verhältnisse zwischen den *Größen* der dazugehörigen Areale und deren *Anzahlen* gestalten: siehe dazu die Figur 2 auf der folgenden Seite. Die Größen der Flächen der 19.328 vorhandenen Taxatareale<sup>21</sup> variieren zwischen 1 (= 10<sup>0</sup>) und 640 (Messpunkten bzw. Polygonen). Zu letzterer Dimension ergibt sich eine Variation zwischen 5.743 (= 10<sup>3,759</sup>) und 1. Es gibt also sehr viele ganz kleine und nur ganz wenige sehr große Areale, deren Flächen sich im Bereich von N-1 bewegen. Erneut hat sich diese Relation bzw. die Gestalt der in Figur 2 sichtbaren Kurve in allen unseren Datenmatrizen nachweisen lassen.

Noch eine sprachhistorische Reflexion zu diesem Sachverhalt: Einer der Haupterträge der klassischen Sprachgeographie bestand ja in der Aufzeichnung einer

<sup>20</sup> Siehe dazu unsere Berichte in Goebel (2014) (zum ALF) sowie Goebel/Smečka (2017) (zum AIS).

<sup>21</sup> Dieser Betrag umfasst nur *linguistisch* relevante Areale. Jene Areale, die nur aus *Nullstellen* (fehlenden Daten) bestehen, werden bei solchen Zählungen nicht erfasst.



Figur 2: Diagramm zur Veranschaulichung der Relation zwischen der Größe der Taxatareale (x-Achse) und deren jeweiliger Anzahl (y-Achse). Datenbasis: 19.328 Taxatareale, die auf 1.681 Arbeitskarten (AK) auftreten, die ihrerseits aus 626 Originalkarten des ALF abgeleitet worden waren; alle linguistischen Kategorien.<sup>22</sup>

omnipräsenten, oft als ‘Kampf’ qualifizierten Konkurrenz zwischen den Einzugsbereichen verschiedener sprachlicher Elemente, wobei alle diese Flächen als historisch gewachsene und somit diachron variable Entitäten angesehen wurden (und zu Recht immer noch werden). Der ‘Sprachwandel’ als solcher vollzieht sich also aus der Perspektive der Sprachgeographie – und damit auch der DM – als ein vielschichtiger und vielgestaltiger Wandel der räumlichen Implantation(en) einer sehr großen Menge sprachlicher Einheiten.

## 5. Zur ‘speziellen Verzahnung’ (SVZ) zwischen den Taxatarealen

Weiter oben (Abschnitt 2) wurde schon erwähnt, dass bereits vor der Publikation des ALF (1902–1910) bekannt war, dass sogar die Areale von linguistisch statusgleichen oder status-analogen Merkmalen nur wenig Neigung zur Deckungsgleichheit zeigten. Dieser Umstand ist als empirisches Faktum besonders dort offenkundig geworden, wo die Zahl der beobachteten Messpunkte besonders groß war.

<sup>22</sup> Verständnishilfe: x-Achse: die Größen der 19.328 Taxatareale schwanken zwischen 1 und 640 Messpunkten bzw. Polygonen. y-Achse (logarithmisch): die Anzahlen jeweils gleich großer Taxatareale variieren zwischen 5.743 [=  $10^{3,759}$ ] (für Taxatareale der Größe 1) und der Menge 1 [=  $10^0$ ] (für Taxatareale der Größe 640).

Nun ragte und ragt in Europa hinsichtlich der Anzahl der berücksichtigten Messpunkte ein Atlas mit sehr großer Deutlichkeit aus der Menge seiner ‘Konkurrenten’ hervor: der „Deutsche Sprachatlas“ (DSA) von Georg Wenker (1852–1911). Ebenso war bzw. ist es ohne jeden Zweifel Wenker, der in einem auch das 19. Jahrhundert umfassenden (imaginären) „Guinness-Buch der Rekorde“ den Spitzenplatz beim *Kartenzeichnen* beanspruchen könnte. Er hat in der Tat ab den frühen 80er-Jahren des 19. Jahrhunderts auf der Grundlage der an ihn zurückgeschickten Fragebögen eine große Menge feinstgeklärter Sprachkarten gezeichnet, wobei ihm der folgende Sachverhalt aufgefallen ist:

Sind so sämtliche Formen, in denen das Wort erscheint, kartographisch verzeichnet, so werden die einzelnen zu Gruppen sich zusammenschließenden Abweichungen<sup>23</sup> durch Linien abgegrenzt, mit verschiedenen Farben kenntlich gemacht und so das Ganze zu einem übersichtlichen Bilde gestaltet. [...] Dann geschieht die Uebertragung in die Grundkarten<sup>24</sup> des Sprachatlas, zu denen ein erläuternder Text hinzutritt. Jedes einzelne Wort wird also ganz unabhängig von allen anderen, selbst von verwandten, zu Ende verarbeitet, dann erst werden seine Grenzlinien<sup>25</sup> und seine verschiedenen Formen verglichen mit verwandten Erscheinungen ähnlicher Wörter. Es ist dies eine Vorsicht, welche erst im Verlauf der Arbeit zum Grundsatz erhoben worden ist. *Anfänglich war ich wie wohl jeder allzusehr geneigt, von der bequemen und nabeliegenden Vorstellung auszugehen, daß verwandte Wörter, etwa Hund und Pfund, Wurst und Durst auch in ihren mundartlichen Eigenheiten zusammenstimmen müßten. Indessen stellte sich heraus, daß dies nicht immer der Fall ist, daß zwar jedes einzelne Wort seine meist ganz festen Grenzlinien besitzt, daß die Grenzlinien verschiedener Wörter dagegen selbst da, wo man es ganz bestimmt erwartet, nicht immer zusammen fallen, sondern bald mehr bald weniger abweichen. Dies allgemeine Ergebnis muß zunächst, gerade wegen seines Gegensatzes zu den bisherigen Anschauungen, nachdrücklich betont werden, bis man sich an diese etwas unbequeme Thatsache gewöhnt hat.* [Kursive: HG]

(Wenker 1889–1897 [2013]:10)

Beim Blick auf das zweidimensionale Schema einer Datenmatrix oder auf die auf einer Isolinien-Synthese kleineren Umfangs sichtbar werdenden Isoglossen-Wülste ergibt sich hinsichtlich der systeminternen Koexistenz verschiedener Taxatarea-

<sup>23</sup> In Salzburger Terminologie: *Taxate*.

<sup>24</sup> Diese ‘Grundkarte’ entspricht den in der Romanistik zur Auswertung der diversen Sprachatlanten benützten ‘stummen Karten’. Schlussendlich umfasste die Grundkarte des DSA mehr als 50 000 Messpunkte.

<sup>25</sup> Wenker bezieht sich auf Umgrenzungslinien, die erst ab dem Jahr 1892 (durch den deutsch-baltischen Pfarrer August Bielenstein, 1826–1907) mit dem Namen *Isoglossen* belegt worden sind.

le zwanglos das Bild einer *dachziegelartigen Überlappung*, einer *Verschachtelung* oder auch einer *Verzahnung*.

Ich habe mich zur terminologisch standardisierten Erfassung dieses Phänomens für die Metapher der ‘Verzahnung’ entschieden. Hier die Äquivalente zur *speziellen Verzahnung* (SVZ) in verschiedenen Fremdsprachen: eng. *special entanglement*, frz. *enchevêtrement particulier*, ital. *intreccio particolare*, span. *entramado especial*, katal. *encreuament específic*. Dazu zwei Bemerkungen:

- a) Zum einen ist der räumlich zu verstehende Umstand der SVZ der Verbreitungsflächen bestimmter sprachlicher Merkmale schon gegen 1875 dem austro-italienischen Linguisten Graziadio Isaia Ascoli (1829–1907) aufgefallen, der dafür damals das Binom *particular combinazione* [*di quei caratteri*] verwendete bzw. kreierte (Ascoli 1876:387).
- b) Zum anderen haben sich viele Linguisten – ganz im Sinn dessen, was Wenker oben anspricht – bis weit in das 20. Jahrhundert hinein geweigert, diese Nicht-Koinzidenz (von Flächen und Isoglossen) als ‘natürliches’ Faktum *tel quel* zu akzeptieren, und vielmehr immer wieder versucht, spezifische Gründe für die als erwiesen angenommene ‘Verletzung’ einer grundlegenden ‘Regel’ (= punktgenaue Koinzidenz vieler Flächen und den dazugehörigen Isoglossen) zu finden oder anzuführen.<sup>26</sup>

Im Gegensatz dazu nehmen wir an, dass neben den vorhin erwähnten Häufigkeiten auch die SVZ zu den systemimmanenten Besonderheiten geolinguistischer Netze zählt und dazu beiträgt, die kommunikativen Funktionen solcher Netze wenn schon nicht grundlegend zu ermöglichen, so doch zu optimieren.

In wissenschaftshistorischer bzw. -systematischer Hinsicht liegt hier ein perfektes Analogon zu den bei der Sprachgesetz-Frage der Indogermanistik manifest gewordenen Dilemmata vor.<sup>27</sup> Mit Blick auf die nur *eine* Arbeitsdimension umfassende Achse der *Zeit* wurde dort unter Rückgriff auf die als Axiom behandelte ‘Ausnahmslosigkeit der Lautgesetze’ angenommen, dass die *Gesamtbeit* aller einem bestimmten Lautkomplex der Ursprache zuzurechnenden Nachfolgeformen nach

<sup>26</sup> Dafür symptomatisch sind die durch W. von Wartburg (1963:22–24; Karte auf Seite 24) angeführten Explikationen zu den auf den ALF-Karten zu *champ*, *chandelier*, *chandelle*, *chanson* und *chaîne* verschieden ausgeprägten Arealen der Erhaltung des Velars *k* (< lat. CA-). Argumentation und Karte scheinen auch in Berschin/Felixberger/Goebel (2008<sup>2</sup>:255–256) auf.

<sup>27</sup> Siehe dazu die Reflexionen von W. Putschke (2001), wo es allerdings darum geht, die prinzipielle Unabhängigkeit der Wenker’schen Unternehmung (später ausgewachsen zum DSA) von der um 1876 angelaufenen Lautgesetz-Debatte zu erweisen. In unserem Kontext ist dagegen unumgänglich, die enge phänomenologische Verschränkung der grundlegenden Dimensionen der *Zeit* und des *Raumes* als ein gegebenes Faktum hin- und anzunehmen.

Durchlauf eines bestimmten Abschnitts dieser Zeitachse *ein und dieselbe Gestalt* haben müssten. Jedoch hat sich immer wieder gezeigt, dass das nicht der Fall war. Einer der Faktoren, die zur Erklärung dieser 'Deviationen' herangezogen wurden, war bekanntlich die Analogie.

Auch im Rahmen der Geolinguistik herrschte zu Beginn der diesbezüglichen Forschungen eine gesetzesaffine Vorerwartung: nämlich, dass die Verbreitungsgebiete *aller* einem bestimmten Lautkomplex der Ursprache zuzurechnenden Nachfolgeformen innerhalb der *zwei* Arbeits-Dimensionen der Sprachgeographie (= des geographischen Naturraumes) *ein und dieselbe Gestalt* haben müssten. Die in der DM heute vorherrschende Erkenntnis, dass dieser Umstand 'naturgegeben' und nicht 'katastrophenbedingt' ist, hat bis zu ihrer festen Etablierung ein rundes Jahrhundert benötigt.<sup>28</sup>

## 6. Noch eine empirische Bestätigung der zitierten Regularitäten: das Fallbeispiel England (SED)

Die eben für Frankreich bzw. die Galloromania anhand des ALF aufgezeigten Regularitäten konnten nicht nur in zahlreichen anderen romanischen Atlanten, sondern auch in all jenen 'Atlanten' nachgewiesen werden, die im Rahmen der zu den Dialekten Englands entfalteten Geolinguistik entstanden sind. Diese ist in historischer und vor allem in methodisch-prozeduraler Hinsicht davon, was man aus der Romanistik kennt, sehr verschieden. In England hat man sich erst ab 1950 zu einer flächendeckenden Erhebung der eigenen Dialekte entschieden. Hauptinitiatoren dieses „Survey of English Dialects“ (SED) waren der in Leeds tätige Dialektologe Harold Orton (1898-1975) und der in Zürich lehrende Anglist Eugen Dieth (1893-1956).

Die durch ambulierende Linguisten an 313 Messpunkten und anhand eines rund 1.300 Items umfassenden Fragebuchs im Wege direkter Befragung realisierten Enquêtes dauerten bis 1961. Erst bei der Veröffentlichung der gesammelten Daten ergab sich ein (riesen)großer (und folgenschwerer) Unterschied zum ALF: Aus Kostengründen wurden die Daten exklusiv in Tabellenform veröffentlicht, wobei die dazu bis zum Jahr 1971 veröffentlichten 12 Bände („Basic Material“) die gesammelten Daten zwar in der originalen Lautschrift, jedoch ausschließlich in

---

<sup>28</sup> Wir lassen dabei den aufgespannten Zeitrahmen mit dem Jahr 1876, dem *annus mirabilis* der Indogermanistik, beginnen und vorläufig bei der ersten DM-affinen Publikation von Jean Séguy (1971) enden: cf. dazu Hoenigswald (1978). Im Jahr 1876 erschienen bahnbrechende Arbeiten der Indogermanisten August Leskien (*Die Declination im Slavisch-Litauischen und Germanischen*), Hermann Osthoff (*Die Frage des Ursprungs der germanischen n-Deklination*) und Karl Brugmann (*Nasalis sonans in der indogermanischen Grundsprache*) sowie des Germanisten Jost Winteler (*Die Kerenzer Mundart des Kantons Glarus in ihren Grundzügen dargestellt*).

A	B	C	D	E	F	G	H	I
Sprachatlas	Anzahl der jeweils vorhandenen Atlaskarten	Anzahl der taxierten Atlaskarten	Anzahl der beim Taxieren tätig gewordenen Linguisten	Anzahl der berücksichtigten Arbeitskarten	Anzahl der Messpunkte	Anzahl der ermittelten Taxate bzw. der Taxat-areale	Mittlere Größe der Taxatareale (= E×F/G)	Spannweite der Polynymie: zwischen 2 und x
<b>ALF</b>	1 421	626	1 (HG)	1 681	641	19 328	55,74	90
<b>England</b>	1 711	1 516	11	1 524	313	16 810	28,37	108
<b>TOT</b>								
AES	424	424	4	424	313	5 838	22,73	39
CLAE (I+II)	315×2 (zwei Taxierungsebenen)	591	2	597	313+1	7 698	24,35	108
LAE	406	388	3	389	313+1	2 839	43,02	21
WGE	251	114	2	114	313+1	435	82,28	7

*Tabelle 1: Kennwerte der dialektometrischen Vermessung des französischen Sprachatlasses ALF und von vier auf der Grundlage der Daten des SED erstellten englischen Sprachatlanten (AES, CLAE I u. II, LAE und WGE)*

der Form von Tabellen enthalten. Diese zeigen für jeden der besuchten 313 Messpunkte alle dazu erhobenen Antworten in der Abfolge der Fragenstellung, wobei die Anordnung der Messpunkte der in England seit dem 11. Jahrhundert kanonisch etablierten Sortierung der *Counties* gehorcht.

Faktisch mussten also – 50 Jahre nach dem ALF und unter Staunen erregender Ausblendung der schlagenden Erfolge der kartenbasierten Gilliéron-Geolinguistik – die englischen Linguisten ab ovo auf das *grundlegende heuristische Aha-Erlebnis* des so suggestiven Sofort-Überblicks über die räumliche Verbreitung vieler geolinguistischer Merkmale verzichten<sup>29</sup>.

Diesem Defizit wurde in der anglistischen Fachwelt dadurch begegnet, dass sich sehr rasch an verschiedenen Stellen Arbeitsgruppen gebildet haben, die die Tabellen des „Basic Materials“ in mühseliger Kleinarbeit nach bestimmten Kriterien durchsucht und typisiert sowie abschließend den Ertrag dieser ‘taxierenden’ Arbeit in der Form großformatiger Bücher<sup>30</sup> veröffentlicht haben. Erst bei der Durchsicht dieser, exklusiv typisiertes Material enthaltenden Kartenwerke konnte sich die ‘geolinguistische Dynamik’ der SED-Enquêtes entfalten.

Nachdem ich schon im Jahr 1996 durch die freundliche Vermittlung von Wolfgang Viereck (damals Bamberg) die EDV-Daten seines zweibändigen CLAE erhalten und unter Mitarbeit meines damaligen Projektmitarbeiters Guillaume Schiltz dialektometrisch auswerten konnte, hat sich angesichts der dabei erzielten vorzüglichen Resultate<sup>31</sup> meine Neugier auch den anderen, nur auf Papier vorhandenen SED-Typisierungen zugewandt. So wurden zwischen 2000 und 2005 unter Mitarbeit von Emese Lörincz, einer überaus engagierten Salzburger ‘Anglo-Romanistin’, die Daten dreier weiterer SED-Typisierungen (WGE 1974, LAE 1978 und AES 1979)<sup>32</sup> in recht aufwändiger Weise<sup>33</sup> maschinenlesbar gemacht und dann di-

---

<sup>29</sup> Innerhalb der Romanistik haben viele Forscher lange vor dem ALF die große Suggestivwirkung geolinguistischer Kartierungen vorausgeahnt und die Erstellung entsprechender Kartierung eingemahnt. Ich frage mich, was eine bei der Anglistik angestellte wissenschaftshistorische Nachschau diesbezüglich ergeben würde.

<sup>30</sup> Dabei kam mehrfach der bei uns Romanisten falsche Assoziationen weckende Titel „(Linguistic) Atlas of...“ zur Anwendung.

<sup>31</sup> Siehe dazu Goebel (1997, 2007) sowie Goebel/Schiltz (1997).

<sup>32</sup> Diese drei Opera erschöpfen nicht die aus dem SED hervorgegangenen Typisierungen. Ihre Auswahl beruhte vor allem auf deren direkter Verfügbarkeit in Salzburg.

<sup>33</sup> Zuerst wurden alle Karten der drei Opera mit großer Auflösung gescannt, dann wurde zu jedem der betreffenden Kartengründe händisch (bzw. über einem Leuchtpult) ein Netz von Prüfpfaden erstellt und abschließend wurden diese Prüfpfade auf dem Bildschirm deckungsgenau auf die Karten-Scans aufgebracht. Darnach wurden von all diesen mit Prüfpfaden versehenen Karten Farbausdrucke in Originalgröße produziert, die ihrerseits die Grundlage der nachfolgenden händischen Extraktion aller Karteninhalte auf Listen darstellten. Zum einen entsprach die Struktur dieser Listen genau den vorhin erwähnten Prüfpfaden und zum anderen wurde im Salzburger DM-Programm

alektometrisiert. Darnach konnten natürlich all diese Daten in variablem Umfang kombiniert bzw. ‘zusammengeführt’ werden.

Die oben stehende Tabelle 1 zeigt die dabei zustande gekommenen Datenstrukturen. Auffällig sind dabei nicht nur die im Schnitt meist feinere Kammerung der englischen Daten gegenüber jenen des ALF, sondern auch die Variation dieser Kammerung zwischen den einzelnen SED-Typisierungen. Das liegt zum einen an den verschiedenen linguistischen Kategorien und zum anderen sicherlich auch an unterschiedlichen ‘Philosophien’ der Atlas-Autoren hinsichtlich der Typisierung der von ihnen bearbeiteten SED-Daten: Hier – wie auch in zahlreichen Natur- und Sozialwissenschaften – pendeln die typisierenden (und damit Daten-vereinfachenden) Zugriffe zwischen den Polen des *lumping* und des *splitting*.<sup>34</sup>

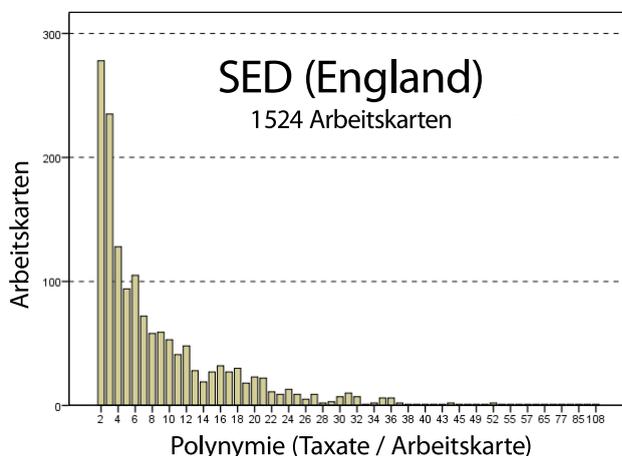
Die mit dem Salzburger Dialektometrie-Programm VDM („Visual Dialectometry“) leicht umsetzbare Zusammenführung der bei der Dialektometrisierung von AES, CLAE, LAE und WGE erarbeiteten Datenmatrizen hat ein respektables Corpus von immerhin 1.525 AK ergeben, das zudem die Kategorien Lexikon und Phonetik in vergleichbarer Dichte abdeckt. Hervorzuheben ist aber die Tatsache, dass die diesen Daten zu Grunde liegende Taxierungsarbeit nicht wie beim ALF nach den Vorstellungen eines einzigen Linguisten (hier: des Verfassers des vorliegenden Beitrags), sondern nach jenen von in toto elf anglistischen Fachkollegen vorgenommen worden ist: Siehe dazu die Kolonne D in Tabelle 1. Dass hier dennoch ‘viele Köche *nicht* den Brei verdorben haben’, soll hier entsprechend gewürdigt werden.

Die auf Figur 3 aufscheinende Kurve auf der nächsten Seite ähnelt weitgehend jener, die auf Figur 1 (zum ALF) sichtbar ist. Die Gesamt-Polynymie der englischen Taxierungen (108) übersteigt deutlich jene des ALF (90). Sie wurde bei den dem Viereck-Atlas CLAE entnommenen (lexikalischen) Daten erreicht. Auch die Kurven der Figuren 2 (zum ALF, siehe Abschnitt 4) und 4 (zum SED, ebenfalls auf der nächsten Seite abgedruckt) differieren voneinander nur wenig. Auffällig ist jedoch die beim ALF (Figur 2) deutlicher als beim SED (Figur 4) ausgeprägte Abflachung des Kurvenverlaufs nach Erreichung der ungefähren Areal-Größe 100. Dies bedeutet zweierlei: a) dass es im Durchschnitt beim ALF mehr großflächige Taxatareale als beim SED gibt; b) dass es im Durchschnitt beim SED mehr kleinflächige Taxatareale als beim ALF gibt.

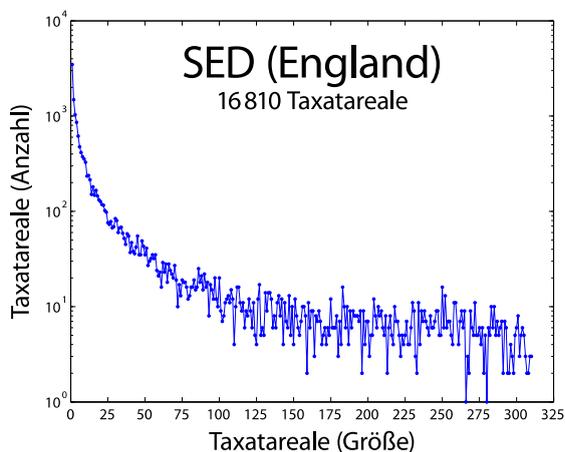
---

VDM eine dazu passende Eingabemaske erstellt. Die eigentliche EDV-Eingabe erfolgte schließlich durch händischen Übertrag des Inhalts der erwähnten Listen in diese Eingabemaske. Den Abschluss der ganzen Eingabeprozedur bildete eine doppelte Korrektur der händisch ‘eingeklopften’ Daten.

<sup>34</sup> Klarerweise erhöht sich im Zeichen des *splitting* die Polynymie der typisierten Daten: Siehe dazu den diesbezüglichen Unterschied (→ Kolonne I in Tabelle 1) zwischen den Polynymie-Spannweiten bei AES und LAE einerseits und dem vor allem lexikalisch orientierten CLAE andererseits.



Figur 3: Histogramm zur Veranschaulichung der Relation zwischen dem Grad der Polynymie (x-Achse) und der Anzahl der entsprechenden Arbeitskarten (y-Achse). Datenbasis: 1.524 im Rahmen der Atlanten AES, CLAE I+II, LAE und WGE aus 1.516 Originalkarten des SED abgeleitete Arbeitskarten (AK); alle linguistischen Kategorien.<sup>35</sup>



Figur 4: Diagramm zur Veranschaulichung der Relation zwischen der Größe der Taxatareale (x-Achse) und deren jeweiliger Anzahl (y-Achse). Datenbasis: 16.810 Taxatareale, die auf 1.524 Arbeitskarten (AK) auftreten, die ihrerseits im Rahmen der Atlanten AES, CLAE I+II, LAE und WGE aus 1.516 Originalkarten des SED abgeleitet worden waren; alle linguistischen Kategorien.<sup>36</sup>

<sup>35</sup> Verständnishilfe: x-Achse: die Polynymie schwankt zw. 2 u. 108 Taxaten pro AK. y-Achse: die Zahl der AK variiert zw. 278 (für 2-nyme AK) und 1 (für 108-nyme AK).

<sup>36</sup> Verständnishilfe: x-Achse: die Größen der 16.810 Taxatareale schwanken zwischen 1 (=10<sup>0</sup>) und 310 Messpunkten bzw. Polygonen. y-Achse (logarithmisch): die Anzahlen je-

Die tieferen Gründe hierfür scheinen mir weniger in den unterschiedlichen Atlas-Konzeptionen (ALF *versus* SED) als vor allem in den unterschiedlichen geolinguistischen Realitäten der Gallo-Romania und der Anglo-Britannia zu suchen zu sein.

## 7. Statt eines Nachworts

Im Sinne des General-Mottos dieser Festschrift („Sprache-Mensch-Maschine“) sei mir eine *maschinen*-bezogene Reflexion gestattet, die letztendlich in eine unbeantwortbare Frage einmündet. Die vorstehenden Darlegungen beziehen sich einzig und allein auf an der empirischen Oberfläche liegende Struktureigenschaften dialektometrischer Datenmatrizen. Zugleich dürfte aber allgemein bekannt sein, dass man in genau diesen Datenmatrizen unter Benützung des ‘normalen’ Methodenkansons der Salzburger Dialektometrie (S-DM) die Existenz optimal geordneter Tiefenstrukturen nachweisen und visualisieren kann. Es besteht natürlich kein Zweifel daran, dass diese ‘wohlgeordneten Strukturen’ von der S-DM nur *gefunden* und keineswegs *erfunden* worden ist. Was da *gefunden* worden ist, ist freilich nichts anderes als das Resultat menschlichen Tun und Lassens im Raum, wobei aber die inneren Mechanismen dieser Tätigkeit, deren Begründung oder gar Zielsetzung nach wie vor im Dunkeln verbleiben.

Nun hat es ja die Menschheit seit alters her immer wieder gejuckt, im Sinne der ewigen *Homunculus-Versuchung* für bestimmte Bereiche kleine Maschinen zu bauen oder wenigstens zu erdenken, die in variabler Autonomie und Perfektion menschlich anmutende Handlungen nachvollziehen bzw. setzen.

Im Lichte dieser *anthropoiden Temptatio* frage ich mich nun, ob und wie es möglich wäre, in Kenntnis der die ‘echte’ basilektale Kommunikation steuernden Gesetzmäßigkeiten von einem ‘Automaten’ eine zunächst inhaltsleer dastehende Datenmatrix dergestalt mit imaginären, aber ‘speziell verzahnten’ Taxaten zu befüllen, dass die nachfolgende dialektometrische Auswertung dieses maschinell-künstlich (und eben nicht empirisch-real) generierten Artefakts ähnlich ‘wohlgeordnete Strukturen’ ergibt, wie diese sich – beispielsweise – bei der Dialektometrisierung der Daten des ALF oder des SED ergeben.

Beim ALF beruhen die dialektometrisch nachgewiesenen ‘wohlgeordneten Strukturen’ letztendlich auf einer ganz speziellen Verzahnung und Verschränkung von mehr als 19.000 räumlich verteilten Merkmalen, beim SED sind das fast 17.000 Merkmale. Nach allem, was wir wissen, hat sich jede dieser Flächen quer

---

weils gleich großer Taxatareale variieren zwischen 3.477 [=  $10^{3,541}$ ] (für Taxatareale der Größe 1) und der Menge 1 [=  $10^0$ ] (für Taxatareale der Größe 310).

durch Zeit und Raum zwar individuell, aber in stetem Konflikt mit anderen Flächen dazu entwickelt, was man heute auf den Karten sehen kann.

Ist es denk- und dereinst vielleicht auch machbar, unter bestimmten sprachtheoretischen Vorab-Annahmen *maschinell* bzw. *algorithmisch* eine Datenmatrix zu generieren, die ‘wohlgeordnete Tiefenstrukturen’ enthält? Wie man weiß, sind derartige ‘Synthesen’ in vielen anderen linguistischen Bereichen bereits sehr erfolgreich, ja sogar mit industriell verwertbaren Erfolgen durchgeführt worden.

## Danksagungen

Ich habe mehrfachen, herzlichen Dank abzustatten:

- für die Generierung der Karte 1: Yves Scherrer, Genf
- für die Berechnung der quantitativen Grundlagen der Figuren 1–4: Pavel Smečka, Salzburg
- für die Erstellung der Figuren 1–4: Werner Goebel, Wien.

## Abkürzungen

AK: Arbeitskarte

DM: Dialektometrie

S-DM: Salzburger (Schule der) Dialektometrie

SVZ: spezielle Verzahnung (cf. dazu Kapitel 5)

VDM: Programm „Visual DialectoMetry“

## Bibliographie

- AES = Kolb, Eduard/Glauser, Beat/Elmer, Willy/Stamm, Renate (Hgg.) (1979). *Atlas of English Sounds*, Bern: Francke.
- AIS = Jaberg, Karl/Jud, Jakob (Hgg.) (1928–1940). *Sprach- und Sachatlas Italiens und der Südschweiz*, 8 Bde., Zofingen: Ringier (Neudruck: Nendeln: Kraus, 1971; Verfügbarkeit im Netz: <http://www3.pd.istc.cnr.it/navigais/> <14.07.2018>).
- ALF = Gilliéron, Jules/Edmont, Edmond (Hgg.) (1902–1910). *Atlas linguistique de la France*, 10 Bde., Paris: Champion (Neudruck: Bologna: Forni, 1968; Verfügbarkeit im Netz: <http://diglib.uibk.ac.at/urn:nbn:at:at-ubi:2-4568> <14.07.2018>).
- Altmann, Gabriel (1985). “Die Entstehung diatopischer Varianten”, *Zeitschrift für Sprachwissenschaft* 4, 139–155.
- Ascoli, Graziadio Isaia (1874). “Il franco-provenzale”, *Archivio Glottologico Italiano* 3.1, 61–120.
- Ascoli, Graziadio Isaia (1876). “Paul Meyer e il franco-provenzale”, *Archivio glottologico italiano* 2, 385–395.

- Berschin, Helmut/Felixberger, Josef/Goebel, Hans (2008<sup>2</sup>). *Französische Sprachgeschichte. Lateinische Basis, interne und externe Geschichte, sprachliche Gliederung Frankreichs. Mit einer Einführung in die historische Sprachwissenschaft*, Hildesheim: Olms.
- Best, Karl-Heinz (2014). "Diversification: Bibliography", *Glottometrics* 28, 87–91.
- Brun-Trigaud, Guylaine/Le Berre, Yves/Le Dù, Jean (2005). *Lectures de l'Atlas linguistique de la France de Gilliéron et Edmont. Du temps dans l'espace. Essai d'interprétation des cartes de l'Atlas linguistique de la France de Jules Gilliéron et Edmond Edmont augmenté de quelques cartes de l'Atlas linguistique de la Basse-Bretagne de Pierre Le Roux*, Paris: CTHS.
- CLAE = Viereck, Wolfgang/Ramisch, Heinrich (Hgg.) (1991–1997). *The Computer Developed Linguistic Atlas of England*, 2 Bde., Tübingen: Niemeyer, 2 vols.
- Christmann, Hans Helmut (1971). "Lautgesetze und Wortgeschichte. Zu dem Satz *Jedes Wort hat seine eigene Geschichte*", in: Coseriu, Eugenio/Stempel, Wolf-Dieter (Hgg.), *Sprache und Geschichte. Festschrift für Harri Meier zum 65. Geburtstag*, München: Fink, 111–124.
- DSA = Wrede, Ferdinand/ Mitzka, Walther/Martin, Bernhard (Hgg.) (1927–1956). *Deutscher Sprachatlas, auf Grund des von Georg Wenker begründeten Sprachatlas des Deutschen Reiches*. 4 Bde. [23 Faszikel mit 128 Karten], Marburg/Lahn: Elwert.
- FEW = von Wartburg, Walther (1922–2002). *Französisches etymologisches Wörterbuch*, 25 Bde., Basel: Zbinden u.a.
- Fischer, Andreas/Ammann, Daniel (1991). *An Index to Dialect Maps of Great Britain*, Amsterdam/Philadelphia: John Benjamin.
- Gilliéron, Jules/Mongin, Jean (1905). *Scier dans la Gaule romane du sud et de l'est. Étude de géographie linguistique*, Paris: Champion (italienische Übersetzung: Lorenzo Massobrio (Hg.) (1990). « *Segare* » nella Gallia romanza meridionale e orientale, Novi Ligure: Grafica editoriale universitararia).
- Goebel, Hans (1984). *Dialektometrische Studien. Anhand italoromanischer, rätoromanischer und galloromanischer Sprachmaterialien aus AIS und ALF*, 3 Bde., Tübingen: Niemeyer.
- Goebel, Hans (1993). „Dialectometry. A Short Overview of the Principles and Practice of Quantitative Classification of Linguistic Atlas Data“, in: Köhler, Reinhard/Rieger, Burghard B. (Hgg.), *Contributions to Quantitative Linguistics*, Dordrecht: Kluwer, 277–315.
- Goebel, Hans (1997). „Some Dendrographic Classifications of the Data of CLAE 1 and CLAE 2“, in: CLAE, Bd. II, 23–32 (siehe oben).
- Goebel, Hans (2000). „La dialectométrie de l'ALF: présentation des premiers résultats“, *Linguistica* 40, 209–236.
- Goebel, Hans (2002a). „Analyse dialectométrique des structures de profondeur de l'ALF“, *Revue de linguistique romane* 66, 5–63.
- Goebel, Hans (2002b). „Sprachatlant: woher? womit? wozu? Einige buntgemischte Reflexionen am Gartenzaun zwischen Romanistik und Germanistik“, in: Anreiter, Peter/Ernst, Peter/Hausner, Isolde/Kalb, Helmut (Hgg.), *Namen, Sprachen und Kulturen. Imena, Jeziki in Kulture. Festschrift für Heinz Dieter Pobl zum 60. Geburtstag*, Wien: Edition Präsens, 257–274.

- Goebl, Hans (2003). „Regards dialectométriques sur les données de l'Atlas linguistique de la France (ALF): relations quantitatives et structures de profondeur“, *Estudis Romànics* 25, 59–121.
- Goebl, Hans (2005). „La dialectométrie corrélative. Un nouvel outil pour l'étude de l'aménagement dialectal de l'espace par l'homme“, *Revue de Linguistique Romane* 69, 321–367.
- Goebl, Hans (2006). „Recent Advances in Salzburg Dialectometry“, *Literary and Linguistic Computing* 21.4, 411–435.
- Goebl, Hans (2007a). „A Bunch of Dialectometric Flowers: a brief Introduction to Dialectometry“, in: Smit, Ute/Dollinger, Stefan/Hüttner, Julia/Kaltenböck, Gunther/Lutzky, Ursula (Hgg.), *Tracing English through Time. Explorations in Language Variation. In Honour of Herbert Schendl on the Occasion of his 65<sup>th</sup> Birthday*, Wien: Braumüller, 133–171.
- Goebl, Hans (2007b). „Dialectometry: Theoretical Prerequisites, Practical Problems, and Concrete Applications (mainly with Examples drawn from the 'Atlas linguistique de la France', 1902–1910)“, in: *Geolinguistics around the World. Proceedings of the 14th NijlA [National Institute of Japanese Language] International Symposium (Tokyo, August 22–23, 2007)*, Tokyo: NijlA, 65–74.
- Goebl, Hans (2009). „Quelques coups d'oeil dialectométriques sur l'Atlas linguistique de la France: structures de surface et structures de profondeur“, in: Dalbera-Stefanaggi, Marie-José/Simoni-Aurembou, Marie-Rose (Hgg.), *Images de la langue: représentations spatiales, sémantiques et graphiques*, Paris: Editions du CTHS, 39–60.
- Goebl, Hans (2010). „Dialectometry and quantitative mapping“, in: Lameli, Alfred/Kehrein, Roland/Rabanus, Stefan (Hgg.), *Language and Space. An International Handbook of Linguistic Variation*, Bd. 2: *Language Mapping* (= Handbücher der Sprach- und Kommunikationswissenschaft [HSK] 30.2.), Berlin: de Gruyter; 1. Teil: 433–457 (Text), 2. Teil (Karten): 2201–2212.
- Goebl, Hans (2012). „Introduction aux problèmes et méthodes de l'« École dialectométrique de Salzbourg » (avec des exemples gallo-, italo- et ibéroromans)“, in: Álvarez Pérez, Xosé Afonso/Carrilho, Ernestina/Magro, Catarina (Hgg.), *Proceedings of the International Symposium on Limits and Areas in Dialectology (LimiAr), Lisbon 2011*. Lissabon: Centro de Linguística da Universidade de Lisboa, 117–166.
- Goebl, Hans (2014). „L'impact de la polynomie des cartes d'atlas sur le résultat de calculs dialectométriques“, in: Polska Akademia Umiejętności. Instytut Filologii Romanskiej Uniwersytetu Jagiellońskiego (Hg.), *Linguistique romane et Linguistique indo-européenne. Mélanges offerts à Witold Mańczak à l'occasion de son 90<sup>e</sup> anniversaire*, Krakau: Polska Akademia Umiejętności. Instytut Filologii Romanskiej Uniwersytetu Jagiellońskiego, 243–260.
- Goebl, Hans (2018). „La face cachée de la géographie linguistique. Bref aperçu sur les « cartes muettes » produites pour l'ALF, l'PAIS et le FEW“, *Revue de Linguistique Romane* 82, 5–63.
- Goebl, Hans/Schiltz, Guillaume (1997). „Dialectometrical Compilation of CLAE 1 and CLAE 2. Isoglosses and Dialect Integration“, in: CLAE, Bd. II, 13-21 (siehe oben).

- Goebel, Hans/Smečka, Pavel (2016). „The Quantitative Nature of *Working Maps* (WM) and *Taxatorial Areas* (TA): A Brief Look at two Basic Units of *Salzburg Dialectometry* (S-DM)“, in: Kelih, Emmerich/Knight, Róisín/Mačutek, Ján/Wilson, Andrew (Hgg.), *Issues in Quantitative Linguistics 4. Dedicated to Reinhard Köhler on the occasion of his 65<sup>th</sup> birthday*, Lüdenscheid: RAM-Verlag, 113–127.
- Goebel, Hans/Smečka, Pavel (2017). „Zur Rolle der Polynymie im Rahmen der Salzburger Dialektometrie. Am Beispiel des italienischen Sprachatlases AIS“, in: Gerstenberg, Annette/Kittler, Judith/Lorenzetti, Luca/Schirru, Giancarlo (Hgg.), *Romanice loqui. Festschrift für Gerald Bernhard zu seinem 60. Geburtstag*, Tübingen: Stauffenburg, 2017, 137–154.
- Hoenigswald, Henry M. (1978). „The *annus mirabilis* 1876 und posterity“, *Transactions of the Philosophical Society* 76, 17–35.
- Jaberg, Karl (1906). „Zum Atlas linguistique de la France“, *Zeitschrift für romanische Philologie* 30, 512.
- Jaberg, Karl (1908). *Spracheographie. Beitrag zum Verständnis des Atlas linguistique de la France*, Aarau: Sauerländer (spanische Übersetzung: *Geografía lingüística. Ensayo de interpretación del “Atlas lingüístico de Francia”*, besorgt von Llorente, Antonio/Alvar, Manuel, Granada: Universidad de Granada. Secretariado de Publicaciones, 1959).
- Jaberg, Karl/Jud, Jakob (1928). *Der Sprachatlas als Forschungsinstrument. Kritische Grundlegung und Einführung in den Sprach- und Sachatlas Italiens und der Südschweiz*, Halle: Niemeyer (Neudruck: Nendeln [Liechtenstein]: Kraus, 1973; italienische Übersetzung: *Atlante linguistico ed etnografico dell'Italia e della Svizzera meridionale, vol. I: L'atlante linguistico come strumento di ricerca. Fondamenti critici e introduzione*, besorgt von Sanga, Glauco/Baggio, Serenella, Mailand: Unicopli, 1987).
- LAE = Orton, Harold/Sanderson, Stewart/Widdowson, John (Hgg.) (1978). *The Linguistic Atlas of England*, London: Croom Helm.
- Malkiel, Yakov (1967). „Each Word has a History of its Own“, *Glossa* 1, 137–149.
- Meyer, Paul (1875). Rezension zu: Ascoli (1874, siehe oben), *Romania* 4, 293–296.
- Paris, Gaston (1888). „Les parlers de France“, *Revue des patois gallo-romans* 2, 161–175 (erneut abgedruckt in: Roques, Mario (Hg.) (1909). Gaston Paris, *Mélanges linguistiques. Latin vulgaire et langues romanes, langue française, notes étymologiques*, Paris: Champion, 432–448).
- Putschke, Wolfgang (2001). „Die Dialektologie, ihr Beitrag zur historischen Sprachwissenschaft im 19. Jahrhundert und ihre Kritik am junggrammatischen Programm“, in: Auroux, Sylvain/Koerner, Ernst Fryderik Konrad/Niederehe, Hans-Josef/Versteegh, Kees (Hgg.) *History of Language Sciences. Geschichte der Sprachwissenschaften. Histoire des sciences du langage. An International Handbook on the Evolution of the Study of Language from the Beginnings to the Present. Ein internationales Handbuch zur Entwicklung der Sprachforschung von den Anfängen bis zur Gegenwart. Manuel international sur l'évolution de l'étude du langage des origines à nos jours* (= Handbücher der Sprach- und Kommunikationswissenschaft [HSK] 18.2.), Berlin/New York: Walter de Gruyter, Bd. 2, 1498–1513.
- Rosenqvist, Arvid (1919). „Limites administratives et division dialectale de la France“, *Neuphilologische Mitteilungen* 20, 87–118.

- Schuchardt, Hugo (1900 [1870]). *Über die Klassifikation der romanischen Mundarten* [Probe-Vorlesung gehalten zu Leipzig am 30. April 1870], Graz: Styria (Neuabdruck in: Spitzer, Leo (Hg.) (1928<sup>2</sup>). *Hugo Schuchardt-Brevier. Ein Vademecum der allgemeinen Sprachwissenschaft*, Halle: Niemeyer, 166–188; Neudruck: Tübingen: Niemeyer / Darmstadt: Wissenschaftliche Buchgesellschaft, 1976).
- SED = Orton, Harold/Halliday, Wilfrid J./Dieth, Eugen/Wakelin, Martyn F. (Hgg.) (1962–1971), *Survey of English Dialects. The Basic Material*, 12 Bde., Leeds: E. J. Arnold (Neudruck: London: Routledge, 1998).
- Séguy, Jean (1971). „La relation entre la distance spatiale et la distance lexicale“, *Revue de Linguistique Romane* 35, 335–357.
- Wartburg, Walther von (1918). *Zur Benennung des Schafes in den romanischen Sprachen: ein Beitrag zur Frage der provinziellen Differenzierung des spätern Lateins* (= Abhandlungen der Preussischen Akademie der Wissenschaften in Berlin, Philosophisch-historische Klasse 1918, Nr. 10), Berlin: Verlag der Königlichen Akademie der Wissenschaften, Reimer.
- Wartburg, Walther von (1963<sup>2</sup>). *Problèmes et méthodes de la linguistique*, Paris: Presses Universitaires de France [übersetzt von Pierre Maillard und Stephen Ullmann nach dem deutschen Original: *Einführung in die Problematik und Methodik der Sprachwissenschaft*, Tübingen : Niemeyer 1962<sup>2</sup>].
- Wenker, Georg (2013 [1889–1897]). *Schriften zum Sprachatlas des Deutschen Reichs. Gesamtausgabe. Band I: Handschriften: Allgemeine Texte. Kartenkommentare (1889–1897)*. Herausgegeben und bearbeitet von Alfred Lameli, Hildesheim/Zürich/New York: Olms, 2013.
- WGE = Orton, Harold/Wright, Nathalia (Hgg.) (1974). *A Word Geography of England*, London/New York: Seminar Press.

