The Handbook of Dialectology

Edited by

Charles Boberg, John Nerbonne, and Dominic Watt

WILEY Blackwell

This edition first published 2018
© 2018 John Wiley & Sons, Inc.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by law. Advice on how to obtain permission to reuse material from this title is available at http://www.wiley.com/go/permissions.

The right of Charles Boberg, John Nerbonne, and Dominic Watt to be identified as the authors of the editorial material in this work has been asserted in accordance with law.

Registered Offices
John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA

Editorial Office
9600 Garsington Road, Oxford, OX4 2DQ, UK

For details of our global editorial offices, customer services, and more information about Wiley products visit us at www.wiley.com.

Wiley also publishes its books in a variety of electronic formats and by print-on-demand. Some content that appears in standard print versions of this book may not be available in other formats.

Limit of Liability/Disclaimer of Warranty
While the publisher and authors have used their best efforts in preparing this work, they make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives, written sales materials or promotional statements for this work. The fact that an organization, website, or product is referred to in this work as a citation and/or potential source of further information does not mean that the publisher and authors endorse the information or services the organization, website, or product may provide or recommendations it may make. This work is sold with the understanding that the publisher is not engaged in rendering professional services. The advice and strategies contained herein may not be suitable for your situation. You should consult with a specialist where appropriate. Further, readers should be aware that websites listed in this work may have changed or disappeared between when this work was written and when it is read. Neither the publisher nor authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

Library of Congress Cataloging-in-Publication Data

Names: Boberg, Charles, editor. | Nerbonne, John A., 1951– editor. | Watt, Dominic James Landon, editor.
Title: The handbook of Dialectology / edited by Charles Boberg, John Nerbonne, Dominic Watt.
Description: First edition. | Hoboken, NJ : John Wiley & Sons, Inc., [2017] |
Series: Blackwell Handbooks in linguistics | Includes index.
Identifiers: LCCN 2017017090 (print) | LCCN 2017029760 (ebook) | ISBN 9781118827598 (pdf) |
ISBN 9781118827581 (epub) | ISBN 9781118827550 (cloth) | ISBN 9781119068419 (pbk.)
Subjects: LCSH: Dialectology–Handbooks, manuals, etc. | Language and languages–Variation–Handbooks, manuals, etc.
Classification: LCC P367 (ebook) | LCC P367 .H46 2017 (print) | DDC 417/.2–dc23
LC record available at https://lccn.loc.gov/2017017090

Cover image: © 2015. Digital image, The Museum of Modern Art, New York/Scala, Florence
Cover design: Wiley

Set in 9.5/11.5pt Palatino by SPi Global, Pondicherry, India

10 9 8 7 6 5 4 3 2 1

Contents

Blackwell Handbooks in Linguistics

This outstanding multi-volume series covers all the major subdisciplines within linguistics today to offer a comprehensive survey of linguistics as a whole. To see the full list of titles available in the series please visit www.wiley.com/go/linguistics-handbooks

The Handbook of Child Language
Edited by Paul Fletcher & Brian MacWhinney

The Handbook of Phonological Theory, Second Edition
Edited by John A. Goldsmith, Jason Riggle, & Alan C. L. Yu

The Handbook of Contemporary Semantic Theory
Edited by Shalom Lappin

The Handbook of Sociolinguistics
Edited by Florian Coulmas

The Handbook of Phonetic Sciences, Second Edition
Edited by William J. Hardcastle & John Laver

The Handbook of Morphology
Edited by Andrew Spencer & Arnold Zwicky

The Handbook of Japanese Linguistics
Edited by Natsuko Tsujimura

The Handbook of Linguistics
Edited by Mark Aronoff & Janie Rees-Miller

The Handbook of Contemporary Syntactic Theory
Edited by Mark Baltin & Chris Collins

The Handbook of Discourse Analysis
Edited by Deborah Schiffrin, Deborah Tannen, & Heidi E. Hamilton

The Handbook of Language Variation and Change, Second Edition
Edited by J. K. Chambers & Natalie Schilling

The Handbook of Historical Linguistics
Edited by Brian D. Joseph & Richard D. Janda

The Handbook of Language, Gender, and Sexuality, Second Edition
Edited by Susan Ehrlich, Miriam Meyerhoff, & Janet Holmes

The Handbook of Second Language Acquisition
Edited by Catherine J. Doughty & Michael H. Long

The Handbook of Bilingualism and Multilingualism, Second Edition
Edited by Tej K. Bhatia & William C. Ritchie

The Handbook of Pragmatics
Edited by Laurence R. Horn & Gregory Ward

The Handbook of Applied Linguistics
Edited by Alan Davies & Catherine Elder

The Handbook of Speech Perception
Edited by David B. Pisoni & Robert E. Remez

The Handbook of the History of English
Edited by Ans van Kemenade & Bettelou Los

The Handbook of English Linguistics
Edited by Bas Aarts & April McMahon

The Handbook of World Englishes
Edited by Braj B. Kachru; Yamuna Kachru, & Cecil L. Nelson

The Handbook of Educational Linguistics
Edited by Bernard Spolsky & Francis M. Hult

The Handbook of Clinical Linguistics
Edited by Martin J. Ball, Michael R. Perkins, Nicole Müller, & Sara Howard

The Handbook of Pidgin and Creole Studies
Edited by Silvia Kouwenberg & John Victor Singler

The Handbook of Language Teaching
Edited by Michael H. Long & Catherine J. Doughty

The Handbook of Language Contact
Edited by Raymond Hickey

The Handbook of Language and Speech Disorders
Edited by Jack S. Damico, Nicole Müller, & Martin J. Ball

The Handbook of Computational Linguistics and Natural Language Processing
Edited by Alexander Clark, Chris Fox, & Shalom Lappin

The Handbook of Language and Globalization
Edited by Nikolas Coupland

The Handbook of Hispanic Sociolinguistics
Edited by Manuel Díaz-Campos

The Handbook of Language Socialization
Edited by Alessandro Duranti, Elinor Ochs, & Bambi B. Schieffelin

The Handbook of Intercultural Discourse and Communication
Edited by Christina Bratt Paulston, Scott F. Kiesling, & Elizabeth S. Rangel

The Handbook of Historical Sociolinguistics
Edited by Juan Manuel Hernández-Campoy & Juan Camilo Conde-Silvestre

The Handbook of Hispanic Linguistics
Edited by José Ignacio Hualde, Antxon Olarrea, & Erin O'Rourke

The Handbook of Conversation Analysis
Edited by Jack Sidnell & Tanya Stivers

The Handbook of English for Specific Purposes
Edited by Brian Paltridge & Sue Starfield

The Handbook of Spanish Second Language Acquisition
Edited by Kimberly L. Geeslin

The Handbook of Chinese Linguistics
Edited by C.-T. James Huang, Y.-H. Audrey Li, & Andrew Simpson

The Handbook of Language Emergence
Edited by Brian MacWhinney & William O'Grady

The Handbook of Korean Linguistics
Edited by Lucien Brown & Jaehoon Yeon

The Handbook of Speech Production
Edited by Melissa A. Redford

The Handbook of Contemporary Semantic Theory, Second Edition
Edited by Shalom Lappin & Chris Fox

The Handbook of Classroom Discourse and Interaction
Edited by Numa Markee

The Handbook of Narrative Analysis
Edited by Anna De Fina & Alexandra Georgakopoulou

The Handbook of English Pronunciation
Edited by Marnie Reed & John M. Levis

The Handbook of Discourse Analysis, Second Edition
Edited by Deborah Tannen, Heidi E. Hamilton, & Deborah Schiffrin

The Handbook of Bilingual and Multilingual Education
Edited by Wayne E. Wright, Sovicheth Boun, & Ofelia García

The Handbook of Portuguese Linguistics
Edited by W. Leo Wetzels, João Costa, and Sergio Menuzzi

The Handbook of Linguistics, Second Edition
Edited by Mark Aronoff and Janie Rees-Miller

The Handbook of Translation and Cognition
Edited by John W. Schwieter and Aline Ferreira

The Handbook of Dialectology
Edited by Charles Boberg, John Nerbonne, and Dominic Watt

List of Contributors

Alexander Adelaar
The University of Melbourne,
Australia

Ronelle Alexander
University of California, Berkeley, CA, USA

Enam Al-Wer
University of Essex, Colchester, UK

Lieselotte Anderwald
University of Kiel, Germany

Peter Auer
Department of German Linguistics,
University of Freiburg,
Germany

Guy Bailey
University of Texas Rio Grande Valley,
Texas, USA

Charles Boberg
McGill University, Montreal, Canada

David Britain
Department of English, University of Bern,
Switzerland

J.K. Chambers
University of Toronto, Canada

Ashwini Deo
The Ohio State University, USA

Jacob Eisenstein
Georgia Institute of Technology,
USA

Hans Goebl
Department of Romance Studies,
University of Salzburg, Austria

Charlotte Gooskens
Center for Language and Cognition,
University of Groningen, Netherlands

Matthew J. Gordon
Department of English, University of
Missouri, Columbia, Missouri, USA

Jack Grieve
Aston University, UK

Damien Hall
Newcastle University, Newcastle, UK

Wilbert Heeringa
Fryske Akademy, Netherlands

Raymond Hickey
Institute for Anglophone Studies,
University of Duisburg and Essen,
Germany

Frans Hinskens
Meertens Instituut (KNAW) & Vrije
Universiteit Amsterdam, Netherlands

Rudolf de Jong
University of Leiden, Netherlands-Flemish
Institute in Cairo

Jacques Van Keymeulen
Department of Linguistics – Dutch,
Ghent University, Belgium

William A. Kretzschmar, Jr.
Department of English, University of Georgia, Athens, Georgia, USA

Tore Kristiansen
Nordic Research Institute, University of Copenhagen, Denmark

Sebastian Kürschner
Catholic University of Eichstätt-Ingolstadt, Germany

John M. Lipski
Pennsylvania State University, State College, PA, USA

Carmen Llamas
University of York, UK

Ronald Macaulay
Pitzer College, USA

John Nerbonne
University of Groningen, Netherlands & University of Freiburg, Germany

Takuichiro Onishi
National Institute for Japanese Language and Linguistics, Tokyo, Japan

John C. Paolillo
Indiana University Bloomington, USA

Dennis R. Preston
Department of English, Oklahoma State University, Stillwater, Oklahoma, USA

Jelena Prokić
University of Marburg, Germany

Stefan Rabanus
University of Verona, Italy

Benedikt Szmrecsanyi
University of Leuven, Belgium

Chaoju Tang
University of Electronic Science and Technology of China, Chengdu, Sichuan, P.R.China

Tullio Telmon
Università degli Studi di Torino, Torino, Italy

Erik R. Thomas
North Carolina State University, USA

Kevin Watson
University of Canterbury, Christchurch, New Zealand

Dominic Watt
Department of Language and Linguistic Science, University of York, UK

Martijn Wieling
University of Groningen, Netherlands

Vladimir Zhobov
Sofia University, Sofia, Bulgaria

7 Dialectometry

HANS GOEBL

7.1 Introduction

In my treatment of dialectometry in this chapter, I will consider the topic from the standpoint of a Romance geolinguist. It seems necessary to emphasize this point, because we can observe a certain dissimilarity in the linguistic geographies of Romance, German, and English linguistics, a fact which sometimes complicates understanding. I would also like to emphasize that my contribution relies completely and solely on data from linguistic atlases. Whenever I subsequently speak of "geolinguistic variation," the reader should be aware that my ideas are based on linguistic atlas data.

Dialectometry might thus more properly be called *atlantometry*. The fundamentally *inductive* character of dialectometry must also be pointed out. The preponderant aim of dialectometry consists in discovering, via the numerical analysis of many concrete patterns,[1] more abstract patterns which would otherwise remain hidden. We do so in order to obtain a systematic insight into the problem of, so to speak, the "basilectal management of space by *Homo loquens*." The primary motives for research in dialectometry are thus linguistic, and have initially nothing to do with quantity. Many dialectometricians have adopted a similar position, including in particular, Jean Séguy, the original creator of the term and the method of "*dialectométrie*" (Séguy 1973: 1).

Historically speaking, the position of Séguy (1914–1973) is that of a single link in a long chain. Obviously, this metaphorical chain is located in France, and it refers to some very important chapters of French spiritual and scientific evolution. Séguy thus has many forerunners whose contributions and importance will be presented below (see also the historical overviews given in Goebl, 2006b, 2013d).

7.2 Historical Background A: From the *Ancien Régime* up the End of the Nineteenth Century

Our historical retrospective starts with the gradual emergence of a "geodetic" conception of the territory of France under the *Ancien Régime*, the political system in force during the period from the sixteenth to the late eighteenth centuries. During this time, a large number of geographers, economists, tax collectors, and military engineers tried not only to measure the size of the territory of France but also to grasp the importance of certain socio-economic

The Handbook of Dialectology, First Edition. Edited by Charles Boberg, John Nerbonne, and Dominic Watt.
© 2018 John Wiley & Sons, Inc. Published 2018 by John Wiley & Sons, Inc.

variables in order to inform the king about the riches of his realm. Witness, for instance, the activity and writings of four generations of the Cassini family (from Jean Dominique Cassini, 1625–1712, until Jean Dominique, comte de Cassini, 1748–1845).

As a result, the general idea developed—obviously in a slow and rather subconscious manner—that the space of France was a kind of machine capable of producing a certain quantity of goods, with gears that meshed according to certain inner principles.

The many reforms of the French Revolution (1789–1799) and the specifically centralistic spirit of the Napeolonic administration (1799–1814) considerably accelerated new concepts of space that were emerging at the time. The most important event in this respect was the abolition of the old subdivision of France into its historical provinces, and the establishment of the Republican grid of departments. The central philosophy of the new grid had two characteristics of key importance from the *metrological* point of view: (a) the equal size and the even distribution of the different departments (whose number initially was 83) over France, and (b) the definition of a canonical *spatial sampling* for further research.

The new grid was in fact used immediately not only by the Napoleonic administration but also by the subsequent regimes (Monarchies of Bourbon and Orleans 1814–1848, Second Republic 1848–1852, Second Empire 1852–1870, and Third Republic 1871–1940).

From the linguistic point of view, the most interesting activity of the Napoleonic initiatives was the standardized collection of a great number of dialectal translations of the *Parable of the Prodigal Son* undertaken by Charles Etienne Coquebert de Montbert (1755–1831) and his son Barthélémy Eugène (1785–1847; see further, Pop, 1950, *passim*). Charles Etienne should not be considered a linguist as such, but rather a statistician in the spirit of the eighteenth century who was inspired by a wide-ranging empirical curiosity.

Painstaking collection of a wide range of empirical data continued further under the re-established monarchy and gave rise to a very sophisticated investigation of the national space of France. The collection of the data was, naturally, based on the departmental grid.

However, the most crucial problem was the quantitative processing and subsequent visualization of the data collected. In this respect, from throughout the nineteenth century one can find scholarly treasures, which unfortunately are little known outside France: see for instance Gilles Palsky's (1996) superb *Des Chiffres et des Cartes*, which not only contains a good description of the general evolution of these investigations but also a great number of excellent photographic reproductions of the respective visualizations.

As a result, one can assume that at the end of the nineteenth century practically all political and intellectual figures in France had good familiarity with graphics and visual representations giving evidence of the geographical arrangement of a very great number of social, economic, and demographic *mechanismes* of national interest.

7.3 Historical Background B: Jules Gilliéron and ALF (*Atlas linguistique de la France*)

Jules Gilliéron (1854–1926) was born in the French-speaking part of Switzerland and moved permanently to Paris at the age of 22 (Pop and Pop 1959: 5–19). He rapidly came into contact with eminent personalities of the Parisian academic scene. From 1883 onward he was charged with the teaching of "*dialectologie de la Gaule romane*" at the École Pratique des Hautes Études, a position in which he continued until his death. What he encountered in Paris from a geolinguistic point of view was a great number of unsolved questions referring to the linguistic subdivision of France in the present and in the past, and many complaints about insufficient data in this connection. His chief intellectual mentor was the great French philologist Gaston Paris (1839–1903). One of the major concerns in the 1870s and 1880s was the question of whether the intellectual concept of dialect had a real counterpart in reality and could therefore "exist." In France, the general opinion denied the existence of dialects but fully recognized the real existence of a great number of single linguistic features. Gaston Paris summarized this fact in 1888 as follows: "Il faut faire la géographie non pas des *dialectes*, mais des *traits linguistiques*" ("We must make the geography not of dialects, but of linguistic traits"). Unfortunately, there was no precise knowledge about the exact geographical extent of such a great number of different dialectal features. This gap would be filled by ALF.

Before starting the work for ALF, Gilliéron exercised his methodological ideas in the Swiss canton of Valais, publishing in 1881 a small phonetic atlas, which could be considered, to some extent, as a forerunner of ALF.

However, what he prepared between 1881 and 1897 was quite different from this little test piece. When the real fieldwork for the future ALF started in 1897, Gilliéron had already elaborated a precise research agenda:

- *Theoretically*: the great challenge was to determine the geographical range of a large number of basilectal geolinguistic features belonging to different linguistic categories (phonetics, morphology, vocabulary, etc.).
- *Practically*: the inquiries should be done *in loco* by contacting bilingual people (French and local dialect) and observing two strict principles. The fieldworker, Edmond Edmont, was told (a) to transcribe only the first answer given by the interviewee, and (b) to avoid any "extortion" of further (multiple) responses. The challenge was therefore to elicit only the *basilectal component* of the multiple competence of the interviewees.

These two constraints had a profound impact on the quality of the ALF data, and guaranteed its perfect commensurability. In just four years (1897–1901), Edmont succeeded in visiting 638 localities spread evenly over the Romance-speaking parts of France, Belgium, and Switzerland, and neighboring areas (the Channel Islands and Piedmont, Italy) where Galloromance varieties were spoken. In his peregrinations he used three questionnaires, starting with a set of 1,421 questions, which was later further enlarged to 1,920 items.

Edmont's astonishing accomplishment was enhanced by the rapid publication of the collected data between 1902 and 1910. As a result, the completed ALF comprised 10 in-folio volumes with full-text maps in which the reader could find, for three geographically coherent sections (series A: whole grid; series B: southern part of the grid; series C: south-eastern part of the grid), the transcriptions produced by Edmont and minimally corrected by Gilliéron.

Another issue of great importance, as we shall see in the next section, was the parallel publication of blank maps (or *cartes muettes*, "mute maps") of the ALF grid (with 638 localities, sites, or points) and their diffusion among interested scholars.

7.4 The Practical and Theoretical Importance of ALF

ALF's contribution to Romance linguistics and philology all over Europe was immediate and substantial. An important factor in this success was Gilliéron's teaching at the École Pratique des Hautes Études. His classes were attended by a large group of upcoming Romanists from all parts of Europe. The courses Gilliéron gave became legendary (Pop and Pop 1959: 53–63). This was true also of his publications, which resonate with a great personal commitment and are full of innovative linguistic ideas (for an excellent example, see Gilliéron 1918).

One of the main pillars of Gilliéronian linguistics was the systematic study of geolinguistic feature areas. While working with the aforementioned mute maps in order to exploit the raw data transcriptions of single ALF maps, Gilliéron (like other Romance scholars, e.g., Jaberg, 1908) noticed that the diffusion areas of different linguistic features could vary considerably according to size, shape, and geographic location. So as to understand and explain this (initially very strange) variability, Gilliéron developed a special methodology called *aréologie*, in which the different *aires* are regarded as results of processes of diffusion, retraction, and resistance, all of which result from the metalinguistic actions of dialect speakers. We re-encounter here the old idea that relations in space are the outcomes of specific human behavior.

In according "metalinguistic responsibility" to dialect speakers, Gilliéron developed a psycholinguistic theory, whereby factors such as linguistic creativity and the management of homonymy had pivotal roles. Note that Gilliéron utilized and analyzed the ALF data by looking exclusively at single atlas maps and avoiding any data synthesis.

Let us, however, return to the use of mute maps, which also played a central part in the new geolinguistic conceptions. The use of mute maps always required the following steps:[2]

- the choice of a classification criterion in order to extract and visualize some specific properties from the raw data of a given atlas map,
- the choice of the visualization mode: signatures that are either *spatial* (i.e., using areas) or *linear* (using isoglosses),
- the projection of the selected signatures (in color or in black and white) onto the blank form of the mute map.

Obviously, the graphic quality of such cartographic exercises could vary considerably: the maps could be produced for personal study or for publication, and they might assume different cartographic forms depending on the geolinguist's drawing ability. What is really important, however, is the fact that following the publication of ALF it was clear that no one would be able to avoid intelligent work with mute maps. In Germanic philology, incidentally, linguistic atlases never offer their data in raw form but instead present them in the form of symbol maps. These are just a particular classification of the raw data,[3] which often remain completely invisible or inaccessible to the user. In such circumstances, it is obviously difficult to develop clear ideas about the classification of atlas data.

Whereas Gilliéron's teaching systematically neglected diachrony, many of his Swiss, German, and Austrian followers used the ALF data for diachronic studies, starting from two questions: (a) what were the sizes, shapes, and geographic locations of ALF feature areas in the past? and (b) is it possible to extract, from medieval data, comparable areal information so as to reconstruct diachronic evolution over the course of two or three centuries?

From our dialectometric viewpoint, Gilliéron, his ALF, and the newborn ALF geolinguistics are highly important for several reasons: for the excellent metrological quality of the ALF data, for the revitalization of the idea that spatial relations depend upon human behavior, and for data classification based on mute maps.

Unfortunately, all these advantages are scarcely known outside Romance philology. Another particularity of Romance linguistics and philology is worth mentioning, that of the tight link established between linguistic geography and a wide range of sub-branches of Romance philology (text philology, historical grammar, lexicology, and etymology).

ALF's example was rapidly imitated (see Chambers' chapter on written surveys, this volume) and applied to other great Romance-speaking domains such as Italy (AIS, 1928–1940, created by the Swiss scholars Jakob Jud and Karl Jaberg), Romania (ALR, 1938–1942, under the responsibility of Sextil Puşcariu, Sever Pop, and Emil Petrovici), and Catalonia (ALC, 1923–1964, under the direction of Antoni Griera). Fortunately, the basic methodological principles and assumptions of ALF were not altered in the process.

7.5 From ALF to Jean Séguy's *Dialectométrie*

After World War II, regrettably, the empirical guidelines observed by Gilliéron when compiling the ALF were slowly forgotten in France. The metrological status of ALF as a kind of "glotto-geodesy" of France was replaced by the rather uncritical search for new, "naturalistic," and linguistically attractive data. The nationwide perspective of ALF stimulated the appetite for analogous but regional perspectives, but the rather unspecific character of the ALF questionnaire engendered the desire for regional questionnaires with (much) more specific items. Furthermore, standardized inquiries were renounced for the sake of more authenticity and naturalism.

One of the first Parisian scholars of Gilliéron, Albert Dauzat (1877–1955), tried at the end of the 1930s to meet all of these requirements in a new research project called *Nouvel Atlas Linguistique de la France* (NALF), attempting above all to maintain the Gilliéronian demand for commensurability in the collected data. For the sake of NALF—which, from the beginning, meant a "family" of linguistic atlases each related to single historical regions such as Normandy, Provence, Gascony, and so on—Dauzat recommended that the different questionnaires should have approximately the same extent, and that a third of the items they contained should be identical.

Owing to the appearance of a new generation of geolinguists who no longer respected (or understood?) either Gilliéron's or Dauzat's principles, the new regional atlases, whose compilation started at the beginning of the 1950s, evolved in a completely different direction. They were characterized by the following principles: a highly regionalized choice of items for (NALF) questionnaires; the establishment of new inquiry grids, very often to the complete exclusion of the old ALF sites; the de-standardization of the data collection by applying guided or completely free conversation with informants, instead of standardized questioning; the conscious elicitation of multiple responses by "squeezing" the multiple competences of the informants; and the use of local dialects during the inquiries, instead of normal French.

It is obvious from the viewpoint of Gilliéronian glotto-geodesy that these data collection methods no longer had the same quality as those of ALF. The new regional linguistic atlases did not evolve toward the production of a second, more detailed layer of glotto-geodetic analyses of different parts of France that would allow comparative insights into the geolinguistic dynamics of a period of about 50 years. Instead, they went in the direction of compiling large-scale geographically stratified regional vocabularies.

At the end of the twentieth century the complete list used for NALF (which was later renamed to the *Atlas Linguistique de la France par Régions*) contained 25 titles, the investigation grids of which cover the whole territory of France.[4] Without doubt, they all have considerably enlarged our knowledge of many French regional dialects. But they cannot be used for any large-scale comparison. Across the 25 questionnaires there are not even ten items which they all have in common. In short, the NALF atlases reflect a completely different geolinguistic methodology that abandons the requirement that the data it produces should be commensurable. It is no longer glotto-geodesy as per ALF, but data collection for its own sake.

One of these new regional atlases was the *Atlas Linguistique de la Gascogne* (ALG) compiled by Jean Séguy (1914–1973). ALG was elaborated and published in six folio volumes between 1954 and 1974. Séguy started ALG fully aware of the directives given by Dauzat, but on the way yielded further and further to what we might call "data collection syndrome." Finally, while trying to exploit the inner structures of ALG's data, he neared desperation, confronted as he was by the endless variability in the size, shape, and geographic location among the different feature areas on ALG's mute maps. This was the moment of his conversion to quantitative thinking and the beginning of his efforts to bring order to this apparent chaos.[5]

Séguy's truly dialectometrical writings consist of two articles (Séguy 1971, 1973) and some interesting pages and maps at the end of the sixth volume of ALG. In these writings he deals—always in a very elementary way, and using rudimentary graphic devices—with some problems of quantification with respect to measuring linguistic distances (not similarities!) between two neighboring sites, for which he analyzed data taken not only from ALG but also from other Romance linguistic atlases. In spite of these methodological weaknesses, Séguy's position is nevertheless seminal, in the following sense. Behind (or below) the apparent chaos of the ALG data and that of other atlases, he conjectured the existence of an underlying primary order which could be explored and expressed in quantitative terms. Séguy assumed the existence of a quantitative relation between geographic and linguistic distances, and conducted simple experiments with it, attempting to apply distance formulae taken from general statistics to ALG data and those from other linguistic atlases. In 1973, he created the neologism *dialectométrie*.

7.6 Regensburg-Salzburg Dialectometry (RS-DM): Theoretical Assumptions and Practical Achievements

My own dialectometric work began at Regensburg University, to which I was affiliated between 1973 and 1982, and continued later in Salzburg. Unfortunately, Séguy's early death in 1973 prevented any collaboration. I started from the following theoretical assumptions.

The continuous changing of geolinguistic feature areas according to their size, shape, and geographic location seems to be a linguistic universal. Moreover, this fact accounts for the (often bemoaned) non-coincidence of isoglosses. It also seems to be the motivation for the claim made by many linguists that "each word has its own history,"[6] which, for the sake of linguistic geography, should be translated "each *feature* has its own *area*." In a given geolinguistic dataset the total number of areas seems to play the same role as the totality of words in a given text. It would thus be interesting to compare the respective findings in quantitative linguistics.

The inner structure of different linguistic atlases (like ALF, AIS, and others) depends on the interplay of a large number of single feature areas belonging to different linguistic categories (phonetics, morphology, lexis, etc.). If there are regularities in the spatial distribution of these areas, they can only be found via the scrutiny of a large set of areas. Given the very visual character of (Romance) linguistic geography, the central heuristics of the new investigations must continue to be essentially cartographic, obviously using quantitative instead of qualitative cartography.

At the same time (1973 ff.), the practical considerations were as follows:

1. There were some new calculation machines called *computers*, which seemed to be useful for the treatment of mass data.
2. Are there methods of quantitative classification in other fields with similar statistical needs and experiences?
3. Is there previous experience with the quantitative mapping of a great amount of geographically-dispersed numerical data?
4. Can the new *computers* also be applied to cartography?

With respect to (1), I succeeded very quickly in establishing collaborations with computer specialists by avoiding any *bricolage* and the relative shortcomings of do-it-yourself solutions. Secondly, I soon learned of "numerical classification" (*numerische Klassifikation*, *classification automatique*), and could even come into direct contact with some of its representatives (e.g., H. H. Bock, R. R. Sokal).[7] It was also very useful to examine interdisciplinary case studies, mainly in the field of biology, economics and quantitative geography (see Haggett 1965). In addition, I came into contact with the German quantitative linguist Gabriel Altmann at Bochum University, and some of his students.

The answer to (3) was positive. In the early 1970s there already existed good handbooks and reviews on quantitative cartography written in German, French and English (e.g., Dickinson 1973), in which the statistical, visual, and cartographic problems associated with the making of quantitative maps were extensively discussed.

As for (4), the use of computers in cartography did not begin until the end of the 1970s. The integration of shadings, hatchings and colors into this process did not proceed without problems.

7.7 Some Methods and Findings of the RS-DM[8]

I will now briefly discuss four important methods used by RS-DM[9] using four cartographic examples.

7.7.1 Step A: Taxatation and Presentation of Plate 1

The raw data consist of 626 original ALF maps from which we have extracted—by means of phonetic, morphologic, and lexical taxatation—1,681 "working maps" (WM) (See Figures 7.1 and 7.2, and Plate 1 (map)).

The main goal of taxatation is the areal classification of the original maps on the basis of specific linguistic criteria. In our case, these criteria belong to Romance historical phonetics and lexicology. In the 1,681 WMs we find 18,047 feature areas (and therefore as many linguistic elements called *taxates*) that vary greatly in size, shape, and geographic location. As to their inner fragmentation, see the histogram in Figure 7.2., which shows that there are many roughly structured WMs but few WMs of fine granularity.

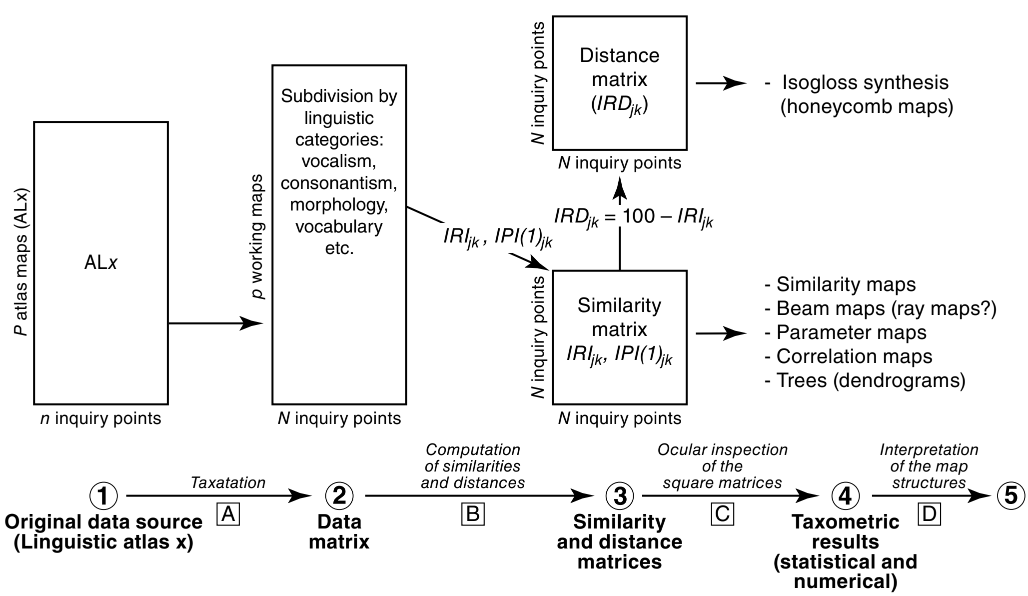

Figure 7.1 Flow chart of the methods used by the Regensburg-Salzburg-dialectometry.

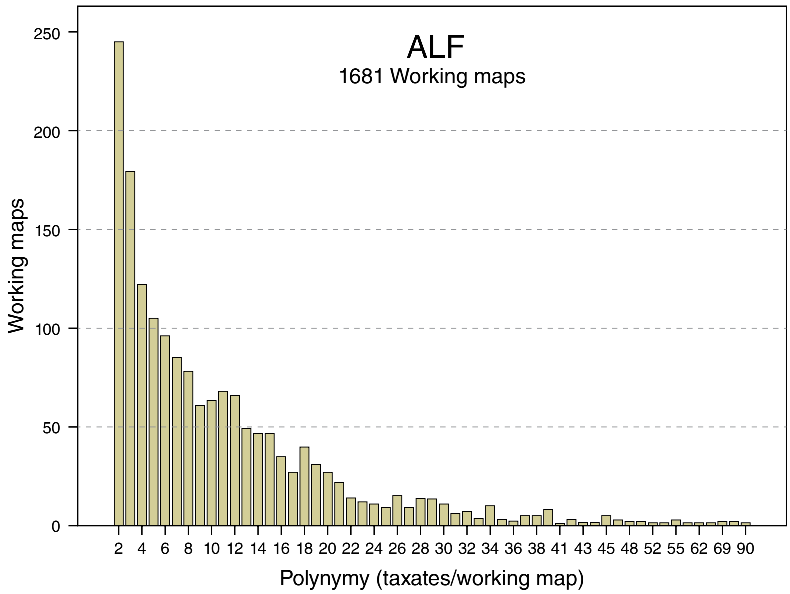

Figure 7.2 Histogram of the total ALF-corpus, showing the granulation and frequency of 1681 working maps (WMs) belonging to all linguistic categories. For better understanding: the granulation of the WMs starts with 2 taxates/WM (valid for 245 WMs) and ends with 90 taxates/WM (valid for 1 WM).

The exponential-like regularity of the decreasing curve has been found in all our dialectometric analyses.[10] It seems to be a direct consequence of the collective dialectal behavior of the French speakers. As early as 1985, Altmann described this situation as the result of the antagonistic interplay of birth and death processes, which were obviously related to the adoption and atrophy of linguistic features, which, in the present case, are associated with areas.

We conjecture that these spatially related regularities have the same law-like status as the well-known *Lautgesetze* (sound laws), discovered by the Neogrammarians of the end of the nineteenth century, whose regularities are related to time.

Plate 1 was established on the basis of the ALF map 18, *l'aile* ("the wing"). It shows 11 different outputs (*taxates* and their *areas*) of the final Latin –A (in the Latin etymon ÁL[A]). This etymon occurs at all 641 ALF sites. The most frequent output (at 350 ALF sites) is the taxate *zéro* (Ø, i.e., the Latin final –A vanishes completely). The other ones are *-à* (at 70 ALF sites) and *-ó* (at 69 ALF sites), and so on, as shown in the legend.

The changing granularity of the WMs is called *polynymy*. Theoretically, it ranges from 2 to N (the maximal number of localities in a given grid). Plate 1 is therefore 11-nym: it is a map of medium granularity from which we can derive, in the present ALF taxatation, 67 other specimens.

The geographic structure of the map in Plate 1 clearly shows the north-south division of the Gallo-Romance domain. It suggests that the Northern taxate 1 (zero) expanded at the expense of *-a* or *-o*, which are closer to their Latin origin. By including diachronic information derived from the different parts of the Gallo-Romance domain, the linguistic interpretation of the map profile can be further refined and completed.[11]

The data matrix in Figure 7.3 shows that every character vector of an atlas site consists of areas of different sizes. It could thus be useful to analyze this variation quantitatively. It is

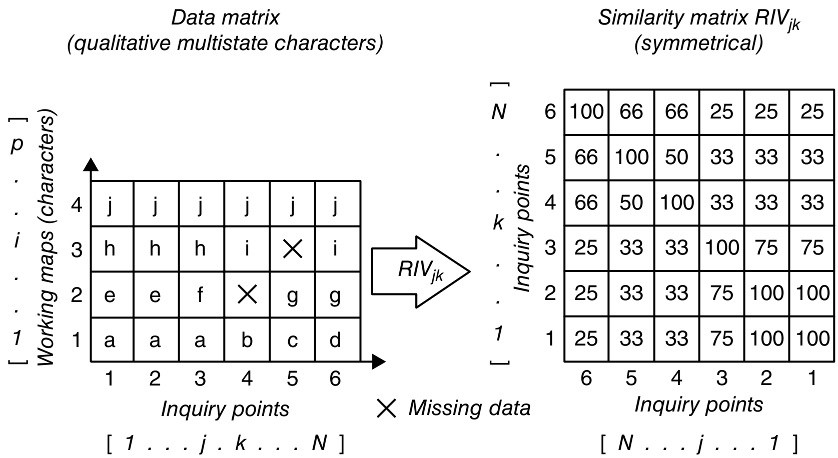

Figure 7.3 Data matrix and similarity matrix. Scheme of calculation of the interdialectal similarities via RIV_{jk} (Relative Identity Value).

necessary, obviously, to organize the data matrix according to different linguistic categories if we wish to compare the different computations with one another.

Since the 1990s we have used VDM (*Visual DialectoMetry*), a computer program created and further improved by Edgar Haimerl, for managing all the steps of the numerical and visual processing of our data.[12] The dialectometric basics implemented in VDM are drawn from our handbook (Goebl, 1984).

7.7.2 Step B: Computation of Similarities and Distances

The main result of step B is the change in the ontological nature of the investigated data, namely from *qualitative* to *quantitatitive* (see Figure 7.3). Statistically speaking, this happens by measuring the similarities between the N locality vectors. As the handbooks of numerical classification offer a variety of similarity indices, one should select an index that fits one's concepts of interdialectal similarity.

In this instance, the Relative Identity Value (RIV_{jk}) has proven very successful. It is calculated using the number of pairwise matchings (also called co-identities, or COI) and the number of pairwise mis-matchings (co-differences, COD) of taxates. RIV_{jk} values range between 0 and 100%, according to the formula at (1):

$$\mathrm{RIV}_{\mathrm{jk}} = 100 \sum \mathrm{COI}(i)_{\mathrm{jk}} / \sum \mathrm{COI}(i)_{\mathrm{jk}} + \sum \mathrm{COD}(i)_{\mathrm{jk}} \tag{1}$$

Table 7.1 shows the meaning of these symbols.

The calculated similarity values will be stored in the (square) similarity matrix (see Figure 7.1, point 3), which can easily be converted into a distance matrix by applying the following formula: *distance* (*RDV*) = 100 – *similarity* (*RIV*). The two matrices are depositories of the totality of the pairwise relations (be they similarities or distances) that exist between the N investigated sites.

Because the square matrices consist of two symmetrical halves and the scores located along the diagonal are irrelevant, the number of the valid similarity or distance values is N/2(N–1).

In the case of ALF, we are faced with 638 original sites that have been augmented by three artificial sites corresponding to the standard languages French, Italian, and Catalan. Thus, we have 641 sites and 205,120 similarity or distance values for further processing.

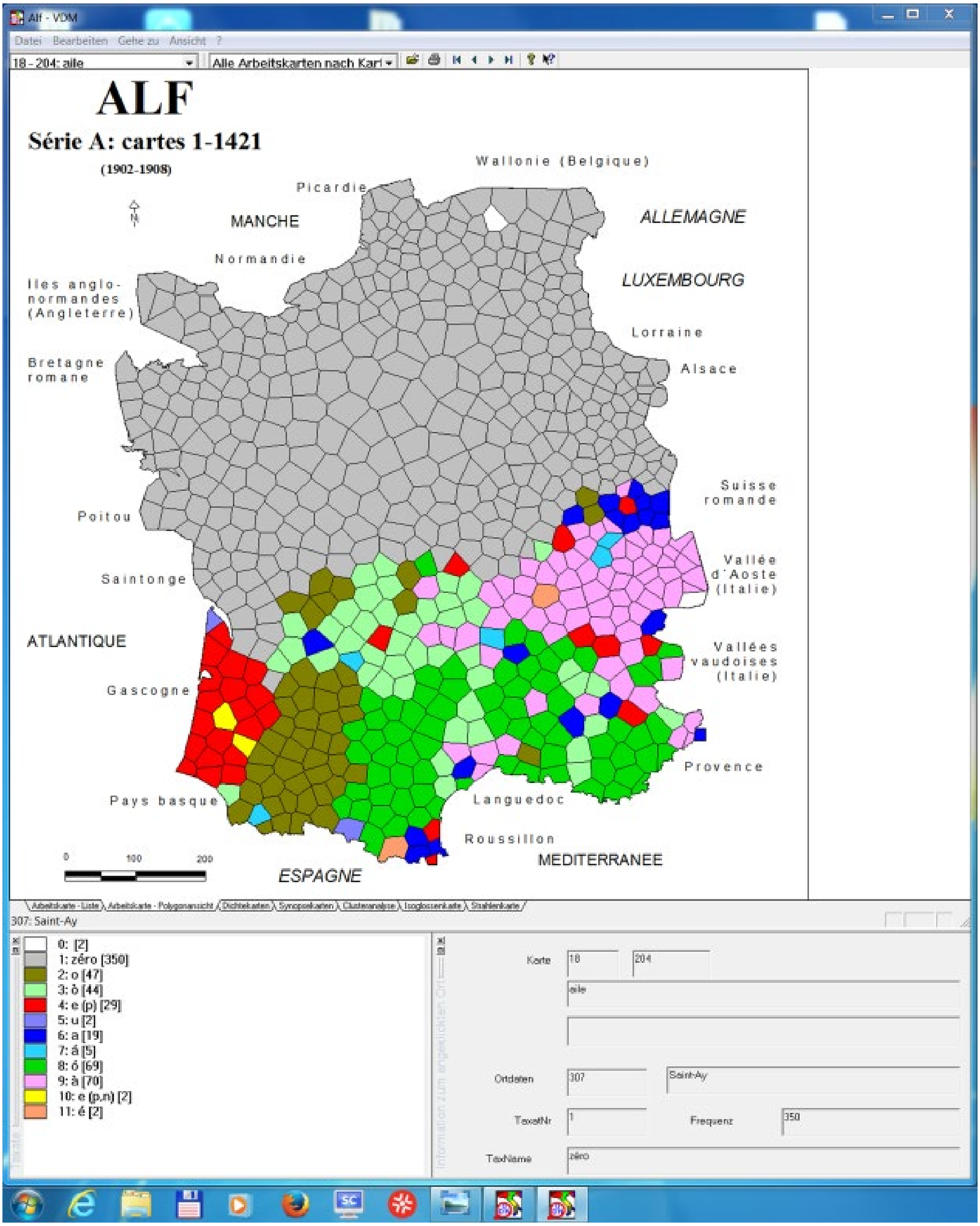

Plate 1 Sample of a *phonetic* working map: spatial distribution of the Gallo-Romance results of final -A in the Latin etymon ÁLA (< Fr. *aile*) "wing" (following ALF 204 aile). Cartographic status: *qualitative* choropleth map. See Section 7.7.1, pp. 129–131. (*See insert for colour representation of the figure*.)

Table 7.1 Meanings of the symbols used in the formula at (1).

Symbol	*Meaning*
$COD(i)_{jk}$	co-difference between two taxates (on the map *i* and for the sites *j* and *k*)
$COI(i)_{jk}$	co-identity between two taxates (on the map *i* and for the sites *j* and *k*)
i	one of *p* working maps
j	reference site
k	site (atlas point) to be compared with the reference site *j*
p	total number of the working maps in the data matrix
N	total number of the sites (atlas points) in the data matrix
RIV_{jk}	Relative Identity Value (between the attribute vectors of the sites *j* and *k*)
RDV_{jk}	Relative Distance Value (between the attribute vectors of the sites *j* and *k*)

7.7.3 Step C: Visualization

According to the cartographic tradition of linguistic geography, the subsequent processing of the calculated data has to be visual. The data output thus moves from numerical to visual. The cartographic status of the new maps or schemes will no longer be *qualitative* (as in Plate 1), but henceforth *quantitative* (as in Plate 2, Plate 3, and Plate 4).

Obviously, this change should be done algorithmically. In this respect, the quantitative branch of thematic cartography offers a series of very useful solutions, which have been partly incorporated into VDM (see Dickinson 1973 *et passim*, and Goebl (1984, 86, 113)).

VDM provides different choropleth and isopleth maps and some dendrographic schemes (trees), which all use colors. Here we discuss only the cartographic details and the dialectometric status of the following three map types: similarity maps (see Plate 2, and Goebl 1984 I, 114), parameter maps (see Plate 3, and Goebl 1984 I, 136) and isogloss synthesis (see Plate 4, and Goebl 1984 I, 183, *et passim*). For the discussion of the remaining map types (beam maps, trees, correlation maps), see Goebl (1983; 1984, 172; 2005a).

7.7.4 Step D: Presentation and Discussion of Plate 2, Plate 3, and Plate 4

7.7.4.1 Similarity Map

From the statistical point of view, every similarity map relies on one of the N (=641) vectors of the similarity matrix (see Plate 2). There are consequently 641 similarity maps that may be generated and compared to one another. Of the 641 RIV_{jk} values, 640 have been visualized. The reflexive score $RIV_{307,307}$ (= 100) has been disregarded: the polygon in question belongs to the reference site (here, ALF P. 307), and it therefore remains blank.

The cartographic prerequisites are as follows. Cartographically speaking, similarity maps are quantitative choropleth maps, which allow us to visualize a quantitative pseudo-continuum. The inner numerical variation of the N vectors of the similarity matrix should be adequately represented by the correspondingly variable visualization.

The spatial repartition of the different graphic steps should be done by appropriate interval algorithms. For optimizing the visual recognition of spatial patterns, having free choice over the number of graphic intervals is indispensable. The base map should be polygonized according to Voronoi geometry (Okabe, Boots, and Sugihara 1992).

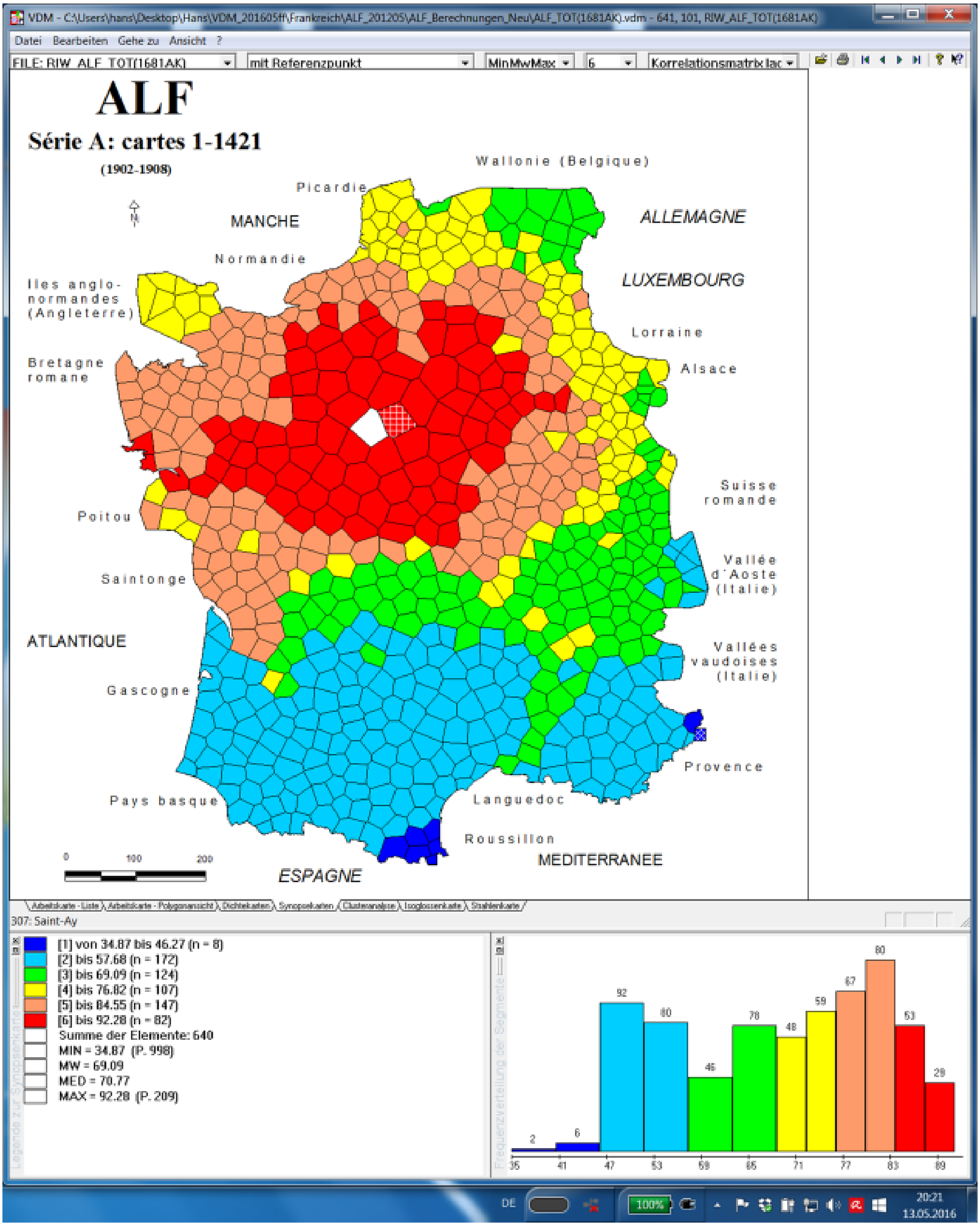

Plate 2 Sample of a similarity map: spatial distribution of the similarity values referring to ALF-point 307 (Saint-Ay, Département Loiret). Similarity index: $RIV_{307,k}$: corpus: 1681 working maps, all linguistic categories; algorithm of visualisation: MINMWMAX 6-tuple. Cartographic status: *quantitative* choropleth map. See Section 7.7.4, pp. 133–139. (*See insert for colour representation of the figure.*)

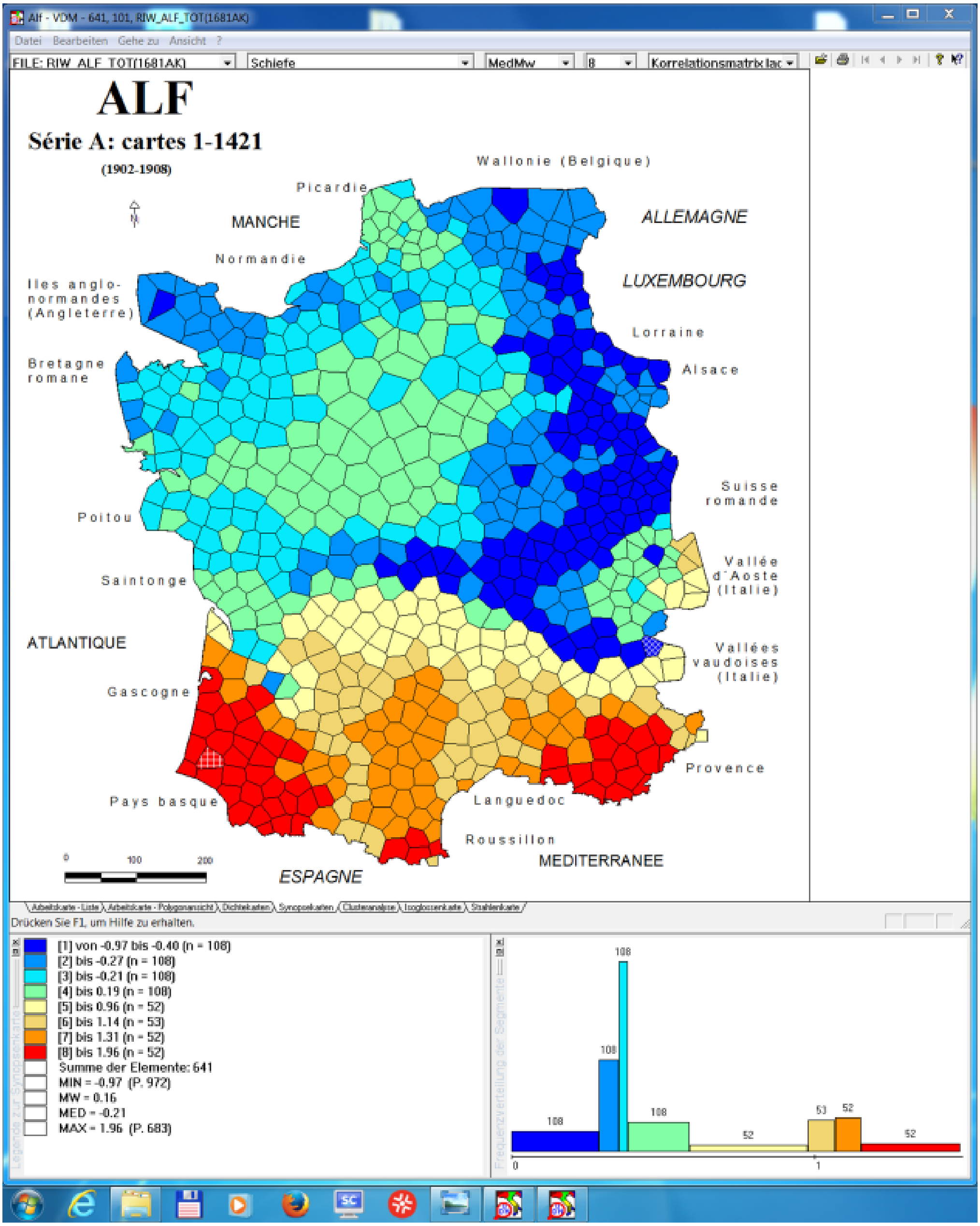

Plate 3 Sample of a parameter map: synopsis of 641 skewness values (according to the asymmetry index of R. A. Fisher). Similarity index: RIV_{jk}: corpus: 1681 working maps, all linguistic categories; algorithm of visualisation: MEDMW 8-tuple. Cartographic status: *quantitative* choropleth map. See Section 7.7.4, pp. 133–139. (*See insert for colour representation of the figure*.)

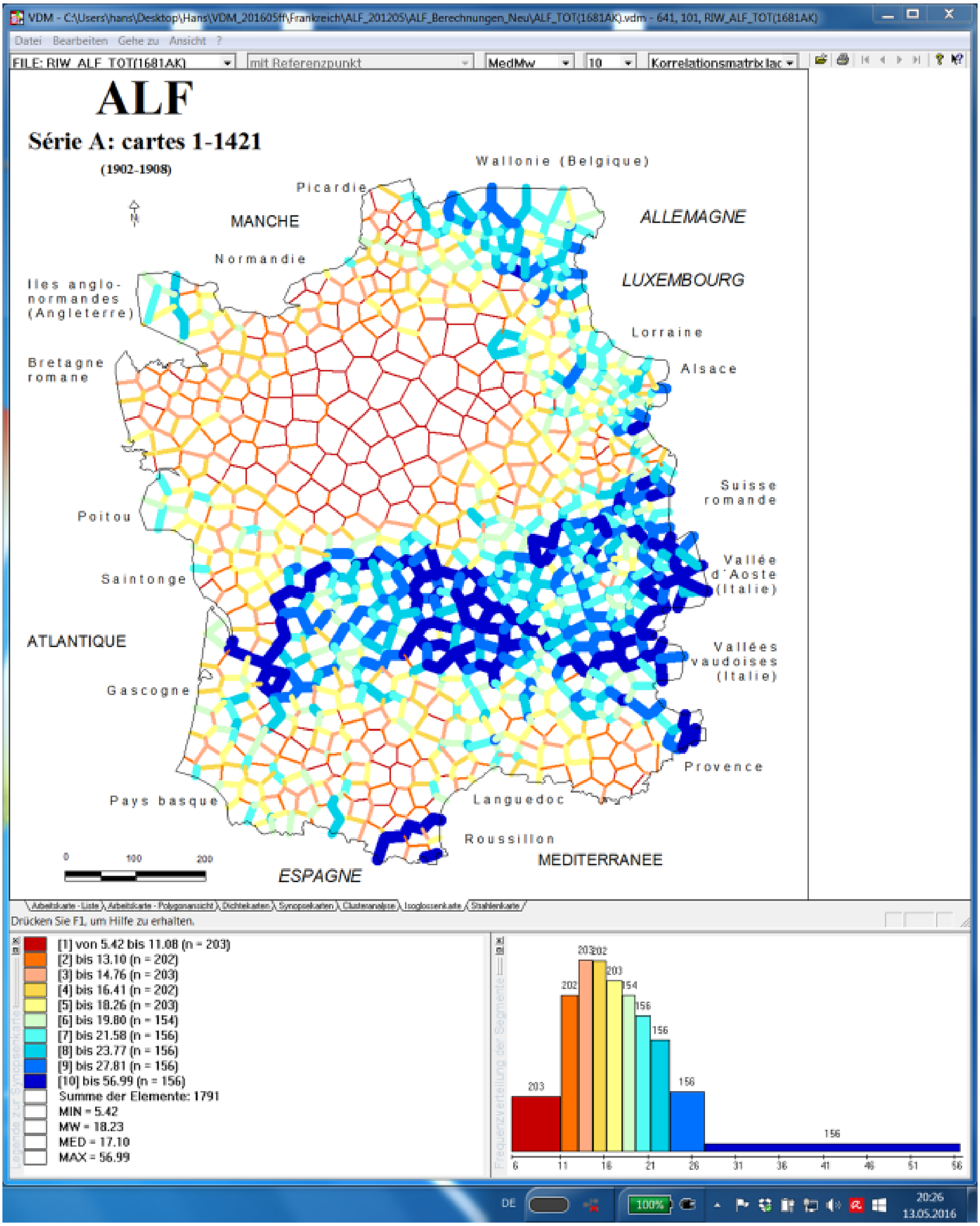

Plate 4 Sample of a interpoint map (honeycomb mode): synopsis of 1791 distance values (according to RDV_{jk}). Distance index: RDV_{jk}: corpus: 1681 working maps, all linguistic categories; algorithm of visualisation: MEDMW 10-tuple.Cartographic status: *quantitative* isarithmic (or: isopleth) map. See Section 7.7.4, pp. 133–139. (*See insert for colour representation of the figure*.)

For all these requirements the VDM program offers adequate solutions. The number of intervals is always even. The interval algorithms used in this article (MINMWMAX and MEDMW) always distribute N or N-1 dialectometric scores on the two sides of the arithmetic mean (cf. Goebl 1984 I, 93). MINMWMAX gives the intervals located above and below the arithmetic mean the same numerical width, whereas MEDMW gives the intervals located above and below the arithmetic mean the same size (i.e., number of polygons). MINMWMAX has been applied to Plate 2, MEDMW to Plate 3 and Plate 4.

In practice, the current use of visualization techniques depends deeply on visual training and experience, and belongs to what is nowadays called "imaging." It is always recommended that the user begin the contemplation of a choropleth map with the polygons of the highest-ranking intervals before proceeding to the polygons of lower-ranking intervals. The choropleth structure of the map shows a regular decrease in the dialectal similarities in space. The same phenomenon can be observed in all the remaining N-1 similarity maps of the same similarity matrix, as well as when using other similarity indexes.

Obviously, the regularity of the decrease in similarity from the reference site to the borders of the map is independent of the personal linguistic experience of the speakers themselves. This decrease must be explained with reference to the synergy of the forces of all language laws that determine the linguistic nature of the 641 locolects of the ALF grid.

The similarity maps have different (geo)linguistic meanings. To a Romance linguist, the stratification of the (warm and cold) colors on Plate 2 is characteristic: the sum of the zones in intervals 4–6 (warm colors) corresponds (with the exception of Wallonia in the North) to the influence area of the *Langue d'Oïl*, whereas the sum of the polygons in intervals 1–3 (cold colors) constitute the catchment area of the *Langue d'Oc* (or *Occitanian*). Moreover, a similarity map indicates the (relational) position of a given locolect in the middle of a given grid.

Note that these analyses can be repeated for different linguistic categories. In this case, the shapes of the various similarity profiles can vary somewhat. However, these differences are mostly rather slight (see Goebl, 2002, 2003, 2006a, 2010, and 2013a).

We have also observed that the overall structures of all our dialectometric visualizations remain practically unaltered once we have amassed a set of 200–300 randomly selected WMs. The same is true by using WM corpora with *little*, *middle* or *big* polynomy (see Figure 7.2). Similarity maps can also be interpreted from a *diffusionist* perspective, as they thereby show—metaphorically speaking—the results of diffusional effects of one element of the whole group.

Another metaphorical interpretation would be a *communicative* one. The algebraic logic of RIV_{jk} corresponds to the network of connections of telephone calls. Because any single co-identity (COI) would represent a single telephone contact and any co-difference (COD) the opposite (i.e., no contact), it follows that every similarity map relying on RIV_{jk} can be interpreted as analogous to the telephone activity of a given subscriber to a telephony service. From this perspective (see Plate 2), one can see that the intensity of the 'telephone connections' of ALF point 307 is rather circular, and decreases rapidly to the North (Wallonia), the South, and the South-East of the grid.

7.7.4.2 Parameter Map

The legend and histogram in Plate 2 give all the information necessary for the statistical understanding of the respective similarity distribution (minimum 34.87; arithmetic mean 69.09; maximum 92.28, etc.; see Plate 3). The shape of the histogram shows a rather symmetrical frequency distribution with two modes. Of course, for the remaining 640 similarity distributions all of these values may be different.

A systematic check of these numerical variations in fact shows that it is possible to interpret them linguistically. In this regard, the study of the symmetry of the similarity distributions is of particular interest. The legend of Plate 2 shows that 336 (= 107 + 147 + 82, or 52.5%) of 640 RIV values lie above the arithmetic mean (69.09). Again using a communicative metaphor, one could argue that the "dialecticity" of ALF P. 307 is related well to the rest of the grid, whereas the contrary would be true if, say, 60% of the RIV values lay below the arithmetic mean. In this respect, one of the most suitable indices we can use to grasp the symmetry of a frequency distribution is its skewness (for the formula, see Goebl 1984 I, 150).

The construction of Plate 3 depends upon the following procedure: (a) computation of the 641 skewness values, (b) mapping of these values, and (c) linguistic evaluation of the new choropleth pattern. What is immediately striking about Plate 3 is its self-explanatory structure: in the North, East, and West the polygons in intervals 1 and 2 (dark and middle blue) form two circular, or pincer-shaped, patterns with a very clear spatial distribution. The large "circle" surrounds the whole *Domaine d'Oïl*, whereas the domain of *Franco-Provençal* (on the South-eastern periphery) is located between the two "jaws" of the pincer, which has its pivot to the West of Lyon. In the South, we have three "bulwarks" (Gascony, Languedoc, and Provence, all in red (interval 6), which are linked together by polygons in intervals 4 and 5. This geographical structure, together with our knowledge of the linguistic evolution of France and some statistical ideas, allows us to confer the following linguistic meaning on the different hatchings and shadings of the map.

- *Polygons in interval 1 or 2 (dark and middle blue)*: these are zones of great "linguistic compromise." That is, they are dialecticities with a high percentage of many large-sized ("mega-choric") linguistic attributes (and their areas);
- *Polygons in interval 6 (red)*: these represent zones of minimal linguistic compromise. They are, in other words, dialecticities with a high percentage of small- ("oligo-choric") and medium-sized ("meso-choric") linguistic features (and their areas).

Note that by the term "linguistic compromise" (German *Sprachausgleich*) we mean a specific ratio of the entanglement of oligo-, meso- and mega-choric areas as a result of local or regional language contact and conflict.

The two circular configurations in intervals 1 and 2 are, on the one hand (referring to the *Domaine d'Oïl*), the visible result of the secular expansion of the linguistic type of the *Langue d'Oïl*, and on the other hand (in relation to *Franco-Provençal*), the visible result of the retreat of the old Latin linguistic heritage of Lugdunum/Lyon, with all its concomitant linguistic consequences, pressed as it was on two sides by the *Langue d'Oïl* in the North and by the *Langue d'Oc* in the South. By contrast, "aggregates" in interval 6 (red) in the South are real zones of linguistic isolation, whereas between them there are some weak flows of linguistic compromise (see the polygons of intervals 4–5).

The choropleth structure of Plate 3—which very clearly indicates dynamic irradiation in the North and a punctual (block-like) resistance in the South—is of great importance for the history of the whole Gallo-Romance domain. Similar structures can be found in all our dialectometric analyses.

7.7.4.3 Interpoint Map

Cartographically speaking, Plate 2 and Plate 3 were based on *areas* (polygons). Their iconic message depends upon the visual interplay of continuous surfaces, whereas the oldest cartographic schemes in linguistic geography, viz. the drawing of isogloss bundles, is based on the combining of *lines* crossing the space discontinuously. The dialectometric modeling of the combined drawing of isoglosses is quite straightforward. Cartographically, the drawing of isoglosses follows the edges of the polygons in the Voronoi base map (see Plate 4). The drawing of isogloss bundles is done by changing the thickness and coloring of the polygon sides. From the statistical point of view, the values to be visualized are *distances*, rather than *similarities* ($RDV_{jk} = 100 - RIV_{jk}$). They can be extracted from the distance matrix according to the neighborship geometry of the Voronoi base map.[13]

However, the line-based isopleth map presents new visual challenges, because the eye has to capture a discontinuous image syntax. In Plate 4 we applied a visualization using ten graphic steps, the interval algorithm MEDMW, and a medium-sized thickness span for the polygon edges. The number of visualized distance values, and therefore polygon edges, is 1,791, a number which corresponds to 8.73% of the total content (205,120 RDV values) of the distance matrix. The taxometrical impact of this visualization is thus rather small. The line-based pattern of Plate 4 is nevertheless highly suggestive.

With regard to the spatial compactness of thick and dark blue polygon edges, it appears rather clearly that there are four hotspots on the map. One is in the South, between the Roussillon and the Languedoc. Another is in the middle of the map, between the Domaines d'Oc and d'Oïl (note the curvature). A third is located at the Eastern border of the map, and the fourth on the Northern border of the Franco-Provençal area, between the Northern domains of Picardian and Wallonian. By contrast, there are zones of little compartmentalization in the central parts of the Domaines d'Oïl and d'Oc (i.e., Languedoc and Provence).

A visual comparison with older French isogloss syntheses (see Rosenqvist 1919 and Ettmayer 1924) reveals a striking resemblance between those pioneering maps (realized, of course, via manual work) and the interpunctual message of Plate 4.

7.8 Final Remarks

The dialectometrical procedures introduced in the present chapter have an exclusively diagnostic and exploratory character, and represent a hybrid compound with elements taken from linguistics, statistics, and cartography. The concurrence of these methods is intended to enhance our knowledge of the structure and function of geolinguistic networks using quantitative means.

Dialectometry, as it has been documented here, is defined programmatically as the *quantitative* branch of classical, atlas-borne and essentially *quality*-oriented linguistic geography. It could be demonstrated, with dialectometrical support, that there exist hitherto hidden and unexpectedly complex spatial patterns, which would imply the existence of genuine "spatial laws" in linguistic atlas data. These patterns exemplify the postulate expressed at the beginning of the chapter concerning the "basilectal management of space by *Homo loquens*," which is one of the numerous semiotic behaviors of our species.

NOTES

1 This is exactly what we call "area" later on.
2 See Jaberg (1908), who gives an excellent introduction in this kind of analysis.
3 The DSA (*Deutscher Sprachatlas*) is representative in this regard. Its author, Georg Wenker, himself converted a good proportion of his data into maps, which later became canonical.
4 See the historical overview in Winkelmann (2001).
5 One should not overlook the intellectual support given to Séguy by his friend and colleague Henri Guiter (1909–1994); see Guiter (1973).
6 See Malkiel (1967) and Christmann (1971).
7 See Sneath and Sokal (1973), Bock (1974), and Chandon and Pinson (1981).

8 See my contributions on dialectometry published between 1981 and 2013. I refer the reader also to the web-based bibliography of my DM writings at https://www.sbg.ac.at/rom/people/prof/goebl/dm_publi.htm (accessed 7 March 2017).

9 Nowadays there are two further dialectometrical research centers worth mentioning, these being the Groningen school (see the contributions of Nerbonne, Heeringa, Prokić, and Shackleton quoted below, and Chapters 20 and 23 of this volume), and the Athens (Georgia) school (see the work of Kretzschmar and Schneider, as well as Chapter 3 of this volume). Szmrecsanyi (2013) and Chapter 18 of this volume are also relevant. Note that in Groningen and Athens the basic assumptions of dialectometry differ from those of Regensburg-Salzburg, mainly by omitting the concept of area and stressing the method of sequence comparison by means of Levenshtein distance.

10 See our DM contributions related to the following domains: France (1984, 2002, 2003, 2004, 2005b, 2006a, 2007, 2010, 2013a), Italy (1981, 1983, 1984), Iberia (2013b), Catalonia (2013c), England (2007, and Goebl and Schiltz, 1997), and German-speaking Switzerland (Goebl, Scherrer, and Smečka, 2013).

11 See Jaberg's (1908) pioneering booklet, and the major synthesis of Brun-Trigaud, Le Berre, and Le Dû (2005), which represents the sum of the Gallo-Romance *aréologie* based on ALF.

12 See http://www.dialectometry.com/dmdocs/index.html (accessed 7 March 2017).

13 Note that the geometric principle of dialectal neighborhood ("interpunctual contiguity") had already been defined by Carl Haag in 1898.

REFERENCES

Altmann, Gabriel. 1985. "Die Entstehung diatopischer Varianten: Ein stochastisches Modell." *Zeitschrift für Sprachwissenschaft*, 4: 139–155.

Bock, Hans Hermann. 1974. *Automatische Klassifikation: Theoretische und praktische Methoden zur Gruppierung und Strukturierung von Daten (Cluster-Analyse)*. Göttingen: Vandenhoeck and Ruprecht.

Brun-Trigaud, Guylaine, Yves Le Berre, and Jean Le Dû. 2005. *Du temps dans l'espace: Explication linguistique de la France (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 Leroux)*. Paris: Comité des Travaux Historiques et Scientifiques (CTHS).

Chandon, Jean-Louis, and Suzanne Pinson. 1981. *Analyse typologique: Théories et applications*. Paris: Masson.

Christmann, Hans Helmut. 1971. "Lautgesetze und Wortgeschichte: Zu dem Satz *Jedes Wort hat seine eigene Geschichte*." In *Sprache und Geschichte: Festschrift für Harri Meier zum 65. Geburtstag*, edited by Wolf-Dieter Stempel, and Eugenio Coseriu, 111–124. Munich: Fink.

Dickinson, Gordon. 1973. *Statistical Mapping and the Presentation of Statistics*. London: Arnold.

DSA. 1927–1956. *Deutscher Sprachatlas, aufgrund des von Georg Wenker begründeten Sprachatlas des Deutschen Reiches in vereinfachter Form begonnen von Ferdinand Wrede, fortgesetzt von Walther Mitzka und Bernhard Martin*. Marburg/Lahn: Elwert, 23 fascicles with 128 maps.

Ettmayer, Karl von. 1924. *Über das Wesen der Dialektbildung erläutert an den Dialekten Frankreichs*. Vienna: Denkschriften der Akademie der Wissenschaften in Wien.

Gilliéron, Jules. 1881. *Petit atlas phonétique du Valais Roman (Sud du Rhône)*. Paris: Champion.

Gilliéron, Jules. 1918. *Généalogie des mots qui désignent l'abeille d'après l'Atlas Linguistique de la France*. Paris: Champion.

Gilliéron, Jules, and Edmond Edmont, eds. 1902–1910 [1968]. *Atlas linguistique de la France* (10 vols.). Paris: Champion. [ALF]

Goebl, Hans. 1981. "Éléments d'analyse dialectométrique (avec application à l'AIS)." *Revue de linguistique Romane*, 45: 349–420.

Goebl, Hans. 1983. "Parquet polygonal et treillis triangulaire: Les deux versants de la dialectométrie interponctuelle." *Revue de linguistique Romane*, 47: 353–412.

Goebl, Hans. 1984. *Dialektometrische Studien: Anhand Italoromanischer, Rätoromanischer und Galloromanischer Sprachmaterialien aus AIS und ALF* (3 vols.). Tübingen: Niemeyer.

Goebl, Hans. 2002. "Analyse dialectométrique des structures de profondeur de l'ALF." *Revue de linguistique Romane*, 66: 5–63.

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: 59–120.

Goebl, Hans. 2004. "Sprache, Sprecher und Raum: Eine kurze Darstellung der Dialektometrie – Das Fallbeispiel Frankreich." *Mitteilungen der österreichischen geographischen Gesellschaft*, 146: 247–86.

Goebl, Hans. 2005a. "La dialectométrie corrélative: Un nouvel outil pour l'étude de l'aménagement dialectal de l'espace par l'homme." *Revue de linguistique Romane*, 69: 321–67.

Goebl, Hans. 2005b. "Dialekte und Familiennamen in Frankreich: Ein interdisziplinärer Vergleich mit den Mitteln der Dialektometrie." In *Gene, Sprachen und ihre Evolution: Wie verwandt sind die Menschen – Wie verwandt sind ihre Sprachen?*, edited by Günter Hauska, 68–99. Regensburg: Universitätsverlag Regensburg.

Goebl, Hans. 2006a. "Recent Advances in Salzburg Dialectometry." *Digital Scholarship in the Humanities (*formerly *Literary and Linguistic Computing*), 21(4): 411–35.

Goebl, Hans. 2006b. "Warum die Dialektometrie nur in einem roman(ist)ischen Forschungskontext entstehen konnte." In *Was kann eine vergleichende romanische Sprachwissenschaft heute (noch) leisten? Romanistisches Kolloquium XX*, edited by Wolfgang Dahmen, Günter Holtus, Johannes Kramer, Michael Metzeltin, Wolfgang Schweickard, and Otto Winkelmann, 291–317. Tübingen:Narr.

Goebl, Hans. 2007. "A bunch of dialectometric flowers: A brief introduction to dialectometry." In *Tracing English through Time – Explorations in Language Variation: In Honour of Herbert Schendl on the Occasion of his 65th birthday*, edited by Ute Smit, Stefan Dollinger, Julia Huettner, Gunther Kaltenböck, and Ursula Lutzky, 133–71. Wien: Braumüller.

Goebl, Hans. 2010. "Dialectometry: Theoretical prerequisites, practical problems, and concrete applications (mainly with examples drawn from the 'Atlas linguistique de la France', 1902–1910)." *Dialectologia,* Special Issue 1, 63–77.

Goebl, Hans. 2013a. "Dialectometry and quantitative mapping." In *Language and Space: An International Handbook of Linguistic Variation, Vol. 2. – Language Mapping*, edited by Alfred Lameli, Roland Kehrein, and Stefan Rabanus, 433–57. Berlin: de Gruyter.

Goebl, Hans. 2013b. "La dialectometrización del ALPI: Rápida presentación de los resultados." In *Actas del XXVI Congreso Internacional de Lingüística y de Filología Románicas*, vol. 6. (Valencia 2010), edited by Emili Casanova Herrero, and Cesáreo Calvo Rigual, 143–54. Berlin: de Gruyter.

Goebl, Hans. 2013c. "La dialectometrització dels quatre primers volums de l'ALDC." *Estudis romànics*, 35: 87–116.

Goebl, Hans. 2013d. "Le Baiser de la Belle au bois dormant où: des péripéties encourues par la géographie linguistique depuis Jules Gilliéron." In *Dialectologie: corpus, atlas, analyses*, edited by Rita Caprini, 61–84. Alessandria: Edizioni dell'Orso [= *Corpus* 12, 2013].

Goebl, Hans, Yves Scherrer, and Pavel Smečka. 2013. "Kurzbericht über die Dialektometrisierung des Gesamtnetzes des 'Sprachatlasses der deutschen Schweiz' (SDS)." In *Vielfalt, Variation und Stellung der Deutschen Sprache*, edited by Karina Schneider-Wiejowski, Birte Kellermeier-Rehbein, and Jakob Haselhuber, 153–75. Berlin: de Gruyter.

Goebl, Hans, and Guillaume Schiltz, 1997. "A dialectometrical compilation of CLAE I and CLAE II: Isoglosses and dialect integration." In *Computer Developed Linguistic Atlas of England (CLAE II)*, edited by Wolfgang Viereck, and Heinrich Ramisch, 13–21. Tübingen: Niemeyer.

Griera, Antoni. 1923–1964. *Atlas Lingüístic de Catalunya* (8 vols.). Barcelona: Edicions Polígrafa. [ALC]

Guiter, Henri. 1973. "Atlas et frontières linguistiques." In *Les dialectes romans de France à la lumière des atlas régionaux*, edited by Georges Straka, and Pierre Gardette, 61–109. Paris: Centre National de la Recherche Scientifique.

Haag, Carl. 1898. *Die Mundarten des Oberen Neckar- und Donaulandes (Schwäbisch-alemannisches Grenzgebiet: Baarmundarten).* Reutlingen: Hutzler.

Haggett, Peter. 1965. *Locational Analysis in Human Geography*. London: Arnold.

Heeringa, Wilbert. 2004. *Measuring Dialect Pronunciation Differences using Levenshtein Distance*. Groningen: Groningen Dissertations in Linguistics, 46.

Heeringa, Wilbert, and John Nerbonne. 2001. "Dialect areas and dialect continua." *Language Variation and Change*, 13: 375–400.

Jaberg, Karl. 1908. *Sprachgeographie: Beitrag zum Verständnis des Atlas linguistique de la France*. Aarau: Sauerländer.

Jaberg, Karl, and Jakob Jud, eds. 1928–1940 [1971]. *Sprach- und Sachatlas Italiens und der Südschweiz* (8 vols.). Zofingen: Ringier. [AIS]

Kretzschmar, William, and Edgar Schneider, eds. 1996. *Introduction to Quantitative Analysis of Linguistic Survey Data: An Atlas by Numbers*. Thousand Oaks, CA: Sage.

Malkiel, Yakov. 1967. "Each word has a history of its own." *Glossa*, 1: 13–149.

Nerbonne, John, and William Kretzschmar. 2003. "Introducing computational techniques in dialectometry." *Computers and the Humanities*, 37: 245–55.

Nerbonne, John, and William Kretzschmar. 2006. "Progress in dialectometry: Toward explanation." *Digital Scholarship in the Humanities (*formerly *Literary and Linguistic Computing*), 21(4): 387–97.

Nerbonne, John, and William Kretzschmar, eds. 2013. "Special Issue 'Dialectometry++'." *Digital Scholarship in the Humanities (*formerly *Literary and Linguistic Computing*), 28(1).

Okabe, Atsuyuki, Barry Boots, and Kokichi Sugihara. 1992. *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*. Chichester: Wiley.

Palsky, Gilles. 1996. *Des chiffres et des cartes: Naissance et développement de la cartographie quantitative française au XIX*e *siècle*. Paris: Comité des travaux historiques et scientifiques (CTHS).

Paris, Gaston. 1888. "Les parlers de France [1888]." In *Mélanges linguistiques*, edited by Gaston Paris, 432–48. Paris: Champion 1909.

Pop, Sever. 1950. *La dialectologie: Aperçu historique et méthodes d'enquêtes linguistiques* (2 vols.). Gembloux: Duculot.

Pop, Sever, and Rodinca Doina Pop. 1959. *Jules Gilliéron: Vie, enseignement, élèves, œuvres, souvenirs*. Louvain: Centre International de Dialectologie Générale.

Prokić, Jelena. 2010. *Families and Resemblances*. Groningen: Groningen Dissertations in Linguistics, 88.

Puşcariu, Sextil, Sever Pop, and Emil Petrovici, eds. 1938–1940. *Atlasul lingvistic român* (4 vols.). Cluj: Muzeul Limbii Române. [ALR]

Rosenqvist, Arvid. 1919. "Limites administratives et division dialectale de la France." *Neuphilologische Mitteilungen*, 20: 87–119.

Séguy, Jean, ed. 1954–1974. *Atlas linguistique et ethnographique de la Gascogne* (6 vols.). Paris: Centre National de la Recherche Scientifique. [ALG]

Séguy, Jean. 1971. "La relation entre la distance spatiale et la distance lexicale." *Revue de Linguistique Romane*, 35: 335–57.

Séguy, Jean. 1973. "La dialectométrie dans l'Atlas linguistique de la Gascogne." *Revue de Linguistique Romane*, 37: 1–24.

Shackleton, Robert. 2010. *Quantitative Assessment of English-American Speech Relationships*. Groningen: Groningen Dissertations in Linguistics, 81.

Sneath, Peter, and Robert Sokal. 1973. *Numerical Taxonomy: The Principles and Practice of Numerical Classification*. San Francisco, CA: Freeman.

Szmrecsanyi, Benedikt. 2013. *Grammatical Variation in British English Dialects: A Study in Corpus-Based Dialectometry*. Cambridge: Cambridge University Press.

Viereck, Wolfgang, and Heinrich Ramisch, eds. 1991–1997. *Computer Developed Linguistic Atlas of England* (2 vols.). Tübingen: Niemeyer. [CLAE]

Winkelmann, Otto. 2001. "Romanische Sprachatlanten." In *Lexikon der romanistischen Linguistik* (vol. 1/2), edited by Günter Holtus, Michael Metzeltin, and Christian Schmitt, 1004–68. Tübingen: Niemeyer.