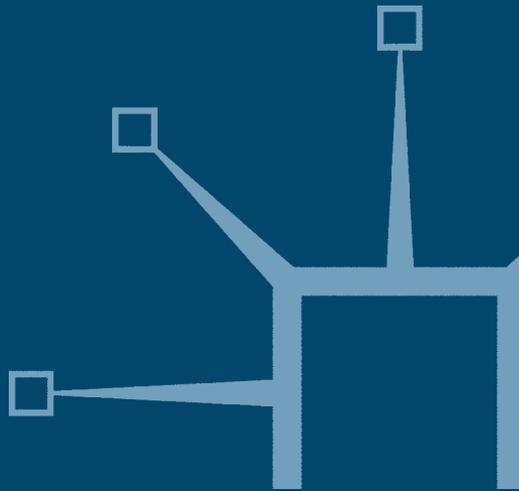


palgrave
macmillan

Working in Language and Law

A German Perspective

Hannes Kniffka



Working in Language and Law

Also by Hannes Kniffka

Elements of Culture-Contrastive Linguistics

Recent Developments in Forensic Linguistics

Soziolinguistik und empirische Textanalyse

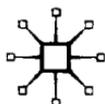
Texte zu Theorie und Praxis forensischer Linguistik

Working in Language and Law

A German Perspective

Hannes Kniffka
Bonn University

palgrave
macmillan



© Hannes Kniffka 2007
Foreword © Roger W. Shuy 2007

All rights reserved. No reproduction, copy or transmission of this publication may be made without written permission.

No paragraph of this publication may be reproduced, copied or transmitted save with written permission or in accordance with the provisions of the Copyright, Designs and Patents Act 1988, or under the terms of any licence permitting limited copying issued by the Copyright Licensing Agency, 90 Tottenham Court Road, London W1T 4LP.

Any person who does any unauthorised act in relation to this publication may be liable to criminal prosecution and civil claims for damages.

The author has asserted his right to be identified as the author of this work in accordance with the Copyright, Designs and Patents Act 1988.

First published in 2007 by
PALGRAVE MACMILLAN
Houndmills, Basingstoke, Hampshire RG21 6XS and
175 Fifth Avenue, New York, N.Y. 10010
Companies and representatives throughout the world.

PALGRAVE MACMILLAN is the global academic imprint of the Palgrave Macmillan division of St. Martin's Press, LLC and of Palgrave Macmillan Ltd. Macmillan® is a registered trademark in the United States, United Kingdom and other countries. Palgrave is a registered trademark in the European Union and other countries.

ISBN-13: 978-0230-55142-8 hardback
ISBN-10: 0-230-55142-4 hardback

This book is printed on paper suitable for recycling and made from fully managed and sustained forest sources. Logging, pulping and manufacturing processes are expected to conform to the environmental regulations of the country of origin.

A catalogue record for this book is available from the British Library.

Library of Congress Cataloging-in-Publication Data

Kniffka, Hannes.

Working in language and law : a German perspective / Hannes Kniffka ;
foreword by Roger W. Shuy.

p. cm.—

Includes bibliographical references and index.

ISBN 0-230-55142-4 (alk. paper)

1. Forensic linguistics – Germany. 2. Writing – Identification – Germany.

I. Title.

KK3938.K58 2007
363.25—dc22

2007022329

10 9 8 7 6 5 4 3 2 1
16 15 14 13 12 11 10 09 08 07

Printed and bound in Great Britain by
Antony Rowe Ltd, Chippenham and Eastbourne

Contents

<i>List of Tables and Figures</i>	ix
<i>Foreword by Roger W. Shuy</i>	xii
<i>Acknowledgements</i>	xv
Introduction	1
Linguistic expert testimony <i>in</i> or <i>pro foro</i> in Germany	4
Range of variation of linguistic expert testimony in the German legal context	9
Overview of chapters and notes on the genealogy of the texts	14
Part I The Interdisciplinary Status of Forensic Linguistics	
1 Forensic Linguistics: Its Relatives and Neighbors – An Interdisciplinary Perspective	27
Introduction	27
Interdisciplinarity: Perspectives and postulates	30
Structural similarities and differences between forensic phonetics and forensic linguistics	38
Part II Forensic Linguistic Research in Germany in the Past	
2 The Linguist as an Expert Witness in German Courts. A View from the 1970s	51
Orientation	51
Applied linguistics and theoretical linguistics	52
Interdisciplinary status of “forensic linguistic” work and the overall aim of linguistic expert opinions <i>in/pro foro</i>	53
Case A: Analyzing the sense of a text passage	54
Case B: A press correction or a repetition of the original statement?	57
Case C: Disputed authorship of an article in the press	58

Case D: Disputed authenticity of the minutes of a works council meeting	59
Case E: Fake edition of a company paper	60
Some methodological characteristics of forensic linguistic expert opinions on authorship attribution (AA)	62
Case I	63
Case II	63
3 Status and Tasks of Forensic Linguistic	
Authorship Analysis	71
Introduction	71
Thesis 1: The notion of “idiolect” in linguistics	72
Thesis 2: The term “forensic linguistics”	73
Thesis 3: Forensic linguistics as an auxiliary science	73
Thesis 4: “Features” as a linguistic term	74
Thesis 5: Methodology of authorship analysis	76
Thesis 6: The need for theoretical reflection	77
Thesis 7: Forensic linguistics as a branch of linguistics	78
Thesis 8: Sub-disciplines of forensic linguistics	79
Thesis 9: Qualitative and quantitative analysis	79
Thesis 10: FL anonymous authorship analysis within a sociolinguistic framework	79
Thesis 11: Dynamic data of speaker-specific behavior	81
Thesis 12: Linguistic “differential diagnosis”	82
Thesis 13: Interdisciplinary perspectives	83
Thesis 14: Documentation of research and casework	84
4 “Shibboleths” as Data of Linguistic Behavior	86
Introduction	86
Origin and history of the concept	89
Descriptive characteristics of shibboleths	90
Salient structural features of shibboleths	92
Working definition of shibboleths and domains of occurrence	101
Summary	107
Part III Ongoing Research in Forensic Linguistics in Germany	
5 Libel, Linguists, and Litigation in Germany	113
Orientation	113

German defamation laws	116
Basic statistics of insults and other offenses	118
Some differences between German defamation laws	121
An exemplaric libel case: <i>Konkubine</i> in today's Standard German (1972)	125
Points of argument	130
What judges and linguistic experts need to learn about one another	133
The status of linguistic expert testimony <i>in/pro foro</i>	139
Outlook and future tasks	141
Appendix I: "Praxiskommentar" (Herdegen 1975)	143
Appendix II: Original text of a forensic linguistic expert opinion (English translation)	144
6 A Heuristic Author and Writer/Typist Taxonomy	149
Introduction	149
Dimensions of variation	150
A heuristic taxonomy of features and variants of authorship and writership of incriminating texts	152
Relevance for applied linguistic and criminalistic classification of text types	162
7 The System and Diagnostic Potential of Orthographic Data in Forensic Linguistic Authorship Attribution	166
Introduction	166
Data of orthographic behavior and anonymous authorship analysis	167
Status and repertoires of data of written language behavior in anonymous authorship analysis	173
Exemplaric description of ambiguous orthographic data in FLAA	182
Summary	190
8 Orthographic Data in Forensic Linguistic Authorship Analysis	192
Introduction	192
Examples of orthographic data in anonymous letters	194
Continua of orthographic behavior and the methodological impact for a "idiosyncrasy coefficient"	208

Text type-specific distribution and constraints of linguistic and orthographic features	225
Summary	233

Part IV Outlook and Future Tasks

9 Language and Law: Some Needs	237
Introduction	237
Linguists and lawyers: misconceptions of each and of each other	240
Problems in forensic interaction: “tricks of the trade”	253
Intrinsic deficits of linguistic basic research and interdisciplinary projection	256
Summary	264
10 Outlook	265
<i>Notes</i>	272
<i>References</i>	274
<i>Index of Authors</i>	282
<i>Subject Index German</i>	284
<i>Subject Index English</i>	286

List of Tables and Figures

Tables

3.1a	Terminological variants for the description of the variation of forms	75
3.1b	Terminological variation for speaker-specific linguistic features	75
5.1	Defamation laws (excerpt) in the German Penal Code	116
5.2	Share represented by attempts with regard to individual offenses (by offense group)	119
5.3	Insults. Total number of offenses recorded for 1987–2003	121
5.4	Libel letters of Mr. A, June–August 1972	125
5.5a	Questionnaire 2a	128
5.5b	Questionnaire 2b	128
5.6	Distinctive contexts for the word <i>Konkubine</i>	131
6.1	Author and writer constellations from a layman’s perspective	152
6.2	Heuristic taxonomy of text products	154
6.3	Distinctive text-profile matrices	160
7.1	Domains of data of forensic text analysis	168
7.2	Deviations in orthography and language	177
7.3	Three heuristic classes of data of orthographic behavior in German	179
7.4	Features of orthographic behavior: excessive and (simultaneous) repeated use of graphemic signs (mainly punctuation)	181
7.5	Styles of orthographic behavior	182
8.1	Classes and types of the “Fugen-s” in German	199
8.2	“Fugen-s” in identical words in the three German-speaking countries	200
8.3a	Spelling of a well-known emblem in German	207
8.3b	Concomitant exceptional spellings	207
8.4	Descriptive and explanatory potential of orthographic data	212
8.5	Different spelling variants of a stereotyped expression	214

8.6	Spacing with parenthesis dashes ± empty space (ES) before and after parenthesis dashes	216
8.7	Low idiosyncrasy coefficients of OD in a written text of everyday communication	220
8.8	Omission marks (three dots) at the end of a quotation at the end of a sentence at the end of a text-paragraph	221
8.9	Results of spelling experiment with German students (aged 21–28)	224
8.10a	Text type-specific contrast(s) of <i>grammatical</i> , <i>textual</i> and <i>semanto-pragmatic</i> features in anonymous extortion and blackmail letters versus defamation and libel letters	226
8.10b	Text type-specific contrast(s) of <i>graphemic</i> , <i>orthographic</i> and <i>layout</i> features in anonymous extortion and blackmail letters versus defamation and libel letters	227
8.11	Textual parameters and measuring instances	233
9.1	Heuristic taxonomy of needs of cooperations in the field of language and law	239
9.2	Lawyers' assessment of language proficiency of a cleaning lady (original wording of statement sent to the court)	249
9.3	Official court transcript of taped notes of witness interrogation on command of language	251
9.4	Fuzzy norms: Data on hyphenization in German in the word <i>5-Jahresvertrag</i>	260
9.5	Early occurrences of the word <i>lebensälter</i>	263

Figures

1.1	The field of linguistics	28
1.2	Forensic linguistics and neighboring sciences	29
5.1	English translation equivalents of German insult and defamation offenses	118
5.2a	Section 185, 'insult' (§185 StGB "Beleidigung")	122
5.2b	Section 186, 'malicious gossip' (§186 StGB "Üble Nachrede")	122
5.2c	Section 187, 'defamation' (§187 StGB "Verleumdung")	123
5.3	Areas of forensic linguistic expert testimony	140

6.1	Hierarchical affiliation of features of authorship and writership	157
8.1	Anonymous threatening letter	195
8.2	Example of an exceptional anonymous extortion letter	202

Foreword

Forensic linguistics, accurately described by Hannes Kniffka as a branch of applied linguistics, is developing rapidly in many countries of the world. Although the International Association of Forensic Linguistics was founded over a decade ago, its meetings and publications are all conducted in English. This not only tends to limit accessibility to non-English speakers, but also makes it difficult for these same English speakers to learn about the developments in forensic linguistics in different legal systems, cultures, and contexts. It is always difficult to keep track of simultaneous developmental stages at the international level. Therefore, not surprisingly, the description of the overall growth and development of forensic linguistics around the world has been considerably less than comprehensive. Although this book makes no claim to be a history of forensic linguistics in Germany, it opens an important door to those unfamiliar with the German language and context, and it helps them compare and learn from the theory and practice of the earlier and more recent work by Kniffka, who founded and has practiced forensic linguistics in Germany since the 1970s.

Most legal systems are fairly comparable but, often, even their small differences can cause confusion to practitioners in different legal jurisdictions. For example, defamation normally is considered a civil tort in some countries but defamation is listed in the German penal code, it apparently can be treated as a crime in Germany. The way that expert witnesses are employed in the USA is very different from the way they work in Germany, where they are often requested and called upon by the Court rather than by the opposing sides; a process that, if adopted in the USA, could alleviate the common complaint that experts are 'hired guns' for whomsoever employs them. While, in US civil cases, depositions of expert witnesses are usually taken and then made available to both sides before trial, this practice is largely unknown in much of Europe. In criminal cases, the techniques and strategies of law enforcement can lead to different uses of linguistic analysis by lawyers. In the USA, for example, there are thousands of cases in which undercover audio and videotape recordings are made by police officers and cooperating witnesses in their effort to capture criminal evidence while it actually takes place. This law enforcement practice appears to be far less common elsewhere in the world. Even the simple expression 'giving testimony' does

not mean the same thing in different countries. In the USA, it conventionally means giving oral testimony at trial, whereas in many other countries, including Germany, it means presenting a written report ordered by a judge or lawyer. These, and other differences in the ways that legal systems and law enforcement view and use forensic linguists, must be taken into consideration in order for forensic linguists to communicate effectively across language and legal system boundaries. The more we learn about each other, the better our mutual understanding and communication will be. *Working in Language and Law: A German Perspective* contributes substantially to that goal.

This book reflects the way forensic linguists are employed in Germany which, as the author points out, is largely to identify authors of documents of anonymous or unknown authorship, to analyze language that will help determine whether defamation has occurred, and to help resolve trademark infringement cases. Over the years, the author has provided his services in a multitude of such cases, yielding a wealth of experience and knowledge to share with linguists in English speaking countries. Needless to say, this is very valuable and much needed information.

Truly expert forensic linguists, as Kniffka and others have said repeatedly, are trained in all aspects of linguistics, including phonetics, morphology, syntax, semantics, pragmatics, discourse analysis, text linguistics, orthography, sociolinguistics and language change. Without competence in these basic tools of linguistics, there is no good reason to identify such witnesses as linguistic experts at all; simply being able to speak a language in no way qualifies a person to analyze it properly. However obvious that may seem, this book reports cases in which experts with little or no background in linguistics were asked to offer opinions about language use. As the author points out, the very fact that this happens tells us that linguists throughout the world have not done an adequate job of making the basic content and value of our field known to lawyers, judges and the general public. Books such as this one offer important steps toward addressing this problem.

Because Kniffka has worked on so many authorship identification cases, this topic emerges throughout his book. Uniquely, he tries to relate his own linguistic analyses of the written material in evidence to the forensic phonetic analyses and handwriting/typing analyses produced by other experts. Not infrequently, the cases described here involve written and spoken data to be analyzed by all three sister sciences. In relating and combining the results of each discipline with the others, the diagnostic potential of the evidence is increased and expanded. In this way,

Kniffka shows how forensic linguists must develop a heuristic typology of empirical research.

Forensic linguistics does not exist in a theoretical vacuum. It begins with sound linguistic theory, but it also requires the cooperation of the neighboring sciences, including sociology, criminology, psychology, law and statistics. This book welcomes and argues in favor of interdisciplinary cooperation as central to forensic linguistic work, all of which leads to the writer's valuable suggestions and models for future research and analysis of language data used in criminal and civil cases. Kniffka is also a strong advocate for mutual understanding and patience among representatives of all the disciplines that work in this area. More specifically, the book points out that linguists have succeeded in making our field's potential contributions known and understood by the legal profession and our sister fields. *Working in Language and Law: A German Perspective* contributes greatly to our finding solutions to this problem.

ROGER W. SHUY
Distinguished Research Professor of
Linguistics, Emeritus
Georgetown University

Acknowledgements

This book has benefited substantially from the interest, help and advice of colleagues and friends in linguistics and forensic linguistics throughout the world. In particular while at international conferences I had the privilege to consult with fellow forensic linguists, who generously let me share their insights and experience in the field. They are so numerous that it is impossible to name them all here; I wish to express my sincere thanks to them all.

I would like to especially thank Roger Shuy, who has not only been a close friend and colleague in forensic linguistics for several decades, but who also kindly read the chapters of this book with the greatest patience and kindness, and shared with me his wisdom in the field on many occasions. I also thank him for helping me iron out my stylistic and grammatical shortcomings.

I am also extremely grateful to Peter French, who originated the first international conferences on forensic phonetics and forensic linguistics in 1990. He gave these conferences an intellectual, professional and personal note that delighted many people, including me, and encouraged them to “join the club” (now called the IAFL).

It is also impossible to name all the people from the “real life” field of law, who inspired, criticized, fostered, enriched and corrected my knowledge in the area of (language and) law.

However, I would like to mention certain people by name who helped me in several areas: Sven Benson, Carole Chaski, Manfred Hecker, Georg Heike, Anita Karpa (BLKA), Wolfgang Löwer, Lothar Michel †, Gerhard Nickel, Francis Nolan, Christian Rohrer (BKA), Thomas Rübner, Larry Solan, Wolfgang Steinke (BKA), Peter Tiersma and Angelika Wöbken (BKA). They have been of great help to me in looking beyond the area of forensic linguistics proper.

Thanks also go to the publisher Palgrave Macmillan, in particular Jill Lake and Melanie Blair, and all others involved in the production process of this volume.

Last, but not least, I wish to express my heartfelt thanks to the people who helped produce the written manuscript of this book, in particular Frank Albrecht, Jan Clasen, Paul Stroh and Jürgen Walburg. Without their cooperation, patience and diligence, this book would not have been produced. All remaining errors are my own.

This page intentionally left blank

Introduction

This volume has three aims:

- (1) It pinpoints some “key questions” and “central issues” of forensic linguistics (henceforth FL), which have been addressed in articles by the author as early as the 1970s and continue to be addressed in forensic linguistic research on both sides of the Atlantic, in Australia, Asia and other parts of the world;
- (2) It documents in English translation articles that have to date only existed in German and hence did not attract adequate attention in the English-speaking forensic linguistic community;
- (3) It reports *ongoing* FL research in Germany, mainly in the areas of anonymous authorship attribution and offenses against (German) defamation laws.

This volume is not intended to be an account of the historical development of FL in Germany, let alone in other parts of the world. Incorporating a historical dimension (mainly in Part II, “Forensic Linguistic Research in Germany in the Past”), it shows the status of FL in the past and at the present time, and the rocky road still ahead of us. The picture gained from this seems to be somewhat clearer than a merely systematic account of the present-day activities in FL (addressed mainly in the Introduction, in Part I: “The Interdisciplinary Status of Forensic Linguistics”, Part II: “Forensic Linguistic Research in Germany in the Past”, and in Part III: “Ongoing Research in Forensic Linguistics in Germany”). It also seems that most of what is said in the early German articles (in Part II) is not at all out of date. Most issues discussed seem to have great relevance for the practical and theoretical concerns of the various fields of FL today.

This is, of course, equally true for other, older articles, such as Judith Levi's report and bibliography of 1982, and Roger Shuy's 1986(a) report on the field of language and law, to name just two.

Another motivation for this book was to give readers at least some clues and basic information on the general location and certain current issues in FL in Germany, thus enabling them to draw comparisons, in particular with the American and British systems. No one researcher, it seems, is equally at home with research on the American, British, and German legal systems, or able to work theoretically and practically in all three cultures. This is yet another indication that we should not ignore research done in languages other than English. There is plenty of evidence that research in FL from "remote" cultures and languages may be of critical importance to one's own research and practical work.

Due to globalization, including the globalization of crime, forensic linguists are bound to work together and have some kind of symmetrical exchange. Obtaining more knowledge and practical experience with FL issues in other countries and cultures is not only a pleasant by-product, but also a requirement of critical importance. The trouble of learning a foreign language and studying legal systems other than one's own is rewarded by practical experience and an extension of knowledge and methods that can be applied to legal and linguistic matters in one's own culture. So, it is strongly hoped that this report on a German (in fact, my own) perspective on FL is of interest to colleagues working in FL and other areas of language and law in English-speaking and other countries and cultures, both in a material sense of an enrichment of information, and also as procedural and methodological feedback for the work in their own language and culture.

The author of this book is in no way capable of giving a full-fledged account of the contrasts between the FL work and the legal system in Germany and, for example, that in the US. This is due to the fact that he knows far too little about the two systems and has a training only in linguistics, not in law, as would be preferable. However, the call for internationality and a more symmetrical exchange of results of the theoretical and practical work of FL is *conditio sine qua non* for the future. Let me illustrate this with one brief example: the author's report on giving forensic linguistic expert testimony in Germany in the 1970s appeared as Kniffka 1981 ("Der Linguist als Gutachter bei Gericht" ('The linguist as an expert witness in court') see Chapter 2). At least three American publications with very similar titles have appeared in the past without discussing or mentioning it in the references, let alone giving a more detailed elaboration on the similarities and the contrasts between the German and the American systems (Rieber and Stewart 1990, "The Language Scientist as Expert in the

Legal Setting”; Levi 1994, “Language as Evidence: The Linguist as Expert Witness in North American Courts”; Solan and Tiersma 2002, “The Linguist on the Witness Stand: Forensic Linguistics in American Courts”). I am stating this *sine ira et studio* and definitely without implying any criticism. I do want to point out explicitly, however, that ignoring research in languages other than one’s own seems, in general, a prevalent virtue of English-speaking colleagues, but that the reverse also holds for those of German tongue – with notable exceptions on both sides.

When I wrote the article on the linguist’s work in and for courts in 1981 (that is, in 1979 and 80) I had absolutely no knowledge of the literature on the subject in any language: there hardly was any. My knowledge of the FL work of others did not expand until 1988, when I started an intensive exchange of ideas with Roger Shuy, which has lasted for some 20 years to the present time. In addition to Roger Shuy, I had only one other personal contact, with Peter French since 1989, which also has lasted to the present day and dates back to a time when the International Association of Forensic Phonetics and the International Association of Forensic Linguistics still held joint conferences.

Without having undertaken any empirical research, I would speculate that close international contact with forensic linguists working in other countries and cultures (for example Arabic, Chinese, Indian and so on) is the exception rather than the rule. It can be assumed that forensic linguistic questions and cases occur quite frequently, and that the need for research in these legal systems is equally great. Moreover, internationality and a more or less symmetrical exchange of results of research and practical work is an absolute necessity. I think we all have to make an attitude adjustment in this respect. If FL today can be considered an established field of applied linguistics in research and practical work, if the result of this work is needed urgently by the judiciary (and probably the public), and if it can also be expected to have some kind of impact on the field of linguistics as a whole, one must work in an international perspective. This book is a modest attempt in this direction.

In the remaining part of this introduction, the following points will be discussed briefly:

- The general status and the overall real life situation of forensic linguistic expert testimony in Germany;
- An incomplete list of “key questions” and “central issues” that have been addressed in forensic linguistic expert testimony in Germany by the author in the last 35 years;
- Some brief comments on the selection of papers as a whole and on the genealogy of the papers forming the chapters of this volume.

Linguistic expert testimony *in or pro foro* in Germany

The selection of papers and areas of expert testimony represented in this volume does not nearly cover *all* areas in which FL experts have given testimony in Germany, or in which the author has given expert testimony since 1973. However, the most central and important areas in which I have given expert opinions are included. These are:

- authorship attribution and disputed authorship
- offenses against German laws of defamation (the sections of the “Beleidigungsrecht”, §§185–200 StGB)
- linguistic analysis of the meaning(s) of a disputed utterance/text or part of an utterance/section of a text
- matters of press laws (in particular “Widerruf” §1004 BGB (‘press revocation’))
- trademark cases.

Only trademark cases (mainly concerning phonetic similarity of product names) are somewhat accidentally missing in this volume. According to my own experience of giving forensic linguistic expert testimony, the areas mentioned also comprise the most frequent cases in which forensic linguists are asked for expert testimony in Germany (for a more detailed discussion, see Chapter 2). For the statistics of forensic linguistic expert testimony in Germany, one can only state that no reliable (and, for the most part, not even basic) statistical data are available to date. The Bundeskriminalamt (BKA), the only federal government institution in Germany that has a working unit for FL in addition to working units for forensic handwriting analysis and forensic phonetics – which are the only ones in some Landeskriminalämtern (LKA) in Germany – has some basic statistics for its own forensic linguistic experts’ activity. It concerns cases of “Schwerstkriminalität” (‘capital crime’) only, however, which represents only a small part of all linguistic expert opinions given by linguists working in universities, in other institutions, and as private experts in Germany.

A word on my own activity as a forensic linguistic expert in the last 35 years or so: I have been consulted in some 500–600 cases of various kinds, complexities, levels and categories (an exact number is difficult to give, since sometimes cases never went to court or were not followed up further because the parties reached an agreement, and so on). In some 350 criminal and civil cases, I sent a written expert opinion to the people

that had requested it; usually courts, investigating authorities, the prosecution and the defense (see below).

In most cases, it was a regular “Forensisch-linguistisches Sachverständigengutachten” (‘forensic linguistic expert opinion’). However, in some cases, I was asked to give an expert opinion on expert opinions given by other people, and to re-evaluate the linguistic data (called “Obergutachten” in German and usually requested by a superior or appellate court in Germany).

I would like to point out that the word “cases” applies to a somewhat large variety of legal procedures and data, some consisting of hundreds of pages of incriminating texts; for example, in a defamation case. Extortion and threatening letters are usually shorter (see Chapters 7 and 8 in Part III), perhaps being as brief as one sentence (for example, instructions as to where to deposit ransom money), or even one disputed word or word meaning in a given linguistic and extra-linguistic context. On some cases, I worked for several months (sometimes together with other FL colleagues); on others I worked just for one or two weeks (which is quite rare).

On average, a linguistic expert opinion roughly takes some four to six weeks of intensive analysis of the incriminating texts (“Tatschreiben”) and the comparison data (“Vergleichsschreiben”), depending on the overall size of the two classes of texts.

This, generally speaking, seems to distinguish forensic linguistic expert opinions from forensic phonetic and forensic handwriting analysis expert opinions, which, due to the nature and size of the data, can generally be done in relatively shorter amounts of time (see Chapter 1). A forensic linguistic expert opinion on an anonymous threatening letter with various suspects, with five pages of comparison data of each, requires that each line and each textual detail of the incriminating letter(s) has to be matched and evaluated against each line and textual detail of the comparison texts. In many cases, this implies an enormous amount of work.

It should also be mentioned that, in some cases in which I was consulted, a brief communication with the party requesting the expert opinion had the effect that a forensic linguistic expert opinion was no longer wanted, since it was pointed out that the chances of reaching safe and firm results that meet the requirements of a court of law were not very high. I have made it a professional habit to sum up the linguistic nature and the odds of a case on the grounds of the data available in a brief statement of the testability (“Begutachtbarkeit”) of the data before actually starting the large-scale data analysis of a full-fledged expert

opinion. This is not meant as an anticipation of the actual outcome of the expert opinion, however. If it were possible to predict the outcome of a forensic linguistic expert opinion in this fashion, one would probably not opt for (and not request) a linguistic expert opinion in the first place. Usually, people requiring an expert opinion greatly appreciate this practice.

A wide range of cases can also be noted with regard to the actual outcome that linguistic expert opinions have, and to the use that people make of them. In several cases I worked on, a thorough, substantial and safe (court-proof) expert opinion was not used in the court proceedings by the people who requested it. This is understandable, for example, in a defamation case where, even if a party could have prevailed on the basis of an expert opinion, the potential harm to a company or a person might have been much greater than any possible gain. In other cases with fewer and less substantial data and a not very promising outlook, expert opinions were used extensively in court, sometimes with success and sometimes not.

In yet a different class of cases, the forensic linguistic expert opinions requested were not used in or for any legal procedures at all. They never made it to court, so to speak, although they were originally conceived and designed for that purpose. In quite a few (mainly civil) cases, the parties reached some kind of agreement and settlement, in which the linguistic expert opinion was used as a basis, or background information, for a solution. This way, the case was not publicly tried or publicized at all. Here lies part of the reason why statistics on and documentation of forensic linguistic expert opinions are, by their very nature, a difficult matter. If, for example, a company suspects a former employee of having written an anonymous defamation letter – say, on the weakness of a particular product – that company would not wish to take it to trial, even if the forensic linguistic expert opinion stated that it was safe to assume that the suspect was indeed the author of the anonymous letter, because exposing that product weakness might injure its reputation.

As far as the general distribution of my own expert opinions in the different areas of FL is concerned, I can only offer a rough estimate. Of the cases that went to court, some 60 per cent of the expert testimony I have given was on criminal matters, about 40 per cent on civil cases. For the cases that did not go to court, the proportion is reversed.

In terms of FL areas, most of my expert opinions were on authorship attribution (about 40 per cent), followed by insult and defamation cases (about 30 per cent), the remainder (about 20 per cent) on text-analysis and text-comprehension cases, including trademark cases. Almost all

expert opinions on authorship attribution were requested by courts or judges (some 70 per cent), or by higher investigating authorities (by the police, some 30 per cent), and hardly any by private parties. In civil cases, there was an approximate ratio of 50:50 of courts and private parties requiring a linguistic expert opinion. As mentioned earlier, these are very rough estimates only.

However, several contrasts between forensic linguistic experts' work in Germany and, for example, in the US may shine through here. It is not the intention of this section to elaborate on the more general differences of the legal systems. Several of the more detailed contrasts between the American and the German legal systems are outlined in later chapters. For example, some differences between the American libel and slander laws and the German "Beleidigungsrecht", in which both libel and slander fall together and other distinctions prevail (see Chapter 5).

There are other basic differences that should be stated, or restated, here since they do not seem to be that well known among linguistic colleagues abroad. As mentioned, a forensic linguistic expert opinion is frequently requested by the court (the judge) in Germany; less frequently by a party. The court or judge is absolutely free in the evaluation of a forensic linguistic expert opinion. The "Freie Würdigung des Sachverständigenbeweises durch das Gericht" ('Free evaluation and assessment of the expert opinion by the court') is a basic constituent of German law. It gives judges a wide range of options at their discretion. The judges may request a forensic linguistic expert opinion or they may not, acting as linguistic "experts" themselves. A judge may even reverse the results of a particular forensic linguistic expert opinion or use it in favor of the opposite side. This has, in fact, never happened to my own expert opinions, but it is theoretically possible.

Another reason why statistics on expert opinions are so hard to come by is the feedback from the judiciary: Although I had explicitly asked courts and judges for some feedback on how the judge used my expert report, I received responses in only a small proportion of cases.

Undoubtedly, a noticeable contrast between the American and the German legal systems is that in Germany in only some 5 per cent of all cases do FL experts actually appear in court. In more than 95 per cent of cases, FL experts provide only a *written* report. The court, in turn, usually quotes parts of the expert opinion – in particular, the conclusion – in the judicial opinion. This is, of course, entirely the court's decision, and I have no feedback on the practices of the judiciary in this respect. Only in a very few cases in which forensic linguistic, forensic phonetic or forensic handwriting analysis is used (and only if there is a particular

strategy and request by the defense lawyers), is an expert ordered to appear in court. In the past, this happened particularly frequently in the RAF (“Red Army Fraction”) terrorist cases in Germany. One of the reasons for the rare appearance of FL experts in German courts seems to be that in Germany there is no such thing as cross-examination as found in an American court. Also, the legal status of an expert (“Sachverständiger”) is not identical (see Chapter 1) to that of a witness (“Zeuge”). If, in a civil case, there is a dispute between parties on a particular linguistic issue, each party usually requests its own forensic linguistic expert opinion, and it is at the court’s discretion to decide which to follow. In some cases, an “Obergutachten” (‘superior expert opinion’) is requested by the court to help it reach a decision.

There is also a distinction specific to the German situation between so-called “Behördengutachten” (‘authorities expert opinions’; expert opinions given by people working in federal or state offices) and “private Sachverständigengutachten” (‘private expert opinions’). It is hard to generalize but, for a somewhat large part of (particularly lower) German courts, it may be that “Behördengutachten” enjoy higher credibility than “private Sachverständigengutachten”. In addition, courts and parties are frequently keen to request linguistic expert opinions from a so-called “authorized and publicly sworn in private expert”, who has to be registered with the court and the local Chamber of Commerce. This expert *has* to respond to the request by a court or other authority to give expert testimony if asked to do so. A private (freelance) expert can choose whether to give expert testimony or not.

As mentioned earlier, only on a very few occasions do FL experts appear in court. In the “Behördengutachten”, there is a particular line at the top relating to §256 StPO (“Strafprozessordnung”, ‘criminal court procedures’) stating explicitly that the expert opinion should be read by an authorized person of the court. The characteristics of the German legal system noted above imply that a linguist is hardly ever exposed to critical questioning, cross-examination and so on by lawyers in court. Rather, they face criticism from fellow linguists. At worst, the judge may not accept the expert opinion and decide either to request a further report (from a different linguist) or simply to ignore it. In a civil case, a linguist giving expert testimony battles with one or more linguists hired by the other side. It seldom happens that a lawyer in a civil or criminal case opposes (in court) the expert testimony given by a linguist on scientific linguistic grounds.

It should also be pointed out here that there is an important distinction in German criminal court procedures. They have a twofold structure of

two consecutive parts, consisting of: (1) the “Ermittlungsverfahren” (‘preliminary (investigative) proceedings’); and (2) the “Hauptverhandlung” (‘main trial’). The first is the domain of FL expert testimony (and rightly so, in my view). Its main effort and value lie in the analysis of relevant data and criteria supporting those of the other auxiliary sciences.

Range of variation of linguistic expert testimony in the German legal context

This brief characterization of areas in which linguists are asked for expert opinions is somewhat unsatisfactory. Rather than merely describing the general areas, it seems more productive to give a list of “key questions” and “central issues” that have actually played a role in the expert testimony I have given. Even beyond this specification, it seems worthwhile to exemplify the continua of shades of different questions and cases of a type that forensic linguistic expert testimony deals with.

I now present a selection of “key questions” and “central issues” of cases in which I have been consulted as a forensic linguist. By “key questions” (henceforth KQ), the original wording of the request for a linguistic expert opinion is meant. In all cases, the questions are rephrased and operationalized by the linguist. In expert opinions, I have habitually pointed out the difference between the wording of a question by linguistic laypeople (judges, lawyers and others) and an operationalizable scientific linguistic question. This did not cause miscommunication between the people requesting the expert opinion and the expert. It generally worked quite well, as long as the explanation given in the expert opinion was detailed enough and easily intelligible. To give an example: The question “Is the word *concubine* an insult in today’s German?” (“Ist das Wort *Konkubine* eine Beleidigung im heutigen Deutsch?”, see Chapter 5) cannot be answered linguistically in any sensible scientific way for at least two reasons: (1) the law does not define the term “Beleidigung” (‘insult’) for good reasons; (2) it is not possible to give an empirically safe linguistic definition of a word X being an insult once and for all (without analyzing the data of the linguistic and extra-linguistic context).

The translation of laymen’s questions into a set of linguistic questions is an intrinsic part of every forensic linguistic expert opinion. In fact, it is one of the most important parts. The best linguistic analysis is not worth much if the people who requested it and the others involved in the trial do not understand it. Ample exemplification for this is given in the chapters in Parts II and III. Under “central issue” (henceforth CI), I will

describe the general heading under which the particular type of case would be customarily classified in FL (as practiced in Germany and abroad, as far as I can see). Three general areas of expert testimony will be distinguished here: first, authorship attribution; second, insults, libel and slander; and third, analysis of text comprehension and meaning. The fact that CI is placed before KQ does not represent the sequence in the analysis; it is used here for practical descriptive reasons only.

Authorship attribution

1. CI: Authorship attribution of a set of anonymous extortion, blackmail, threatening letters and of comparison data letters of a given suspect (on the basis of non-linguistic criminal evidence)
 KQ: *Is X (the known author of the comparison data letters) the author of the incriminating anonymous letters?*
2. CI: Forensic linguistic authorship attribution of a set of incriminating anonymous handwritten or typed letters, collaborating with or aiding the examination given by handwriting experts and document analysts, typewriter script experts, and so on
 KQ: *Is the writer/typist of the incriminating anonymous letters also the author (text originator) of the incriminating anonymous letters? Or vice versa: is the author of the incriminating anonymous texts also the person who has written them by his/her own hand or typed them?*

An enormously large variation of possible (and actual) combinations is discussed in Chapters 6, 7 and 8 in Part III.

3. CI: Authorship attribution of incriminating anonymous letters in collaboration with forensic phonetics experts' analysis of anonymous phone calls
 KQ: *Is the person who made the anonymous phone calls the same person who wrote/authored the incriminating anonymous letters? Or vice versa: is the person who wrote/authored the incriminating anonymous letters also the person who made the anonymous phone calls?*

Since, depending on the nature of available data, much depends on the interdisciplinary *cooperation* of the various supporting sciences involved, this is of particular importance in gaining information that can be of use in criminal matters in general. It goes without saying that this applies to all sciences employed as auxiliary sciences in a case.

4. CI: Forensic linguistic authorship attribution of incriminating anonymous letters with disguised authorship: linguistic criteria are needed that show that person X, who is assumed to have written the incriminating anonymous letters on the basis of non-linguistic evidence but purposefully disguised her/his authorship, has indeed authored and written the incriminating anonymous letters.

KQ: *Is X, who is assumed to have written and authored the incriminating anonymous letters, to be confirmed as the author by means of linguistic analysis rather than Y, who functions as the official or explicitly named author of the anonymous incriminating texts? In short: is X simply faking or imitating Y's authorship to disguise her/his own?*

5. CI: Forensic linguistic authorship attribution in a case of disputed multiple authorship

KQ: *Which of the three suspects on trial, A, B or C is the author of a set of anonymous extortion letters?*

In this case, three people were tried for several crimes, including armed robbery and extortion, which they committed together. There is a set of anonymous extortion letters which suspect A (a wall painter) confessed to have written himself according to the dictation of suspect B (a physician); A admittedly is the writer, but says not he but B is the author; B contests this and says A himself wrote and authored the letters (C undisputedly is not involved in this part of the crime).

In this case, I was not able to come up with any linguistic data solving the question of multiple authorship.

6. CI: Forensic linguistic authorship assessment in a different (civil court) case of multiple authorship

KQ: *For a joint venture literary translation in which, undisputedly, two authors (translators) have participated, a disagreement arises about who has done more work (and deserves more money); the one who supplied a first rough translation into German, or the one who supplied a poetic, idiomatic German text version.*

Insults, libel and slander and other defamation offenses

7. CI: Forensic linguistic text analysis (analysis of the linguistic and situational context) of a particular word or phrase in letters written by a person living in the same house as the plaintiff

- KQ: *Is the word X (Konkubine), as used for a neighbor's partner in letters addressed to the owner of the house, an insult ("Beleidigung", "Verbalinjurie", 'verbal injury') in German (in 1972)?*
8. CI: Forensic linguistic analysis of a dialectal word or expression in a special linguistic and situational context
KQ: *Is it an insult in German to call another driver Sie Ochse! ('you ox!') in a traffic accident?*
9. CI: Analysis of non-verbal gestures and hand signs
KQ: *Is it an insult in German to form a circle with both arms above your head meaning Sie sind ein Riesenarschloch! ('you are a huge asshole!')*
10. CI: Analysis of a joking or sarcastic use of a derogatory expression
KQ: *Can the expression X be taken as a joke or a sarcastic remark rather than seriously (as an insult) in a given linguistic and situational context?*

Disputed meaning of a text passage or utterance

11. CI: Forensic linguistic determination of the range of meaning(s) in a particular advertisement: can one exclude meanings X_2 , X_3 , X_4 for a text fragment X and confirm by a linguistic analysis that *only* meaning X_1 is possible?
Legal background: if only a meaning X_1 is possible, the plaintiff's charges are justified and the defendant may be forced to retract a particular advertisement.
KQ: *Can it be proven that a text fragment X only has meaning X_1 , rather than also having meanings X_2 , X_3 and X_4 ?*
12. CI: Forensic linguistic determination of the pragmatic status of a particular phrase (slogan) and its obligatory allusions, associations, connotations, and so on
KQ: *Can an expression (Wir gehören zur Familie) be used by a social minority group (gay activists) without creating associations to the advertisement slogan used by a large industrial company in Germany?*
13. CI: Determination of the similarity status in phonetics, semantics and language use of a particular word or expression
KQ: *Does the word X used for a car obligatorily or necessarily create associations to the name of a given magazine, or are both perceptually (sufficiently) distinct?*

As can be seen from the examples given, the list could easily be extended. It illustrates that there is an enormously wide range of questions that FL experts are asked to help clarify. In fact, to complicate things even more, there is sufficient reason to assume that a continuum of varying specifications exists for each of the issues mentioned. This can be illustrated in more detail by the examples given below for disputed multiple authorship. It also illustrates that, even in the same narrow area of linguistic expert testimony, no two court cases are exactly alike. The cases listed below exemplify a continuum of different nuances of issues to be elaborated on by the FL expert. Each case is different from any other.

Multiple authorship can be quite a different matter, if one assumes just two basic dimensions of variation, the writer-/typist-ship and the authorship of a (single) text product.

- (a) Multiple authorship can mean (as described in case 5 above) that one of several suspects involved in a crime admits that he has handwritten an extortion letter, but contests that he is also the author of the text and claims that another person has dictated it.

CI: Authorship attribution in a case of disputed multiple authorship (and non-disputed writership)

KQ: *Is the writer of the anonymous extortion letter also its author, or was the text dictated by someone else?*

- (b) In another case, the investigating German authorities are concerned that part of an anonymous extortion letter has originated from members of organized crime, and request a linguistic expert opinion to exclude this. The FL expert is expected to apportion the anonymous authorship in regard to different authors.

CI: Heterogeneity of text portions in a threatening letter

KQ: *Can it be linguistically determined who has written different portions of a threatening letter?*

- (c) In the final written examination of law students in Germany (“1. juristische Staatsprüfung”), two sisters show a very poor performance (grade F). This is the officially supervised final written examination (“Klausur” of the “1. juristische Staatsprüfung”). However, both do very well indeed (grade A) in their “Hausarbeiten” (‘law theses’), which are to be written at home within six weeks. Both theses were not only written very well, but also in a remarkably

similar fashion. The law school examination board suspects that a third author with considerable experience in the law profession is the true author of the two “Hausarbeiten”.

CI: Forensic linguistic authorship attribution in a case of illegal intertextual similarity

KQ: *Are there linguistic clues as to whether both sisters can be excluded as the authors, and a third person must be attributed the authorship of the two theses? Furthermore, is there linguistic evidence that this author is one and the same person?*

(d) An example of yet another variant of disputed multiple authorship is a negative forensic linguistic authorship attribution; that is, the exclusion of authorship of an article appearing in the media.

CI: A letter to the editor appearing in a daily newspaper is quoted as authored by a director of a local hotel, who denies the authorship.

KQ: *Are there linguistic features indicating that the hotel director has indeed not written this letter, supporting his claim that he would not write in this way?*

This selection of a continuum of disputed multiple authorship cases could easily be supplemented by dozens of similar cases, which illustrates that the judicial saying “each case is different from any other” also applies to the forensic linguistic state of affairs. Moreover, it highlights the large range of variation in each particular issue and question listed in “real life”.

In my opinion, this is part of the arsenal of the problems and challenges that forensic linguistic expert testimony faces, particularly in the area of various types and dimensions of authorship attribution. It also illustrates the amount of professional experience needed to work as an FL expert – and the fascination of this type of work as an applied linguist.

Overview of chapters and notes on the genealogy of the texts

In the following, some comments on the selection of the chapters of this volume as a whole and a brief characterization of the texts are offered.

To begin with, the selection of papers as such is in no way representative of my work in the field of FL, let alone of FL in Germany as a whole. Although I can attest that there was a lively practice of forensic linguistic

expert testimony in Germany and documentation of it in publications as far back as the early 1980s (see Kniffka 1981), this book does not try to give a detailed account of the historical development of the field in Germany.

The chapters in this volume share a common feature: to date, none of the texts was available in English. All texts have been revised for this work, which also contains several new pieces.

Chronologically, the texts of the three chapters in Part II are considerably earlier than those in Parts I and III. The texts in Part II date back intellectually as far as the early 1970s. The very first is the pioneering paper “Der Linguist als Gutachter bei Gericht”, which appeared in 1981. Most of the texts of Part II were presented in German at forensic linguistic, forensic phonetic and handwriting expert conferences in the late 1980s and early 1990s in Germany. They were presented, in particular, at the Mannheimer Tage zur Schriftvergleichung organized by the late Lothar Michel in Mannheim, Germany, and at the 1988 meeting organized by the BKA in Wiesbaden, Germany. Parts of Chapters 2 and 3 were also presented in English, at the first conference of the International Association of Forensic Phonetics (IAFP) organized by Peter French in York, UK, in 1990. A brief description of this early history of FL conferences in Germany and the UK is available on the website of the International Association of Forensic Linguistics (IAFL) run by Susan Blackwell and Jess Shapero at the University of Birmingham.

The texts of the English translations of the earlier papers in Part II vary from the original German texts to a degree. In some instances, the original German text was shortened. In most, it was extended. No substantial changes in the texts of Part II were made. The texts are included here for documentary reasons, to show that many of the ideas that are claimed to have arisen in the 1990s and 2000s are actually older. Since the 1970s, they have played a substantial role in the theory and practice of FL expert testimony in Germany.

The chapters in Part III (“Ongoing Research in Forensic Linguistics in Germany”), and also Chapter 1 in Part I, are more recent in their textual genesis than those in Part II. Chapter 5 (“Libel, Linguists and Litigation”) was written for a presentation at the ILA Conference at the John Jay College of Criminal Justice in New York in April 2005. A German version of Chapter 7 and, in part, Chapter 8 appeared in the German linguistic journal *Osnabrücker Beiträge zur Sprachtheorie (OBST)* in 2003. Chapter 8 was presented in a new and extended version in English at the seventh IAFL Congress in Cardiff, UK, in July 2005. So, in brief and oversimplified, the chapters forming Part III are entirely new texts and were written specifically for this volume.

Parts of Chapter 9 in Part IV (“Language and Law: Some Needs”) were presented at the AILA Conference at Jyväskylä, Finland, in 1996. There has been an online version of it since 1998. This is a revised version, and the first version to appear in print.

A brief summary of the main points of each paper and its textual genealogy now follows.

Chapter 1 (“Forensic Linguistics: Its Relatives and Neighbors. An Interdisciplinary Perspective”) in Part I (“The Interdisciplinary Status of Forensic Linguistics”) gives an overall characterization of the field of FL as a branch of applied linguistics and as one of the auxiliary forensic sciences. In so doing, it provides a heuristic typology of where to position FL as a branch of applied linguistics, and pinpoints some of the major research areas and goals that have been established so far, drawing mainly on the modern literature; in particular, Kniffka (1981; 1990a; 1998; 2000b), Levi and Walker (1990), Shuy (1993a; 1998a; 1998b; 2000; 2002b; 2005; 2006), Solan (1993), Solan and Tiersma (2005), and Tiersma (1999). The main purpose of the paper is to elaborate on the major structural differences and similarities of the four neighboring sciences: FL, forensic phonetics (FP), forensic handwriting analysis (FHA) (officially called “*Gerichtliche Schriftvergleichung*” in German) and document analysis. It is assumed that FL can most adequately be described by a demarcation of its position between the two closest neighboring fields, the sister disciplines FP and FHA.

A closer look at, and a more detailed analysis of, the interrelations and the methodological similarities and differences are of critical importance for the development of all three fields in the future. This also is motivated by the “real life” fact that, in many criminal cases, all three sciences are involved due to the nature of the data available. An interdisciplinary perspective on more remote auxiliary forensic sciences may also be very helpful; not in terms of shallow ad hoc analogies from other sciences but in terms of a solid and sober reflection (for example, on modern quantitative methods of the natural sciences).

Part II (“Forensic Linguistic Research in Germany in the Past”), mainly Chapters 2 and 3, covers the historical development and research tradition of FL in Germany. Chapter 2 (“The Linguist as an Expert Witness in German Courts. A View from the 1970s”) is a translation of the section on anonymous authorship attribution in the 1981 paper “*Der Linguist als Gutachter bei Gericht*” (“The Linguist as an Expert Witness in Court”). The English translation is but one fifth of the total text of the original. The paper was written on the basis of some 25 legal expert opinions in cases in which I had been consulted at the time. It summarizes, in brief,

an exemplaric selection of these cases – lacking, for example, a description of the experiences and also the “suffering” I went through while writing it. There was no precedent of a linguistic report available to me on forensic linguistic expert testimony. Nor did I have any knowledge of or contact with colleagues, abroad or in Germany, working along the same lines (my first contacts with Roger Shuy, then at Georgetown University, USA, and Peter French, then at York University, UK, date back to 1988 and 1989, respectively).

The original 1981 paper gives a systematic survey of FL activities in German courts to date and the types of interaction between forensic linguists and the judiciary. The excerpt of it translated here tries to position FL as an auxiliary science in the criminal investigation of anonymous authorship. It also relativizes the tasks, possibilities and limitations of forensic linguistic authorship attribution. It warns against the nonsensical notion of “linguistic fingerprints”, which were advertised as some kind of methodological miracle weapon by a self-proclaimed forensic linguistic “expert”, inexperienced and non-professional in his claims. The 1981 paper states clearly: “Es gibt keine ‚idiolektalen Fingerabdrücke‘, und wenn es sie gäbe, hätte die Linguistik wohl kaum eine Möglichkeit, sie exakt zu messen” (“There are no “idiolectal fingerprints”, and if there were, linguistics would not have the means to measure them in an exact fashion”) (Kniffka 1981: 598). As if this were not enough, several people inside and outside linguistics were then talking about “die linguistischen [sic] Fingerabdrücke” in publications in Germany, which actually would mean ‘the fingerprints of linguists’ (rather than ‘idiolectal fingerprints’).

The paper also states that a forensic linguistic authorship analysis is not able to reach the same stringency and accuracy as a phonetic speaker analysis, let alone an analysis of DNA or of other kinds of samples by natural sciences. This was stated explicitly since, at the time, people were having and giving exaggerated expectations vis-à-vis forensic linguistic authorship analysis. In many cases, an FL expert was called in after the analyses by other (more exact) sciences had failed. The paper explains that it is also an ingredient of real life requirements that one has a realistic view about the possibilities and limitations of forensic linguistic authorship attribution.

Chapter 3 (“Status and Tasks of Forensic Linguistic Authorship Analysis”) is a translation of a paper that appeared in the German journal *Kriminalistik* in 1990, which addresses practitioners; in particular, the police. It is a summary of a paper read at the 1988 meeting “Forensisch-Linguistischer Textvergleich” organized by the BKA in Wiesbaden, Germany (it appeared in press in *Bundeskriminalamt* (ed.) 1989: 205–36).

In short, the 1990 paper in *Kriminalistik*, of which the present version is a translation, summarizes the practical aspects and concerns of forensic linguistic authorship attribution. It does so by clarifying and defining some FL terms, and by describing the method of authorship attribution in layman's language. As stated in Chapter 2, analyzing dynamic language behavior (as represented in frozen text products) and stating co-variation and feature configuration on a large scale have been salient notions of authorship attribution up to the present day, rather than looking at single static linguistic data and assigning any identifying potential to them. The paper also clarifies that generalizations in forensic linguistic authorship attribution always and only concern the text product under investigation itself, not the totality of text products a particular anonymous author has written, can or will write, or even the total personality of an anonymous author. It goes on to say that the notion of a profile for the writing of an author may have some relevance in the future but should perhaps be limited to psychological analysis at present.

It may also be of interest to comment on the historical context of the 1988 BKA meeting and the 1989 volume that originated from it. The meeting was organized (unexpressedly) to rebut the unjustified criticism by a charlatan forensic linguistic "expert", who had given testimony in a number of spectacular cases and was making much ado about it in the press. Also, the BKA wanted to counteract criticism it had faced in terrorist cases for linguistic expert testimony given by members of the BKA who did not have linguistic training as such. What is interesting about this is the fact that it took the criticism of an FL charlatan (who did not attend the meeting) in the press to initiate the first ever FL conference on authorship attribution – in Germany and internationally. It is also worth mentioning that the large majority of the participants at the conference had not given forensic linguistic expert testimony at all.

This seems to represent an overall characteristic of FL, at least in Germany at that time, and perhaps also of other branches of applied linguistics: only if there is an obvious urgent need will people show some actual and public concern. Sometimes it seems that FL has been perceived as an ad hoc necessary evil rather than as a branch of science needing substantial and systematic support, at least as much as theoretical linguistics.

Chapter 4 ("‘Shibboleths’ as Data of Linguistic Behavior") is not really an FL paper. It is included here as one example of general linguistic basic research that seems of critical importance, at least indirectly, for forensic linguistic authorship attribution. It is based on, but not really a translation of, a German paper titled "Schibboleths. Philologische Bestandsaufnahme und Gesichtspunkte zu ihrer soziolinguistischen Analyse", which

appeared in the journal *Deutsche Sprache* in 1991. It accounts for the historical and philological evidence that the much-quoted notion of “shibboleth” entails, starting from the use in the bible (Judges 12: 5–6) up to its recent practical uses; for example, in the Lebanese civil war.

My interest in this matter came about because of the (mostly futile) desire to find some linguistic literature on the question of idiolect, beginning with Edward Sapir’s fascinating 1927 paper “Speech as a Personality Trait”. It turned out that shibboleths prevail in the area of group-specific rather than individual or idiolectal language behavior. The paper relates the older (and mostly rather traditional) work to modern approaches in linguistics and adjacent sciences, notably to the articles collected in Scherer and Giles “Social Markers in Speech” (1979), and develops a heuristic taxonomy and terminology for what could be called shibboleth-type features of linguistic behavior.

This reflects my conviction that sociolinguistics and psycholinguistics will probably be the major fields from which FL can, and should, draw its methodological and theoretical ingredients most prevalently.

The wording “Ongoing Research in Forensic Linguistics in Germany” for Part III is somewhat inadequate. First, because the chapters in Parts I and II refer to or imply ongoing research. Second, because it refers only to my own research and omits all research undertaken by other German colleagues. This simply results from necessity. There have been very few publications by German forensic linguists of which I am aware, and the number of people working in this field with an international perspective in Germany still seems rather small.

Chapter 5 (“Libel, Linguists and Litigation”), based on parts of the 1981 paper, was written for the international ILA Conference at the John Jay College of Criminal Justice in New York in 2005. It is entirely new in content and structure, and tries to take the international audience into account. It gives an exemplaric description of an insult case tried in German courts, concerning the use of the word *Konkubine* (‘concubine’) in German in the early 1970s, which was the very first expert opinion the author gave (for the Oberlandesgericht Köln, a German superior court, in 1974). The article elaborates on the salient features of German defamation laws, the linguistic intricacies of the use of the disputed word in 1972, and also the convincing arguments that made the court follow the author’s conclusion.

What really convinced the court that the word *Konkubine* was used in a pejorative rather than neutral sense was a detailed analysis of data from different text types and everyday communication situations in German, rather than a full-fledged theoretical deduction in the terminology of

contemporary linguistics. The paper further elaborates on the necessity of adapting the wording of a forensic linguistic expert opinion to a level that a judge, as a non-linguist, can fully understand, the avoidance of any juridical and legal jargon in the linguistic expert opinion, and several other pedagogical matters, which are of critical importance for the success of a linguistic expert opinion in court. As Roger Shuy, Geoffrey Pullum, myself and many others have expressed time and again, the linguistics of libel and slander is one of the most fascinating, promising and also treacherous areas of the application of linguistics. As Pullum (1991) states:

I have spent a little time studying the laws of libel and slander from a linguistic standpoint. I find it a fascinating but rather scary topic. For anyone who has an ounce of concern about guarantees for freedom of expression, it is highly thought-provoking. But most people, linguists included, know relatively little about it. (1991: 94).

What Pullum says about the general background of linguists in these areas certainly holds for the 1970s, and even today. Semanticists and pragmatists, generally speaking, do not work in the area of defamation laws. Moreover, they hardly know what is going on in the field at all. This is difficult to understand, let alone justifiable, since in many cases the exact semanto-pragmatic analysis of defamation is of much greater real life concern than most analyses made by semanticists and pragmatists. Also, according to my own experience, the application of the semantic and pragmatic analysis in an FL context has great interest and pedagogical effect for students of linguistics. In my view, there is no better and more adequate justification of the real life applicability of linguistics than this.

Chapter 6 (“A Heuristic Author and Writer/Typist Taxonomy”) was also written especially for the present volume. Starting with the empirical fact that one cannot just take for granted a one-to-one relation of text author and writer/typist of an anonymous incriminating letter, this paper gives a more systematic account of the possible combinations of authorship and writership of a written text product, and develops a preliminary heuristic taxonomy to model the dimensions of variation that have to be taken into account. Five dimensions of variation are defined and examined for their real life occurrence. In addition, a set of binary features for a more systematic description of authorship and writership is drawn up, which can be used in combination with matrices to allow a more adequate description of types of incriminating text products.

For matters of illustration, some ten binary features are defined (for example, [\pm id] for identity or non-identity of author and writer of a text; [\pm dis] for disguised or non-disguised authorship). This allows a classification of several of the more frequently occurring types of incriminating text products. Needless to say, this is just an exemplaric study that needs further revision, including many more features and dimensions. It seems that this is a viable start to achieve a more systematic, and theoretically more adequate, description of the variation of text products that occur in incriminating texts. It also seems possible to extend this model of authorship and writership beyond the scope of anonymous incriminating texts to describe dimensions of variation of texts concerning incriminating and non-incriminating texts and text types.

It is hoped that this type of description will help to achieve a more adequate empirical and theoretical account of incriminating texts that may be of practical use for the investigating criminalist right from the start. It can also make it possible to describe some particular types of anonymous incriminating letters in a more structured and adequate way. To note just one example: there is no empirical and theoretical study to date of the occurrence of disguised authorship and/or writership. "Disguised" can mean different things, depending on the particular data constellation – not just anonymity of the name of the author and/or writer or the use of a pseudonym instead of the real name. It can mean several forms and layers of non-revealing and hiding one's name. One might be able to specify what "disguise of authorship" versus "disguise of writership" (that is, handwriting) amounts to in more adequate empirical terms. "Disguise of authorship" could be defined as a deliberate choice of language behavior, different from that which would be one's own customary way of speaking or writing in communicative situations in texts of a comparable nature, in order to conceal one's authorship.

"Disguise of writership", which seems to apply to handwriting only, or at least more than to typed script, can be defined as "the deliberate alteration of one's customary handwriting in order to conceal one's writership of a text". This has been addressed by handwriting experts for several decades; see, for example, the discussion of kinds and techniques of disguise of handwriting in Michel (1982: 178–206). This has never been related to the findings of forensic linguistic authorship analysis as outlined above. It seems that future research can gain substantially from elaboration of the theoretical and empirical impact given in such structural relations. The interrelation of FL and handwriting data also needs

further systematic analysis for the description of types of orthographic behavior (see Chapter 8).

Chapter 7 (“The System and Diagnostic Potential of Orthographic Data in Forensic Linguistic Authorship Attribution”) is, in part, a translation of a paper in German that appeared in the linguistic journal *OBST* (66: 85–116) in 2003. It is based on some older papers (Kniffka 2001), but the English version was entirely revised, rewritten and enriched by several additions on text types, characteristics of incriminating texts, and the interrelations of grammatical and orthographic behavior.

There has been some vagueness, or even confusion, about the role of orthographic data in authorship attribution. The paper discusses the theoretical basis of data and reassesses (in a sociolinguistic fashion) dynamic features and styles of orthographic behavior that are worth taking into consideration in anonymous authorship attribution as well. Orthographic data can never function as sufficient evidence alone, but frequently function as evidence of data of written texts indirectly relevant for authorship attribution.

Chapter 8 (“Orthographic Data in Forensic Linguistic Authorship Analysis”) is an extended version of a paper read at the seventh IAFL Congress held at Cardiff, UK, in July 2005. It gives a detailed account of the various classes and types of orthographic data, and their descriptive and explanatory potential in cases of anonymous extortion and threatening letters in which the author was asked to give expert testimony in the last 30 years or so. The data are all taken from real life cases. The selection of cases is such that the enormous breadth of orthographic data are symptomatically represented. The paper goes on to elaborate on the theoretical status of continua of features of orthographic behavior, introduces the theoretical notion of an idiosyncrasy coefficient of orthographic behavior and discusses its methodological impact for the analysis of authorship and writership.

Finally, the paper gives an exemplaric, empirically based contrastive analysis of the text type-specific distribution of linguistic and orthographic features, contrasting the text type-specific grammatical, textual and semanto-pragmatic features of anonymous extortion and blackmail letters with those of defamation letters. These it correlates with the text type-specific contrasts of graphemic, orthographic and layout features in the two text types.

The paper is the first systematic account to date of the interrelation of data of linguistic and orthographic behavior in a particular text type. It seems that an evaluation of large corpora along the lines described here

will yield valuable results also for the practical analysis of text types usually occurring in anonymous incriminating texts.

Chapter 9 (“Language and Law: Some Needs”) opens Part IV of this volume (“Outlook and Future Tasks”). It is a slightly revised and updated version of a paper read at the AILA Conference in Jyväskylä, Finland, in 1996. It models the various constellations for needs research on the basis of a detailed analysis of intra- and interdisciplinary misunderstandings, miscommunications and misconceptions, of the judiciary and of linguistics each and about each other. This chapter is, in short, a summary of the real life experience the author has gained in some 35 years of giving expert testimony in and for German courts. It is left to the reader to notice which patterns of interdisciplinary behavior between linguistic and law people uniquely apply to the German scene, which apply to the American and/or British context, and which, if any, are “universal” in the sense of occurring in similar variants everywhere in the (Western) industrialized world in interactions of linguistics and legal practitioners. In my view, it is always revealing to acquire an outside perspective of one’s own culture by looking at another.

In Chapter 10 (“Outlook”), some of the recommendations and caveats for future work are briefly characterized. They reflect the view of the author that FL has found its place in the field of applied sciences and in applied linguistics, and is ready to make some terrific progress, if a few conditions of international and interdisciplinary cooperation are met. The most important of all is a much wider and more intensive exchange of information and experience across countries and cultures.

The concepts and topics discussed in this volume cover a time span ranging from the 1970s to the present time. Although this is only a small selection of FL papers that have appeared in German since 1981, people who do not read German are offered access to a somewhat more adequate notion of the development of FL in Germany in recent decades.

Note to readers on conventions of graphic representation followed in this volume:

Linguistic (“object language”) data are represented in *italics* throughout this book, e.g. the German word *Konkubine*.

Linguistic and legal terminology in German and English are given in double quotation marks (“...”), as are original quotations.

The meaning of a word, phrase or sentence is indicated by the use of single quotation marks (the German word *Konkubine* means ‘concubine’), as is customary in linguistics.

This page intentionally left blank

Part I

The Interdisciplinary Status of Forensic Linguistics

This page intentionally left blank

1

Forensic Linguistics: Its Relatives and Neighbors – An Interdisciplinary Perspective

Introduction

This chapter gives a brief account of forensic linguistics as an auxiliary science in forensic matters, and its structural and factual relationship with and its position among other sciences. The metaphorical terms in the title are used as an approximative indication of the multifold relations in which forensic linguistics can be viewed.

The term “relatives” refers to other branches of applied and/or interdisciplinary linguistics, to such areas as language teaching, language acquisition, sociolinguistics, psycholinguistics, clinical linguistics, corpus linguistics and others. It is assumed that FL has the closest and most salient ties, and the greatest level of scientific import and export with these fields. It also refers to areas and approaches of linguistics other than applied and/or interdisciplinary linguistics, such as theoretical and general linguistics, (non-applied) forms of psycholinguistics, text linguistics, computer linguistics and other fields.

“Neighbors” refers to adjacent and neighboring auxiliary forensic sciences, some neighbors being less remote than others. For the former, one would think of forensic phonetics (FP, forensic phonetic speaker recognition and identification, disputed words), and forensic handwriting analysis (FHA, in German “Gerichtliche Schriftvergleichung”, including graphometry and other fields). They are, without question, the most important and salient sister fields of forensic linguistics, from a theoretical, and – even more so – from a methodological and a factual perspective. This results from the fact that in many real life forensic cases these three auxiliary sciences are involved together, not infrequently giving a helping

hand and supporting the work of the other fields to a considerable degree. It is merely a matter of definition whether one would consider phonetics a branch of linguistics or not, but this is of no relevance here. One can consider forensic phonetics, forensic linguistics and forensic handwriting analysis as three autonomous, independent sciences, which work in close theoretical, methodological and practical contact and cooperation. The differences and interrelations between the three will be elaborated on below.

As has been discussed time and again (see, for example, Kniffka 1981; 1990b; 1996a; 1998; Levi 1982; Levi and Walker 1990; Shuy 1993a; 1998a), forensic linguistics in the most generally accepted definition is but one branch or subfield of the wide field of “language and law” (see Levi 1982). Some people use the term “forensic linguistics”, in a very wide definition, nearly synonymous with “language and law”. For the description of inter- and trans-disciplinary relations envisaged here, a much narrower perspective has been chosen. If forensic linguistics is roughly defined as linguistic experts’ analyses of language data for a multitude of purposes and addressees, including *in* and *pro foro*, only one out of several branches of forensic linguistics has been chosen for comparison here; the area of forensic linguistic “anonymous authorship analysis” or, as it is more commonly known in the English-speaking world, forensic linguistic “authorship attribution”.

A simplified illustration of some basic structures and interrelations is given in Figure 1.1, for the field of linguistics, and in Figure 1.2, for forensic linguistics within the entire field of auxiliary forensic sciences.

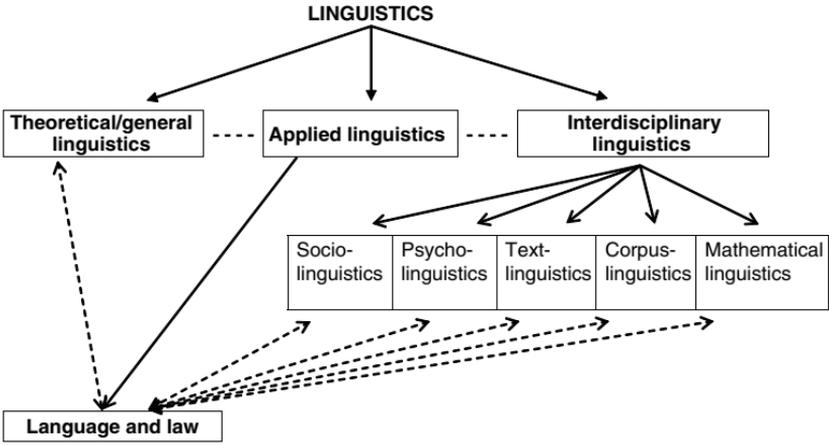


Figure 1.1 The field of linguistics

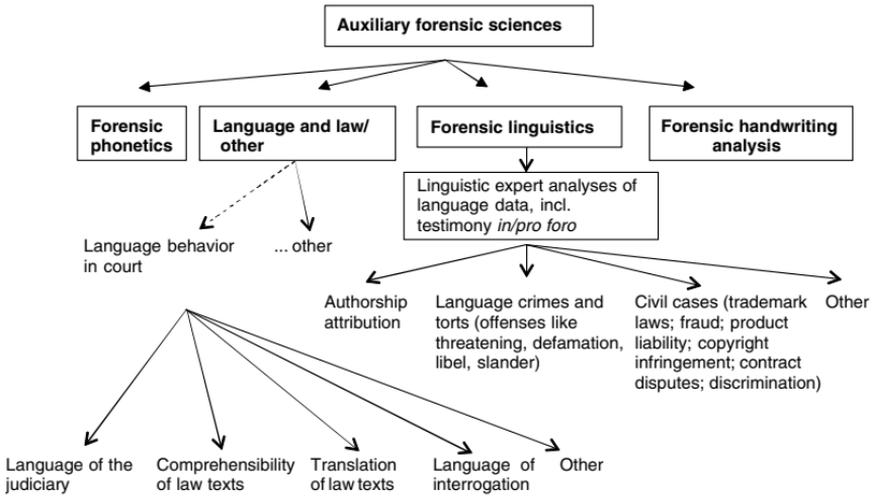


Figure 1.2 Forensic linguistics and neighboring sciences

This is merely a simplified overview, allowing much room for further specification, alternative terms, and structures.

Figure 1.1 does not seem to need further explanation. Straight arrows are used for what may be called a direct hierarchical affiliation of a particular field of research and/or practical work with a mother field. A dotted line with arrows in two directions illustrates a (more or less direct) reciprocal influence of a field on a higher level upon a field on a lower level and vice versa.

The various subfields – sociolinguistics, psycholinguistics and text linguistics – strongly influence the work done in the entire area of language and law. On the other hand, the spectrum of linguistic analysis undertaken in forensic linguistics may show some (in)direct effect on the field of theoretical linguistics, and also the fields of interdisciplinary and applied linguistics as a whole. An example would be the notion of “idiolect” and the analysis of varieties of a language in narrowing down an individual variety in authorship attribution.

In the following sections below, the structural similarities and common denominators of the auxiliary forensic sciences neighboring forensic linguistics will be discussed and, in an exemplaric fashion, also the structural differences and divergences in theory and method (between FL, FP, and FHA).

Considering the relatively large amount of literature on each of these elements, it is somewhat surprising that the literature on the interrelations,

structural similarities and differences of the three sciences (FL, FP, FHA) is rather small. It also seems more than accidental that, of all three sciences involved, scientists from handwriting analysis seem to be the ones that have been most interested in the interdisciplinary status and relations their science has with the other two (see, for example, Blum 1990; Hecker 1990; Michel 1982; 1990). Interestingly enough, forensic handwriting analysis is the science with the oldest and longest tradition as an auxiliary forensic science (for details, see Michel 1982).

There is no question that a look at the similarities and differences, particularly in terms of methodology, is of critical importance for the theory and practice of each of the three. The same holds for the more remote neighbors in some respects. The three sister sciences FL, FP and FHA deserve special attention for systematic and methodological reasons and, first of all, because of the almost trivial fact in forensic cases that, in real life, they frequently co-occur. In many cases in which I have given expert testimony, my work was undertaken upon the request of a scientist from FP or FHA. The close ties existing in practical real life applications are reason enough to take a closer look at the interrelations between the three sciences.

Interdisciplinarity: perspectives and postulates

This section discusses some general principles, perspectives and postulates that apply more or less to all auxiliary forensic sciences, particularly those concerned with the analysis of data of individual people and, more specifically, to the three sister sciences FP, FL and FHA. The overall approach is to elaborate on the common ground, the basic assumptions, precautions and concerns that are the same or similar.

It is a stereotype, or even a commonplace, in FL, FP and FHA that general postulates for interdisciplinarity are spelled out. At the same time, it is very rare that people working in one of the fields actually act according to this postulate and include systematic provisions for the other field(s). The general postulate is to little avail for an improvement of the situation. What is needed is a detailed exemplification and explanation as to why and how interdisciplinary relations have to be examined and described systematically.

In principle, linguists could have known this before, since there are several examples in the history of linguistics where previously strict dichotomies were shown to be detrimental rather than useful to the field. Halle's (1954) discussion is an example, showing that phonetics and phonology can profit considerably by envisaging each other's

research directions, main descriptive aims and the methods used. Another example is the even more short-sighted separation of the (interdisciplinary) study of language variation (in the old days) from systemic linguistics. Later research in sociolinguistics in the 1960s showed that it is possible, and necessary, to include a thorough syntactical, phonological and semantic analysis in work on language variation (for example, Labov 1966; 1970; 1972b; 1975).

To set up specified hypotheses and to give a detailed analysis of arguments is also necessary because of a widespread aversion of the (German) judiciary against “interdisciplinary experts”. This was justified in the old days, perhaps even in the 1950s, but is certainly outdated today. The pledge for an interdisciplinary exchange of ideas and systematic consultation does not imply that the established boundaries of sciences should be torn down. No experienced forensic linguist would act as a forensic handwriting expert or as a forensic phonetician if s/he does not have the appropriate training. An acceptable approach to interdisciplinarity requires that experts recognize their limits within and of their specialized field. What is important is that they also realize the possibilities and potentials of the other fields. If the data require consultation with scientists of another field, they must do so.

To sum up, one could say that the status of forensic linguistics, still somewhat in its infancy, does not result from the fact that there is too little (linguistic) specialization and too much interdisciplinarity when giving expert testimony. Occasionally the reverse holds: there is too much specialization combined with too little interdisciplinary orientation by linguists and, I suspect, also by the scientists of the other auxiliary forensic sciences.

Forensic phoneticians, for example, have long realized that an intensive implementation of sociolinguistic and psycholinguistic research hypotheses and methodological advances, including areas such as language attitudes, foreigner talk and many others, can be of great benefit to forensic phonetic analysis proper. On the other hand, it has been shown also that too narrow a restriction to the data of one's own discipline is, as a rule, detrimental rather than helpful in a forensic analysis.

It should be mentioned here that the term “interdisciplinarity” is a sweeping simplification. One would have to distinguish several layers and levels of interdisciplinarity. The most salient, in the context discussed here, are those between language sciences and law sciences. These have been discussed quite extensively in the literature (see Kniffka 1981; 1990a; 1998; 2000a; Levi 1982; Levi and Walker 1990; Shuy 1993a;

1998a; Solan 1993; Solan and Tiersma 2005; Tiersma 1999). Much less attention has been paid to the interdisciplinary and inter-field relations between FL and FP, FL and FHA, FL and forensic psychiatry, FL and forensic chemistry, including DNA analysis, however. The question of interdisciplinary relations between auxiliary forensic sciences certainly would be worth a separate systematic study.

How can the major perspectives and postulates of interdisciplinarity be described adequately? What exactly does “interdisciplinary perspective of forensic linguistics” mean (including the demands of practical cooperation in real life forensic cases)? The following observations can be made:

(1) As pointed out, a structural distinction between different layers, levels and types of interdisciplinary relations can be made (see Kniffka 1989: 224 ff.) between (a) real life application A and real life application B (German “Praxisfeld” A und “Praxisfeld” B); (b) between applied science A and applied science B; (c) between basic research A and basic research B.

Level (a) concerns the practical work, such as criminalistic investigation, and the collection of raw data by the linguist; (b) refers to the facts of applied linguistics (forensic or non-forensic) and applied science(s) of other disciplines; (c) this concerns linguistic and criminological basic research. Not only would one have to describe the horizontal relations on each level, but also the diagonal relationships of various kinds between levels (a), (b) and (c). The discussion below (and in other chapters in this book) gives ample illustration for this. Such a classification will, very likely, be of interest to practitioners. It may also help in a systematic scientific reflection to be able to recognize certain urgent desiderata much more quickly and effectively.

As mentioned before, there is an enormous amount of literature on language and law today. There is a comparatively small amount of research, however, on expert testimony and the intermediate and interdisciplinary co-occurrence of expert opinions in the various auxiliary forensic sciences. Documentation of forensic cases in which expert opinions of different sciences were given and of their outcome would be of prime importance.

(2) Interdisciplinary relations can also be distinguished in a different sense. It is possible to describe the relations between different forensic and non-forensic sciences and auxiliary sciences. It would be useful to specify the differences between and similarities of, for example, forensic linguistic expert opinions and expert opinions in employees’ evaluations, school certificates, medical and technical diagnostics and so on.

The interrelations between forensic linguistics and (general) linguistics do not need to be discussed because they are always present and of great concern to the forensic linguist (possibly, at some time in the future, also for the general linguist). It would definitely be helpful for forensic linguistics to elaborate on the differences and similarities between this field and, for example, forensic psychiatry on the one hand, and forensic linguistics and forensic phonetics on the other. In addition, the various kinds of “intra”-relationships between different sub-branches of forensic linguistics, such as authorship attribution and semantic analysis of defamation, would be of great importance. The underlying hypothesis of all this is that it is not only worthwhile, but also of great practical and systematic help to view interrelations in a more systematic and detailed way. This is not to infer that direct (simplistic) analogies between different fields could or should be made. On the contrary, the differences between the various sciences involved would be more clearly contrasted. Most certainly, the gain in methodological insight would be a valuable result. This in turn may be of considerable heuristic benefit in dealing with real life data. One such methodological or procedural similarity worth analyzing systematically is the process of giving expert testimony, and the several steps involved therein (as described in detail in Kniffka 1981; 1989, and also in some chapters in this volume).

There is no question that, with reference to general background knowledge, the steps that the forensic psychiatric expert takes when giving testimony in court could also be suitable for use by the forensic linguistic expert in terms of general methodological insights and vice versa.

(3) The need for an interdisciplinary perspective is, almost trivially, constituted by the fact that there are several different arenas where forensic linguistic topics and questions are discussed; such as (in the German system):

- (a) (Written) discussion with the judiciary and the judge ordering an expert opinion;
- (b) Discussion with lawyers;
- (c) Discussion *in foro* presenting an expert opinion;
- (d) Discussion in a class of law students;
- (e) Discussion in a (forensic) linguistics class;
- (f) Discussion at linguistic conferences;
- (g) Discussion at language and law conferences;
- (h) Discussion in legal journals;
- (i) Discussion in linguistic journals and, eventually, *nolens volens*;
- (j) Discussion that a case and the work of the experts of various kinds attracts in public opinion and in the media.

The point here is that discussion in any of these situations occurs more or less in isolation from the others. It requires particular effort to secure an exchange of opinions, or even to generate any kind of exchange of knowledge.

In this context it should be mentioned that, generally speaking, (German) linguists do not read legal journals such as *NJW*, *Archiv für Kriminologie* and so on. The judiciary and criminologists on the other hand, as a rule, do not read *JVLVB*, *Language*, *IJSL*, *Word*, *Language in Society* and so on. There is hard empirical evidence (also cited in the literature, see Kniffka 1981; Shuy 1993a; 1998a) that arguments of the judiciary, on the one hand, and linguists, on the other, would look very different if there were some kind of understanding of the main discussions in the journals of the other science. Linguists would have to change their own assessment, and would learn that assessment by the judiciary is a factor worth taking into consideration when appearing in court. The same holds, *mutatis mutandis*, for the judiciary.

Examples of this include the comments made by a former presiding judge of a higher German court concerning the general relationship of judges and forensic experts. They were made some 20 years ago, but seem to be of considerable importance even today:

It is certainly much easier to comply with an expert opinion on the basis of a reasonable justification than to try to refute it with conviction. So one [the judge], as a rule, complies with it, in particular since the highest [German] court's jurisdiction has, in fact, led to considerable restrictions on the principle of free evaluation of proof concerning expert opinions.

Thus, it is not surprising that, according to an empirical investigation, in 95 per cent of all cases, judges followed expert opinions without entering into a substantial discussion and critical evaluation of the expert opinion and, instead, resorted to empty phrases – in many instances, presumably, with a guilty conscience – that the court had supported the convincing argument of the expert by virtue of its own judgment. (Sendler 1986: 2909 (my translation))

Forensic expert opinions can generally be criticized by the judge under three conditions only: where the expert has based his analysis on incorrect facts, if s/he gives expert testimony in a random, non-factual and non-pertinent way, or if s/he has used the wrong methods of analysis. (Sendler *ibid.*: 2910, quoting F. Werner 1971: 316, *Recht und Gericht in unserer Zeit*)

Judges in particular have had quite some experience with scientists and expert opinions. One can only be surprised, time and again, by what is being offered as science, and what very diverging results scientists are able to produce on a seemingly strict scientific basis in a strictly scientific method of deduction. (Sendler *ibid.*: 2912)

It is incontestable that, in many areas, many judges are neither in a position to examine and control expert opinions in a scientific manner nor equipped to follow the scientific argument presented. Consider the complicated chemical analyses, measurements and calculations in the area of pollutant emission laws contained in expert opinions concerning questions of safety in the area of nuclear technology. (Sendler *ibid.*: 2909)

(Note: In this quote, all footnotes concerning further literature have been omitted.)

These statements are worth reading for linguists *and* legal practitioners for yet another reason: different levels of accessibility of the results of a science for laymen or people without education in a particular field is presupposed. Sendler (1986) does not quote forensic linguistic expert opinions in this context. It is quite possible that, in the view of the judiciary, these are generally considered more accessible than, for example, expert opinions in forensic chemistry and nuclear physics. A background in linguistics could help the judiciary understand that this impression is unjustified and misleading. In fact, linguistic and/or psychological expert opinions are no more easily accessible than forensic expert opinions in the natural sciences. Being a competent speaker of German in no way enables one to be a competent linguistic or forensic linguistic expert, in the same way that knowing how to read and write does not enable one to act as a handwriting expert in court. In short, all those participating should know more about the other sciences involved, in particular the science of law.

(4) Yet another complex dimension of interdisciplinarity that concerns general scientific postulates is the economy, simplicity and consistency of a scientific description – which hold for all sciences involved. What is important in this context is that a new science, such as forensic linguistics, can learn from older, established forensic sciences such as forensic phonetics, forensic handwriting analysis and the other forensic sciences.

(5) Another formal parallelism in almost all forensic sciences is the fact that the results are required to be given in established probability scales, which should help the court interpret the results of forensic expert opinions according to more or less exact criteria: the famous yet

problematic statements “with probability”, “with high probability”, “with probability bordering on certainty” and so on.

Some well-known general tendencies deserve special scientific attention across discipline boundaries; for example, the fact that novices in a particular science generally tend to overrate and overstate probabilities. There have recently been some new developments and suggestions in terms of additional specifications of probability scales in percentages in Germany; such as Köller *et al.* (2004): interestingly, a volume produced by FHA experts working in German police institutions (see p. 44ff.).

(6) There are other facts of common ground in methodology in the three forensic sister sciences (and possibly even other forensic sciences); for example, the need to work with features and quantifiable properties gathered in a valid and reliable scientific fashion, which applies to all forensic sciences in basically the same way. There may well be further specifications of this common ground; however, the general need to define quantifiable measuring of incidents, features, elements and categories of analysis applies to all (including the forensic) sciences.

Another question is the empirical investigation of intra-subject and inter-subject variation, which has to be considered in a particular expert opinion and also in linguistics in general. Michel (1982: 40) stated that, for handwriting analysis, stronger changes in handwriting are more especially noticeable in younger and older people. Middle-aged people, on the other hand, show a relatively great regularity and invariance of handwriting, unless specific exogenous or endogenous conditions cause distraction. One should ask whether there are any corresponding age-related similarities in linguistic behavior – a question that has been neglected in empirical (forensic) linguistics. It is plausible that young people (of post-pubescent juvenile age) show a considerable variation in linguistic behavior, as do people towards the end of their life, whereas the age span roughly from thirty to sixty years could be considered a time of considerable regularity and invariance of linguistic behavior. This may present a wide variation in different speaker populations according to education, profession, exposure to different language genres and exposure to communication standards. There may perhaps also be some kind of intra-speaker variation and developmental change in linguistic behavior in middle age of which we are as yet unaware. On the other hand, it seems possible – even necessary – that forensic linguistic attempts be made to define typical linguistic behavior for particular life cycles in all age ranges.

(7) If it is possible to define speaker-specific features of linguistic behavior in particular life cycles, it may be interesting to examine which

co-variants exist with intra-speaker variation of voice and voice changes, and also of the handwriting of a speaker. Rather than assume simple analogies among the three, it seems worth noting co-variation patterns that may help to consolidate the facts of one science with those of another, particularly since, in many real life forensic cases, data from all three sciences have to be taken into consideration. In addition, inter-speaker parallel structures could be analyzed.

(8) It may also be a useful element of interdisciplinary research to examine the extent to which a particular type of language behavior (such as disguise) can be stated for all three areas – voice, language and handwriting – and perhaps a typology of concomitant features of this type of behavior could be elaborated. In other words, one could ask which disguise of voice, if any, goes together with which disguise of language use (for example in syntax, lexical choice, using a repertoire of a foreign language) and of handwriting (for example in using the left hand by persons normally writing with their right hand and so on). There have been no linguistic, phonetic or handwriting analysis research projects to date that examine these facts in a systematic way. It would be tremendous progress indeed to obtain more reliable empirical data as to which forms of disguise usually occur together and which, as a rule, occur in complementary distribution. One would probably be able to set up continua within each science for which (not necessarily exactly parallel) findings could be found for voice, language and handwriting disguise in a (more or less) corresponding combination. These are heuristic assumptions that need extensive testing and validation. It is an empirical question whether such parallels exist and what forms they take. There are many other behavioral syndromes of types of disguise, which are worth being analyzed. One point of interest would be how easily disguise can be accomplished in voice, language and handwriting, and which differences occur. Furthermore, how long and how consistently can a disguise of voice, language and handwriting behavior be handled by a particular subject, and which categorical differences and parallels can be stated? Last but not least, how easily can this be described on solid scientific grounds?

(9) Yet another promising question with regard to disguise has to be analyzed from an interdisciplinary perspective. Which assessment should disguise be given in the final evaluation of an expert opinion, which should include a substantial documentation of those features of disguise that were easily detected by forensic phoneticians, linguists and handwriting analysts and those that were not? Based on a substantial corpus of the success rates for different types of disguise and their successful

analysis, one might even be able to set up a hierarchy of general difficulties of the analysis and, also, of the frequency of occurrence of disguise in different types of offenses. It might also be useful to look into which types of disguise and which overall constellations of disguised behavior occur in total in criminal and non-criminal types of communication texts.

To sum up, it seems that the similarities and divergences of the relevant data in voice, language and handwriting behavior are of great importance in determining the various levels, elements and functions of behavior in the three sciences, which details would go undetected if such an interdisciplinary perspective were not pursued.

(10) Finally, one must make clear the most basic and important justification, suggested by the practical data of real life cases, for an interdisciplinary perspective of forensic phonetics, forensic linguistics and forensic handwriting analysis. In many cases involving incriminating texts – such as anonymous extortion letters, threatening letters, defamation letters and so on – data are required to be examined by all three sciences. In several extortion cases I worked on, there were written data that required analysis by a forensic linguist, phone conversations that required analysis by a forensic phonetician, and hand-written data that required analysis by a forensic handwriting expert and document examiner. Frequently there was a successful cooperation with the outcome that, indirectly, the data of one science were confirmed by the findings of another.

Structural similarities and differences between forensic phonetics and forensic linguistics

Structural differences

In this section, the contrast between two closely related forensic sciences, FP and FL, is described. Differences between and similarities of these two sciences are discussed. However, to refer to “differences” and “similarities” is somewhat imprecise. It would be more appropriate to discuss the different and distinguishing aspects viewed from the perspective of FL: this view would then have to be supplemented by a perspective from FP.

It would also add to the adequacy and validity of the description, if the *tertium comparationis* of differences and similarities were stated for each of the smaller elements of central questions, methodological maxims, features and measuring categories, which is beyond the scope of this chapter. Similarly, a discussion of basic research and the general status of FL and FP as sciences cannot be offered. The particular contrastive aspect

chosen here is forensic phonetic speaker recognition and forensic linguistic anonymous authorship analysis. It is analyzed in relation to the practical work that forensic phoneticians and forensic linguists perform when giving expert testimony. It also seems that this perspective of the real life ingredients of both fields has been more neglected to date than the theoretical aspects.

What are the features distinguishing FP and FL (among others)?

(1) A difference of prime importance is to be noted in the purpose and the data investigated by both sciences: spoken language data, voice data of a speaker as an object of phonetics with the aim of speaker recognition and/or speaker identification in recordings or directly perceived by hearers, are iconic and symbolic data. Linguistic data, such as written texts, are only symbolic data. This is one of the sources, if not the main criterion, that makes forensic phonetic data empirically "harder" than forensic linguistic data.

(2) There is a field known as instrumental phonetics, but none known as instrumental linguistics. Linguists do not have a comparable arsenal of technical facilities, support and aids at their disposal, such as spectrographs, X-ray and palatograms of sound production and so on. There are no "textographs" with which one could codify, analyze and measure linguistic behavior and textual data in a sufficiently precise way. The instruments that do exist in FL (such as computers with statistics programs, concordance and collocation programs, corpora and so on) are by their very nature somewhat restricted, due to the ingredients of the observed topic and also the range of applications. It is not yet possible, for example, to quantify elements of different types of meaning in an adequate and precise manner, let alone compare the data with those of other types and kinds of texts. The same holds for some syntacto-semantic categories, and even more so for text linguistic and text type-specific pragmatic elements. Measuring forensic linguistic elements, features and entities is, generally speaking, much more analytically demanding, difficult and extensive by its very nature.

(3) The last point represents another real life difference between FL and FP. FL expert opinions concerning anonymous authorship of incriminating texts (such as extortion, threatening and defamation letters) are, as a rule, much more time consuming than either FP expert opinions on speaker recognition or FHA expert opinions. Linguists have to deal with a relatively larger and more heterogeneous amount of textual data in incriminating and comparison letters than forensic phoneticians deal with, *mutatis mutandis*, in a voice analysis of incriminating phone calls

and comparison data. In principle, this also seems to hold for the analysis of samples of handwriting data by the FHA expert. For both sciences, FP and FHA, a sufficient amount of literature is available dealing with these questions.

What matters is that forensic linguists, generally speaking, need larger sets of data of naturalistic, non-experimental texts and much larger text corpora than the two other sciences. As a result, FL expert opinions are, generally speaking, usually longer than FP and FHA expert opinions. Judging from many hundreds of FL expert opinions that I have seen or written myself, I tend to think that FL expert opinions are rarely shorter than about 50 to 80 A4 pages in narrow spacing. Generally speaking, FP and FHA expert opinions tend to be much shorter.

Comments from phoneticians and handwriting experts confirm that, in most cases in the two sciences, their expert opinions take relatively less time than an FL expert opinion. I have never been able to work out and formulate an FL expert opinion in less than a minimum of one week. Not infrequently, several weeks or months have been necessary. Invariably those who request an expert opinion have extremely unrealistic expectations with regard to how long such an opinion takes to construct. It is almost stereotypical in a request for an FL expert opinion that the judiciary, police or other authorities in Germany ask for a “swift information on a tendency” or “a preliminary notice concerning the result”, including even the exact percentage of probability of the results of the expert opinion. The stereotypical response of the forensic linguist in such situations is that no reliable results can be reached in such a short amount of time.

(4) A further difference can be noted concerning the degree of general acceptance of results and the establishment of FP and FHA on the one hand, and FL on the other. In all the investigative police institutions in Germany – in particular the Bundeskriminalamt and the Landeskriminalämter, the federal and the state criminal investigation institutions – FHA and FP are staffed considerably better than FL, and have a longer working tradition. Only the BKA employs forensic linguists, of whom at present there are two. German academia shows a similar picture. Until very recently, no German university had a fixed program for teaching FL, let alone offered an MA in this field. The MA and PhD theses with an FL topic were usually undertaken in departments of Linguistics or German. Very few colleagues in applied linguistics work as linguistic experts in courts to date, and almost none of these colleagues dedicate their major research activities to the field of FL. As with the question of general acceptance in court, there seems to be a lesser degree of ignorance

of the fields of FP and FHA than of FL within and outside of German universities. This means that people are even less aware of FL than FP.

The differences between FL, FP and FHA, in terms of the practical work of giving expert testimony and concerning the acceptance in academia and institutionalization in the investigating authorities, may appear trivial and irrelevant at first sight. There is sufficient evidence, however, that they are of great importance for an adequate discussion of the theoretical and practical constituents of the forensic sciences in question. Some of the differences discussed have even more in-depth and complex differences as a consequence.

(5) The differences in purpose and data mentioned in (1) suggest that there is a somewhat binary key question at stake in FP and FHA respectively. Is the voice of the incriminating data identical to that of the comparison data or not? Is the recorded voice sample of a suspect from the same speaker as the recorded extortion call or not? Is the handwriting in an incriminating document identical with the handwriting in the comparison data? Is it from the same or a different writer?

There is a broad spectrum of possible combinations, manipulations, additional (later) changes, insertions and omissions in an authentic array of data. Questions of authenticity of recordings of a voice and of handwriting samples are extensively discussed in both fields. In FL authorship analysis, there is also a binary key question: that of whether the anonymous incriminating texts (“Tatschriften” (TS)) and the comparison texts (“Vergleichsschriften” (VS)) have been written by one and the same author and/or writer/typist (this distinction representing yet another complication). Are they products of different authors and writers/typists? Is the author of the TS and the VS identical or not? Is the writer/typist of the TS and the VS identical or not?

However, expert analyses of real life cases under real life conditions very frequently do not deal with binary questions of this kind. There is a much more complex continuum of involvements of one or more authors and writers. Generally speaking, a scalar co-authorship of a nature yet to be determined can be taken as the most general and frequent (unmarked) constellation. Whereas simultaneous multiple origination of a product is a marked feature in itself in the analysis of voice and handwriting – for example, if someone directs someone else’s hand when writing a text – incriminating texts as objects of FL, as a rule, do not represent a marked constellation in terms of multiple authorship. One cannot extract features in typed texts that indicate whether they have been written by a single author and writer, one author and a different writer, or several authors working simultaneously as a team on a

text as a whole or on different parts of a text in different constellations. Neither can one analyze for features of a writer who, for example, initially made handwritten notes and someone else who typed the final version. Nor can one detect whether a typed text has been dictated at gunpoint and, possibly, has been changed knowingly or unknowingly by the writer/typist; whether there are several (chronologically different) versions of a text, or a combination of portions of a text produced at different times and by different authors that may have been put together later by one or more authors/text editors.

This is a strongly simplified illustration of real life dimensions of variation that one encounters with written text products or that have to be taken into consideration in authorship attribution. In an even more basic simplification, one can say that FP and FHA do not seem to have a situation directly corresponding to (simultaneous) multiple authorship in FL. They certainly do not have to view this as a frequently occurring type of everyday textual data constellation. Due to the problems that cases concerning multiple authorship normally entail, I have resorted to framing the question relating to author identity of the TS and the VS as: "Is the author of the comparison data texts in one way or another involved in the authorship of the incriminating texts?"

(6) There is yet another challenging difference between FP and FL. It concerns the role that the judgment of laymen in speaker recognition and identification can play in FP. There is considerable documentation of cases in which laymen, as long-time acquaintances of a person, were able to recognize the voice of a speaker with an amazingly high percentage of accuracy (for example, the identification of the broadcast voice of the murderer of two policemen in Hannoversch Münden, Germany, by close friends). It seems that there is no corresponding fact in forensic linguistics: laymen are usually *not* able to recognize/identify the text products and also hand-written material of a person they know well to an equally successful degree. There are, however, some notable exceptions to this; for example, the identification of the Unabomber's written manifesto by his brother (personal communication of Roger Shuy). Judgments by laymen concerning authorship identity of anonymous incriminating and comparison textual data, as a rule, do not achieve the same degree of success. At least, there are not many actual cases described in the forensic literature in which laypersons' judgments of anonymous texts or anonymous handwriting samples have led to similarly impressive investigative results (for a more in-depth discussion see the literature on forensic phonetics).

In summary, lay people seem to be able to recognize and identify anonymous voice samples (broadcast in the media) with greater validity and reliability when relating them to a speaker they have known well for a long time, than they are able to relate anonymous text products or handwriting samples (for example, published in a newspaper) to an author or a writer that they know well.

There is a considerable amount of literature on features and factors responsible for the relatively high level of achievement of lay people to identify phonetic samples as opposed to their relatively low level of ability in FL and FHA. To date, a thorough empirical, detailed and scientifically satisfying explanation of this fact, and also the question as to which feature configurations are responsible for this, has not been offered. In this regard, there is an urgent need for interdisciplinary cooperation. Further study is required to identify how forensic phoneticians' discoveries and explanations can be made accessible in forensic linguistics; that is, how laypersons' judgments in matters of forensic linguistic authorship analysis can best be supported. At any rate, no simple analogies in one way or the other are possible.

(7) Undoubtedly, forensic phoneticians have a more powerful and empirically sound arsenal of instruments and methodology of analysis at their disposal than forensic linguists. As in the case of laypersons' judgments, there may be quite a few other respects in which forensic linguists could learn from forensic phoneticians. Doherty and Hollien (1978: 1ff.) showed that several parameters/vectors (such as long term power spectra, fundamental frequency, speaking time and so on) worked well in producing significant results for normal speech data created under ideal lab conditions. Not so, however, when the speech data were distorted speech; that is, when they implied distortion and/or disguise of speech. It would be worth checking whether this, *mutatis mutandis*, could also be a useful perspective for forensic linguistic analysis, leaving open the question as to whether they represent a structural similarity or a difference.

(8) A structural parallel/similarity between FP and FL can probably be justified in an even more general sense. Taking the résumé by Doherty and Hollien (1978: 7) as a starting point, one could investigate how and to what extent it applies in basically the same way. "[...] It should be noted that the exact number of parameters/ vectors necessary to develop an effective speaker identification procedure is not known" (*ibid.*). This may no longer apply to FP today: I leave the decision to colleagues in FP. For FL, it is undoubtedly true. The only thing we know for sure is that we have not as yet identified a sufficient number of parameters and

vectors that would enable us to give a full-fledged, scientifically safe description of authorship attribution.

The analysis of speech distortions as discussed by Doherty and Hollien (1978) may also be taken as an incentive for forensic linguists to look into text classes and text types, and the text type-specific distribution of syndromes of errors and mistakes more closely. This may help to identify the results of various stress factors and kinds of disguises of authorship on more solid grounds.

To sum up, it is necessary and worthwhile to examine in each particular instance whether the results, methodology and overall approach of another forensic science can provide incentives and challenges for one's own. All existing differences and asymmetries between the two sciences FL and FP taken into account, and rejecting *ad hoc*, naïve and global analogies between the two fields, one should still be looking for a fruitful cooperation and exchange of ideas. The gain in methodological insights is usually much greater than any possible loss.

Structural similarities of FP and FL with regard to practical expert testimony

The (incomplete) list of structural differences given above could create the false impression that there are hardly any structural similarities and common ground in the forensic work of FP and FL. At least as far as the analysis of anonymous voices and text products in real life is concerned, the opposite is true. There is enormous similarity in general layout, methodology and practical work. It may be adequate to say that the differences stated primarily concern the general status of FP and FL as *sciences*, and that the structural similarities are revealed more clearly in the light of the every day professional *practical* work of forensic phoneticians and forensic linguists giving expert testimony. In the following, a very incomplete, subjective and unsystematic list of structural similarities, coincidences and elementary parallels of FP and FL is given.

It should be noted that this list does not intend to amplify the structural similarities, elementary parallels and coincidences of the three fields, or to state that the similarities outweigh the differences. The intention is to state some of the structural similarities explicitly to ensure that an interdisciplinary exchange and increase in interdisciplinary awareness is worthwhile for all the three sciences involved.

(1) From a scientific point of view, there is no one simple question in most forensic cases as posed by the (lay)persons who request the expert opinion (for example, "Who is the author?"; "With which probability is

X the author?"). One of the jobs of forensic phoneticians, forensic linguists and forensic handwriting experts is to perform the twofold "translation process" from the laypersons' question into an operationalized scientific set of questions, and to re-translate the results back into a manner that lay people can understand. The overall success of an expert opinion depends on the scientific and professional stringency of the analysis *and* on the translation of the results into accessible language. It should be noted here that the fact that there is not just one question, but a larger set of operationalized scientific questions, applies to each particular area within FL; such as authorship attribution, text pragmatic analysis of what is said, clues to what is meant and so on. It seems that this also holds, in principle, in FP and FHA.

(2) Another inter-science invariant of significant methodological importance is the fact that isolated single features do not ever allow valid statements concerning speaker-specific behavior. What counts is that there are bundles and configurations of dynamic features (see Kniffka 1981) that can be stated for a larger set of variables which, in principle, vary independently of each other; such as orthographic behavior, syntactic behavior, paragraph division of texts, lexical choice, punctuation, register, code switching, variants of word formation, jargon of various fields and so on. If and how the four basic dimensions of variation – orthographic behavior, grammatical (including lexical) behavior, stylistic text pragmatic behavior, and text phenotypical or lay-out behavior – have structural correspondences in FP and FHA is a matter to be discussed by representatives of the two sciences.

(3) It seems that in the three sciences, in particular in FP and FL, data of language system *and* language usage have to be taken into account. This is an assumption generally accepted in FP and in FL. There is an encompassing sociolinguistic perspective of analysis to be chosen, including data of sociolinguistics and the sociology of language in the technical sense.

(4) The procedures of scientific analysis are basically the same in all three fields as far as individual communicative behavior is concerned. Data of incriminating and of comparison texts are to be sampled and analyzed, as far as possible in an objective methodology in quantifiable terms. Unavoidably, some subjective interpretations come into play in the three sciences. In all three, the difference between description and explanation and/or interpretation of the results must be distinguished. The final evaluation of the data and the expert's recommendation to the people who requested the expert opinion should follow a general standardized procedure and canonical formulation. The scales used for the

evaluation of the data in terms of probability must be standardized and written in a form that laymen and non-scientists understand.

A step forward has been made by the recent publication by handwriting experts Köller *et al.* (2004), who give a thorough and in-depth discussion of the problems involved in probability statements and scales in FHA expert opinions. It would seem that these can, in principle, be applied with almost no changes to the field of FL.

(5) Experts in all forensic sciences are frequently asked by people requesting an expert opinion, as well as by the general public, about the possibilities and limitations of their fields. It is essential to systematize and reflect the limitations much more openly and thoroughly than has been done to date. This might cause the judiciary and others not to request an expert opinion, but that seems a lesser evil than creating exaggerated expectations, which unfortunately frequently happens in Germany.

(6) There is yet another common ground between the three forensic sister sciences and, perhaps, forensic sciences altogether. Without cynicism, one can say that mistakes, errors and blunders of methodology, conceptualization and professional standards occur in all forensic sciences. Some of those that occur more frequently (and which are of most relevance in the three sciences) are:

- (a) Assuming authenticity of the comparison data without obtaining or being able to obtain detailed information about them;
- (b) Non-verification of quotations and representations of texts of incriminating and comparison data (for example, in court transcripts), causing artefacts and failures to be created;
- (c) Premature generalizations of a hypothesis or a tendency (frequently requested by the people who order the expert opinion) on a non-quantitative basis;
- (d) Non-explication of deviations and exceptions (for example, of a tendency) and/or non-discussion of alternative hypotheses;
- (e) Confusion of the sample and its size with the author's behavior in total, or even the author's or speaker's personality as a whole (cf. Heike 1990);
- (f) Non-scientific or non-professional misinterpretation of features and feature values because of theoretical, methodological and other deficits;
- (g) Unrealistic upgrading of the probability level(s) in the final evaluation (a mistake frequently made by beginners who portray greater certainty of the results than they actually have);
- (h) Using non-addressee oriented formulations and special jargon in court.

(7) An almost incidental general requirement of forensic sciences is the necessity for experts to have a solid scientific training in the particular field, meaning a thorough study and professional experience in forensic phonetics, linguistics, handwriting analysis, psychiatry/psychology and so on. There is no legitimization for someone to act as an FL expert based on his/her command of his/her mother tongue alone. Without academic and scientific training, one cannot give expert testimony. The same holds for FP and FHA and all other auxiliary forensic sciences.

(8) All auxiliary forensic sciences have to state their results in a way that lay people (including the judiciary) can understand. Forensic experts also must refrain from trying to do the judge's job. According to German law, the judge is free to evaluate expert opinions in any way s/he sees fit (the so-called "Prinzip der freien Beweiswürdigung von Sachverständigenbeweisen").

(9) The obligation to supply general information about the basic possibilities and limitations of one's field and the limitations applying in the particular case should be addressed to the judiciary and also to colleagues in the field. Large sections of academic linguistics have *no* solid and detailed knowledge of the applications of the field. Theoretical and applied linguistics are, perhaps particularly in the German academic system, frequently worlds apart from each other and almost dogmatically divided. Many pure/theoretical linguists are not even aware of the fact that there *is* something like FL. Similarly, many people in psychology may not know that there is an established discipline for FHA ("Gerichtliche Schriftvergleichung"), and still keep confusing it with (and are actually using the term) "graphology". There is an urgent need to establish firmly the field of applications of linguistics to real life, including law, in the curriculum of German universities.

(10) In FL in Germany and, possibly to a lesser degree, in the two sister sciences, there is frequently merely guesswork as to which classes and kinds of properties of texts, linguistic features and values of features are of critical importance and worth being operationalized for authorship attribution: this general desideratum should also be met by people working in FP and FHA. It is not justifiable to come up with *ad hoc* recommendations. An "expert" X suggests that causal conjunctions in German, such as *da* and *weil*, should be looked at, because "they represent different types of speakers" (sic); another "expert" Y says that word order is "where it's at" for authorship attribution. Yet another "expert" Z says that syntax does not amount to anything, but that the lexicon does. Yet another "expert" might proclaim the reverse opinion. This is non-scientific gossip. In addition, it is redundant since each real life case is

different from all others, and since all features of linguistic behavior and adjacent areas, such as orthographic and phenotypic behavior, also have to be taken into consideration. Which of these features are of critical importance and relevant in a particular case is an empirical question.

(11) For all forensic sciences, a general postulate seems to be applicable without major differences: there should be an official record of expert opinions given in FL, FP, FHA and other fields, documenting which effect an expert opinion had in a particular real life case of law (cf. Kniffka 1990a; 1996a). This would presuppose, however, having documentation of all expert opinions given in a field such as FL to date, including the major errors and mistakes that were made, and faulty (professionally deficient) expert opinions (“Fehlgutachten”), given by people (inside and) outside the experts’ fields of FL, FP and FHA. It seems that a more detailed account of such data, the formulation of professional standards for FL experts, and a thorough discussion of the field in the academic public can help reduce the number of deficient linguistic expert opinions considerably.

Part II

Forensic Linguistic Research in Germany in the Past

This page intentionally left blank

2

The Linguist as an Expert Witness in German Courts. A View from the 1970s

Orientation

This is, in part, a translation of some sections of my 1981 paper “Der Linguist als Gutachter bei Gericht” (Kniffka 1981) (“The Linguist as an Expert Witness in Court”), which is the first attempt nationally and internationally, as far as I know, to give a systematic account of FL expert testimony in and for German courts. It appeared in a German Festschrift, and has mostly been overlooked, in particular by the English-speaking world. I have added data from some of the cases in which I gave expert testimony in the 1970s, other than those characterized briefly in the Introduction and in the third section of this chapter. The main concern is to document what was done in Germany in the early stages of forensic linguistics. This period of FL activity, roughly from 1972 to 1990, is, at times, not paid proper attention.

This chapter also illustrates, if only implicitly, where the field has gone since that time in authorship attribution (henceforth AA) and other areas of linguistic expert testimony. The evaluation is left to the reader. Much of what was stated some 25 years ago is still of critical importance to the forensic linguist’s work today. For the translated sections, the original (German) text of 1981 has been mostly left unchanged. By adding summaries of other cases, it was entirely rewritten, however.

There are three subsections. The first concerns the relationship of “applied” and “theoretical” or “pure” (branches of) linguistics. The second describes the basic classification given in the 1981 paper of FL expert opinions on the basis of the cases the author had been working

on up to 1979. It also focuses on the fact that both branches of linguistic expert opinions – (a) how a text/an utterance is meant, and (b) who the author of an anonymous text is – each represent continua rather than clear-cut different areas of questions (and cases) in one direction only. The third subsection elaborates on the methodological implications of AA analysis more precisely. As sketchy as they may be, the main findings of 1981 still seem to hold today.

Applied linguistics and theoretical linguistics

This chapter is oriented toward practical examples of real life linguistic expert opinions given *in/pro foro* in Germany by the author in the 1970s (1972–80).

The field of applied linguistics is not helped much by declarations of intent such as “This is an interesting matter one should look into more closely ...”. Progress can be reached, step by step, through an open-minded discussion of theory and method in relation to data oriented applications on neighboring sciences and practical fields (“Praxisfelder”) that are constituted by them. Theorizing far away from or above practical work is somewhat redundant in an applied discipline. Not so, however, is a systematic reflection of the basic underlying assumptions and methodological postulates.

Within the field of linguistics – possibly more so within Germany than beyond – there is an obvious and, at the same time, largely unjustifiable distinction between “pure” or “theoretical” sciences on the one hand and “applied” sciences on the other. “Pure”/“theoretical” linguists take pains to give substantial semantic analyses of words and sentences, and consider this to be a salient part of their work. Only rarely do they deal with the everyday “lowlands” of practical empirical semantic analysis as, for example, when the meaning of a word and its use in everyday standard German is a disputed matter in a court case. There is no justification for theoretical linguists to shy away from such jobs or to criticize applied linguists for neglecting the theoretical and systematic interpretation and validation of their methods and results.

Taking into consideration the fact that specialization and division of labor is necessary in all fields, including linguistics, some incidental (but nonetheless sometimes unobserved) postulates seem imperative. In an applied discipline, it is impossible to undertake research without the corresponding pure or theoretical branches. Applied linguists must also be pure or theoretical (general) linguists, at least to the extent that they use theoretical concepts, terms, testing methods and instruments of

analysis. A mere heuristic linguistic analysis, using only common sense and the blue sky above, is non-scientific. On the other hand, mere theorizing without any prospect of empirical validation, without any reference to natural language behavior in everyday communication situations, may be called scientific but has little to do with linguistics as an empirical study of natural language behavior. It is almost absurd that in a research seminar people take great pains to discuss meaning descriptions, but that in a context in which the exact description and analysis of meaning is of significantly greater and more critical importance, both as far as the empirical construct validation and the theoretical foundation is concerned, this effort is, generally speaking, not made.

What can applied linguistics do about these misunderstandings of the field? The first, and perhaps the most important, aim is to persuade so-called “theoretical” linguists that – to a far greater degree than they appear to realize – our field is being approached for help by other academic fields and their practitioners. One such concern is the area of “language and law”. Furthermore, they must be persuaded that applied linguistic tasks are, in principle, in no conceivable way less scientific than, for example, the field known as “theory of grammar”.

Second, people working in the wide area of law need to be informed in a detailed and exemplaric way about what linguists do and how linguistics can, as an auxiliary science, be of great relevance and benefit to their work.

Interdisciplinary status of “forensic linguistic” work and the overall aim of linguistic expert opinions *in/pro foro*

Any classification of FL expert opinions starts with the practical questions of actual concern to the people who have requested them. Questions such as: Who requests a linguistic expert opinion, why and when? To what ends can a linguistic expert opinion be used (or not be used), generally speaking? Which kinds of expert opinions can be distinguished on the basis of actual requests addressed to the FL expert?

In order to reach helpful answers to these questions, a rather encompassing inquiry is necessary. Here, a few exemplaric comments and examples must suffice. Also, the overview of the types of linguistic expert opinions (see Introduction p. 4) should be consulted (note that the cases referred to below are different from those on pp. 62–8).

For FL expert opinions, very broadly, at least two classes can be distinguished:

(1) FL expert opinions that concern the analysis and evaluation of a spoken or written speech event or text. In such cases, the key questions are: What is meant? How is a text meant? How is a text to be understood? This can imply or presuppose a question of the actual form of that text. In some cases it is not even clear what is said, since part of the text may be missing; part of it may be illegible; the text, as such, is ungrammatical, or it is ambiguous in syntactical and lexical ways, so that there are several ways of interpreting it.

(2) FL expert opinions that concern the analysis of the authorship of an anonymous or partly anonymous utterance/text. The major question here is: Who has authored (and possibly with whom have they authored) the text in question?

The first class we call “Gutachten zum Verständnisnachweis” (‘FL expert opinions for a proof of understanding’). The second class we call “Gutachten zum Autorschaftsnachweis” (‘FL expert opinions for authorship attribution’).

Case A: Analyzing the sense of a text passage

Here is a somewhat complex example of a linguistic expert opinion on proof of understanding.

In an edition of a trade journal by publisher Y (“Informationsbrief eines Branchenpressedienstes”), the following text passage appeared: *Schon heute dürfte klar sein, daß die Märkteanstrengungen nicht nachlassen werden, Markenprodukte führen zu wollen, die ihnen auf legalem Wege verwehrt sind.* (‘It should be clear by now that the attempts of wholesalers to get hold of brand name products that they are legally forbidden to obtain will not cease.’)

In previous editions of the trade journal published by Y, the restriction of the distribution of brand-name products has been discussed in detail, including the legality of the refusal of brand name producers to have their products distributed to a wholesaler.

The juridical context that the linguistic expert opinion faces is: Wholesaler X interprets the passage, cited above in context, as being likely to give an uninfluenced average reader the impression that X had acquired brand-name products illegally. X claims that this reading (‘X has

acquired brand-name products illegally') is a false claim, and that it is the only possible understanding of the text passage. It violates the laws of fair competition. Publisher Y is considered a competitor by X (and by some courts), which means that stricter laws apply than those for press products. X sends a written warning to Y not to repeat this claim and requests a legal undertaking from him in this respect. Y sends X the requested undertaking and explains that X's reading of the passage was not the reading they had intended. Y states: *damit sollte keinesfalls gesagt werden, daß der Einkauf von ... Ware nicht legal war, sondern daß es unserer Meinung nach vertriebspolitisch für diese Hersteller legal ist, sich ihre Abnehmer selbst auszuwählen, also z.B. X nicht zu beliefern.* ('this statement was not intended to say that buying the brand-name products was illegal. Rather, that it was legal for the brand-name producers to distribute their products to customers of their own choice and that it was legal not to distribute them to X.' (abbreviated translation)). Y states, in short, that X is 'trying to buy brand-name products that they are forbidden to obtain legally, the refusal of distribution of products having been stated as legal by a court'.

X files charges against Y for the lawyer's fee for the written warning. Y is only obliged to pay the costs if it is Y's own fault that caused the expense. Blame can be attributed to Y if X's interpretation ('X has received brand-name products illegally') is the only possible interpretation and if Y's reading is considered unfeasible.

Y orders a linguistic expert opinion (original wording): *wie der unbefangene Leser die Textpassage versteht, insbesondere, ob er sie im Sinne von X verstehen muß oder sie (mit gleicher Wahrscheinlichkeit) im Sinne von Y verstehen kann* ('how the uninfluenced average reader understands this text section, in particular, whether he has to understand it according to X's interpretation, or whether there is an equal probability that the text section could also be understood according to Y's interpretation').

The linguistic expert opinion comes to the conclusion (translated and abbreviated): 'The interpretation claimed by X cannot be considered to be the only one possible. Y's interpretation is also possible, even if it is slovenly and imprecise in expression. Both interpretations are not atypical for the general manner of expression used in this particular text product (trade journal).'

The district court dismisses X's action against Y. Y files an appeal with the next higher court. In essence, X's lawyers argue that 'this text was addressed to the casual reader' ("der flüchtige Leser"), as opposed to the linguist (supposedly the "careful reader"), and that the text analysis of the latter was not relevant. They continue that the average superficial

reader did not need to be a linguist to find out the meaning of this text and that it was clear from the linguistic context that the wholesalers, in which X was included, had in fact previously acquired brand name products by illegal means.

The superior court (Oberlandesgericht), the court of appeal, dismissed the action, thereby following the evaluation made by the linguistic expert opinion. X had to pay the costs of the legal procedure.

The case shows that linguistic expert opinions can be of great help in the interpretation of legal and business texts. It is also of interest as one (!) example of the ways that the judiciary (in some cases) views the function of linguistic expert opinions. It is not intended here to make a generalization about the judiciary as a whole. The assumption of the judiciary being a homogenous group is as unrealistic as that of the linguists.

In case A, the meaning of a text section as a whole is at stake. The question is what that particular use of language X implies concerning the general usage of the German language by the people communicating. The object of the linguistic expert opinion in this case is to make a scientifically sound statement about the two interpretations of a sentence; more exactly, whether there are linguistic arguments for the fact that only one reading is possible and that any other is not. Case A also illustrates, in comparison to the case referred to in Chapter 5 (the status of the word *Konkubine* ('concubine') in German in 1972), that linguistic expert opinions concerning proofs of understanding do not cover only one type of question as to what is intended. It ranges from the meaning of one particular word in a given linguistic and situational context to the meaning and the sense of an entire utterance and text paragraph, including the text pragmatic status as a whole. In other words, it is not merely the denotative referential semantics of an utterance or a text that is disputed here. In addition to this, the text pragmatic status and the communication act as such are to be taken into consideration, and are the objects of the linguistic expert opinion and the legal procedure. Not only what is said and communicated to someone, but also the fact that something is communicated in a particular, possibly illegal, way is at stake. This means that linguists are not only asked to help analyze the denotative meaning of an utterance. The meaning is undisputed and juridically clear. What is not clear is the question as to what it means if such an utterance or text is made public, stated repeatedly and so on. If the communicative status were clear, there would hardly be any legal procedures in the first place; neither would there be a need to order a linguistic expert opinion.

Case B presents a slightly different scenario. Here also, not only is the meaning of one particular word in dispute, but also the text pragmatic status of a text passage as a whole.

Case B: A press correction or a repetition of the original statement?

This case concerns the old practice that, in press corrections, the correction itself is preceded by a repetition of the original message, although the original message is being corrected. In this situation, in essence, the disputed statement is being repeated rather than corrected. There are many humorous examples of the type: *Meine kürzlich geäußerte Feststellung, dass Frau X Beine wie das Brandenburger Tor hat, ziehe ich hiermit mit großem Bedauern zurück.* ('My statement made the other day that Ms X has legs like the Brandenburg Gate I herewith retract with great regret.')

In case B, publisher Y writes about company X that it has stopped to distribute a particular sales model, and that the financial authorities had refused to recognize this model formally. X orders an injunction that Y is not allowed to repeat this statement. In the next edition, Y varies the topic by simply adding that X has ordered Y by injunction not to say anything negative about the model. This is also met by an injunction from X.

In a third edition, Y states all these facts together; that X had ordered an injunction forbidding Y to say anything negative about the sales model, which had been reported before, adding that, on the request of X, they wish to clarify that the injunction only forbade them from publishing the statement that X had ceased distribution according to the old sales model and that the financial authorities had refused to recognize this model.

X presses charges against Y, saying that rather than a correction it was a repetition of the original statement, violating the injunction not to repeat the statement. Y, on the other hand, argues that it meets the standard of a press correction in the classical form, which, in order to make a linguistic reference, had to repeat the original statement for the reader.

The linguistic expert opinion requested by Y is to help in clarifying whether it was indeed a repetition of the original statement and, if so, whether it was intentional (which the court has to examine according to the injunction). The trial court decides in favor of Y, the appeals court in favor of X. A more detailed account of this rather complex case cannot be given here.

One can summarize that linguistic expert opinions concerning proof of understanding can refer to a large variety of legal issues and sections of law. In case A, the court has to decide whether there is some form of illegal competition (referred to in section 683 BGB (German Civil Code) in connection with section 670 BGB). In case B, a legal problem area represented by the terms “injunction”, “revocation” and “press correction” (section “Widerruf” 1004 BGB) is given. Oversimplified, it can be said that, in these cases, legal obligations are spelled out for linguistic behavior after (false) statements and claims have been made. Linguistically, one could say that the matter to be analyzed by the linguistic expert opinion is itself a juridical matter: the object of the linguistic expert opinion is also the object of the juridical procedure and the verdict. This is another distinction from linguistic expert opinions in the realm of AA.

Case C: Disputed authorship of an article in the press

In an Austrian political magazine, two texts appear. By text type, they would be closest to commentaries. Both have supposedly the same author, whose name appears underneath the two texts. The author himself designates one of the articles as authentically his, but denies authorship of the second text and says it was in fact authored by the editor of the magazine, at his own discretion and without the author’s consent and knowledge. He states in writing that he has nothing to do with the authorship of the second article. This is confirmed by an affidavit taken under oath by the editor of the magazine, who died in the meantime. The affidavit confirms that it was not the author whose name appears in the magazine but the editor himself that had in fact written the second article. Both articles display the same (right-wing) political convictions. They do not share any striking linguistic similarities, coincidences or features, however. Nor do they have any characteristic or otherwise remarkable errors, mistakes or deviations from each other. The author orders a linguistic expert opinion to obtain scientific proof that he did not author the second article.

The result of the linguistic expert opinion is, in brief, that a *non-liquet* applies: on the basis of the data and with the methodological instruments available to linguistics, no safe analysis can be presented to state in a court-proof fashion any evaluation of the question of authorship. The data do not allow any empirically sound results. The linguistic expert opinion further states that, in a text product such as this, it is by definition very difficult to determine the authorship. A text appearing in the

press is generated by a long and complex production process in which several people, among them potential authors or co-authors, editors and copy-editors, are involved who may have contributed to the final version appearing in print, "leaving their mark(s)" more or less analyzable for the forensic linguist.

Case D: Disputed authenticity of the minutes of a works council meeting

An employee in an industrial plant has been laid off. The works council has given its approval and both members of the works council have affirmed this in writing in the minutes of the meeting ("Protokoll"). The plaintiff (the dismissed employee) contests the authenticity of the minutes taken in a complaint that he files with the local court. His point is that the text of the minutes – in particular, in terminology and vocabulary – reveals the authorship of a trained legal expert (lawyer) and could not have been written by a person not trained in law, which neither of the members of the works council are. Hence, the minutes themselves and the decision to dismiss him were invalid: it was hence illegal to lay him off.

A linguistic expert opinion requested by the plaintiff is to compare in detail comparison data written by the suspected lawyer and the two members of the works council. The plaintiff's intention is to substantiate the doubts about the authenticity and validity of the authorship of the minutes as officially announced.

In this case, the result of the linguistic expert opinion is again a *non-liquet*. The comparison data do not allow any significant findings and valid results concerning the similarities between the lawyer's comparison texts and the wording of the minutes. The amount of data given in the disputed minutes and the comparison data are so scarce and heterogeneous in text type that no valid analysis could be administered. Even if there had been more and better data available, the linguistic expert opinion argues that an analysis that a certain author could not have written the text of the minutes – here, someone who is not trained in law or who does not know legal terminology sufficiently well – is not possible on safe grounds. In fact, such a constellation could result from yet different authorships; for example, that the text had indeed been formulated by the members of the works council, but had been edited and corrected by the "Hausjurist", the full-time lawyer employed by the company for such cases (which would not have been a violation of law at any rate).

The two cases C and D are both examples reported for a *non-liquet* result. Case E is an example of a *liquet* of a linguistic expert opinion.

Case E: Fake edition of a company paper

In a large company, copies of a company paper (“Mitarbeiterinformation”) are regularly distributed to the employees on a weekly or monthly basis. They pick them up at the entrance gate of the plant. One morning, a fake paper, not edited and authorized by the company, is lying in the newspaper stand at the entrance of the plant. It is quite difficult to recognize at first glance that this is a fake edition. The layout, the distribution and wording at the head of the company paper and all other features that the eye captures have been designed so that it seems real and authentic – seemingly intended by the anonymous author or authors. Only in the course of further and more detailed reading of the main story, the reader notices that this is a fake story with a somewhat negative connotation, stating very negative news of the company and its managers. The information given, whether factually correct or not, reveals insider knowledge. The company orders a forensic linguistic expert opinion to find out whether the texts – two in the same edition of the paper – were written by a former member of the editorial staff of this very company paper. The expert opinion is to prove whether the former editor is indeed the author of these two texts. The company provides a considerable amount of comparison data, which undisputedly are original writings of the former editor. It should be mentioned here that the linguistic expert witness has no way of examining whether the fake unauthorized texts are in fact defamation texts, as the company claims, or whether they report facts, which the anonymous author(s) infer(s) (this is a juridical question to be decided by a judge, not by a linguist).

The result of a thorough linguistic analysis of the two long anonymous texts in the fake edition and of the comparison data of the former editor is, in short, that the former editor is indeed the author of the two texts. The level chosen on the standard probability scale in use in Germany by most linguistic (and other) expert witnesses is: *mit hoher Wahrscheinlichkeit (kann angenommen werden, daß X der Autor ist)* (‘with high probability (it can be assumed that X is the author)’). On several linguistic levels, including grammar (morphology, syntax, lexicon, word order), and linguistic and text-pragmatic features, several idiosyncrasies, “special usages”, errors and deviations can be found to such a large extent in an exactly parallel fashion in the two anonymous texts in the

fake edition of the paper and in the comparison data of the suspect that identical, or at least partly identical, authorship can be assumed. This case is indicative of the “maximum” data for a scientifically sound linguistic analysis of the question of anonymous authorship, the other end of the continuum of cases of multiple authorship (see p. 9ff.) so to speak.

This could best be stated in terms of a rather large and elementary parallelism of the components of the texts in terms of an Ethnography of Communication framework: the comparison texts – authentic and authorized editions of the company paper of a date shortly before and after that of the fake edition – are designed in a very similar fashion and show more or less the same features for the components setting, addressees, sender, wording, text paragraphing, text types, layout, placement on the front page and so on. It is possible to formulate a working hypothesis that the number and type of coincidences or similarities of two texts in many components or categories – which can, in principle, vary independently from each other – are a reliable indicator for a satisfactory analysis of authorship. There is evidence from other cases on which I have worked that confirms this. There is need for further empirical testing to determine which general validity this working hypothesis has. In addition, it needs further verification by a larger corpus. It should also be checked as to what extent the reverse holds. Possibly one can make generalizations that, whenever a larger amount of ethnographic components show diverging features and structural differences, a successful analysis of disputed authorship cannot be made. Account should also be taken of the fact that there is a substantial amount of comparison data of the suspected author comprising two types: texts (of various types and denominations) undoubtedly written by the suspect himself, and other authentic editions of the company paper edited by the suspect. On this basis, with so many similarities, several comparisons and tests can be administered, which allow reasonably safe results.

In the case under consideration, the suspected author, confronted with the summary of the linguistic expert opinion, made a confession that he was indeed the author of the two disputed texts, and he was convicted accordingly.

It should be added here that anonymous and disputed authorship does not necessarily imply a criminal context or, more precisely, a case tried in a criminal court (in which the anonymous texts, as such, represent a crime of extortion, blackmail, defamation and so on). Some cases in which disputed authorship is at stake are tried in a civil court. I was once consulted as a linguistic expert witness in the case of a translation from a foreign language, in which a well-known cartoon was translated

into German by two translators, one responsible for the rough translation, one for the poetic, “polished” translation into idiomatic German (see p. 11). Following the death of one of the translators, a claim was made by the surviving translator that she had not been paid sufficiently well and deserved more money. The linguistic expert was to find out which translator had done how much work. The job could only be performed in a few instances where intermediate results of translations were available and also commented on in writing. In the majority of the disputed texts, no such written sources were available, however, which led to the conclusion that a *non-liquet* had to apply in this case.

Some methodological characteristics of forensic linguistic expert opinions on authorship attribution (AA)

In linguistic expert opinions concerning AA, language behavior does not itself represent the offense (as, for example, in insult and defamation cases). The question of whether an offense occurred or did not occur does not depend on the evaluation of the linguistic behavior. Linguistic expert opinions concerning AA supply the services of an auxiliary science for criminalists and investigative authorities to analyze (limit or exclude) the “authorship potential” of an anonymous text. Partly because of the current state of linguistic research, which has only very recently started to analyze speaker-specific linguistic behavior in a systematic way, and partly because of the nature of the linguistic questions to be analyzed, the production of linguistic expert opinions on AA has even greater difficulties than on proof of understanding. Linguistic expert opinions concerning AA, in most cases, obtain their validation and persuasive potential in connection with other data, from fields adjacent to FL, such as forensic phonetics and handwriting analysis, and, even more so, from other sciences such as chemistry and medicine. In particular, it is seen in relation to data gathered by the investigating criminalists. This implies that linguistic expert opinions on AA alone are of little avail.

A very brief exemplaric illustration of the methodological implications and the limitations that apply to linguistic expert opinions on AA is given below by two expert opinions. Case I is quoted as an example of what a linguistic expert opinion cannot achieve (a *non-liquet* expert opinion). Case II shows what can be done (which, in addition, shows that FL may also exonerate, not simply accuse, a suspect).

Case I

Three defendants are on trial, accused of a series of crimes (armed robbery, trespassing, blackmail and extortion) that they allegedly committed together. One of them is an office worker, another a painter, the third a physician. The latter, in the confession of the second, is named as author – the person who had dictated certain extortion letters – which he himself (the painter) had written down on paper. There is no dispute that the physician himself did not write the letters. The physician, in turn, orders a linguistic expert opinion to prove that he should be excluded as the author (text originator) of the extortion letters. The expert opinion comes to a *non-liquet* result: a linguistic expert opinion may be able to supply probability criteria that a particular speaker is with some likelihood a possible author (on the basis of sociolectal, variety-specific, register-specific and idiolectal features), depending on the particular data of the text in question. In this case, certain specific formulations that occur in the extortion letters could, in fact, make a physician an even more plausible author (for example, the use of the word *Leib* for ‘belly’ instead of *Bauch*: in colloquial German spoken in the Rhineland, the latter is the only frequently occurring unmarked form). The entire analysis is not empirically safe, however. It is not possible, in a scientifically reliable way, to exclude one out of three possible authors merely on the basis of socio-economic, educational and other parameters. The question as to whether a speaker of higher educational attainment and higher social class is generally more capable of imitating other people’s language styles more easily than speakers of a lower level cannot be touched upon here. A physician may (!) have less difficulty imitating the speaking and writing behavior of a non-academically trained worker if he wishes to hide his authorship specifically. There are also examples for the reverse constellation, however.

In other words, if there is more than one author of more or less equal socio-cultural characteristics, only the people involved know certainly who the author of that text section is. This question of multiple authorship also applies to texts outside the criminalistic realm; for example, to the ‘fathers of the German constitution’ (“Die Väter des deutschen Grundgesetzes”).

Case II

In a large company in the Rhine area, numerous anonymous defamation letters are sent to various employees and clients, in which criticism

and allegations – including delicate internal information about the director of a section – are discussed. The letters reveal insider knowledge of the company. Officers in the company, on the basis of other (non-anonymous) letters and activities in the past, suspect a particular employee as the anonymous author. He has a psychiatric record and had written anonymous letters of complaint to a former employer years before. He is from the local area. Because of the damage being caused to the company by the letters – which are also being sent to people outside the company, including clients – a linguistic expert opinion is ordered (in the original wording): *Das Gutachten soll prüfen, ob X als Verfasser der anonymen Briefe in Frage kommt* ('to examine whether X is a possible author of the anonymous letters'). The suspect is a middle-aged clerk with a middle school (rather than a high school) diploma, who had never left his hometown Cologne for any length of time. He does not know any foreign language.

The expert opinion states, summarized very briefly, that the incriminating anonymous letters show some exceptional spelling and language behavior (such as excessive use of hyphenation, in particular with numbers, certain syntactic, lexical and textual errors) indicating that the author could be a non-native speaker of German. Some of these are anglicisms, such as (non-idiomatic) German *diese Dinge werden mit anderen Maßnahmen begegnet* ('these things are met by [...]'), non-native and non-authentic quotations of German proverbs *Eine Krähe hackt einer anderen nicht in die Augen* instead of the idiomatic *Eine Krähe hackt einer anderen kein Auge aus*, and other data of that nature. On the basis of a thorough linguistic analysis of the incriminating letters and comparison data written by the suspect from Cologne, it was possible to exclude him as author and writer with a high degree of probability. The company was asked to search for other potential authors, who were non-native speakers of German. Two non-native speakers of German were found in the section of the company in question; one of them a German American, the other a native speaker of French. Several sets of comparison data of the two suspects were then analyzed. A particularly large sample of texts by the German American was examined because of the relatively large number of anglicisms in syntax, semantics and phraseology. At the end of this analysis, the German Kriminalpolizei ('Criminal Investigation Police Department') was asked to monitor the employee. Eventually he was caught in the act of writing yet another anonymous letter to the company. In this case, the linguistic expert opinion was able, with a high degree of probability, to exonerate a suspect of being the author and writer of the incriminating texts in question, which was subsequently proven by solid criminalistic evidence.

The most important tasks, potentials and limits of FL expert opinions on AA can be stated as follows: the success of an FL expert opinion on AA (eventually in connection with other, non-linguistic evidence) depends primarily on the nature, quantity and quality of the data available, along with the painstaking care that the linguist applies in the analysis. Obviously, extortion letters and anonymous letters of various kinds very rarely show features that will truly identify a speaker. Also, language as a means of communication – particularly in such interactions – shows, unsurprisingly, more inter-personal than intra-personal characteristics. The job of linguistic speaker/writer identification is a very subtle, difficult and time-consuming endeavor. There is no reason to have exalted expectations, nor to think that FL authorship attribution is generally fruitless and not worth anything. The decisive question is whether the anonymous text product and the comparison data of the potential authors are analyzed in an encompassing systematic way, in regard to the system of linguistic behavior underlying it and to the methods of validation of data. The first and most important requisite is, of course, that there are enough appropriate incriminating texts and comparison data available.

The term “linguistic behavior” used here should not be given too narrow a definition. Verbal strategies for the code and “phenotypic” strategies for the graphemic level (including orthography, layout and make-up) and data of all the other ethnographic components (see Gumperz and Hymes 1972) have to be taken into consideration with equal care as they can provide clues to a syndrome of speaker-specific linguistic behavior that might be relevant in one way or another for the holistic analysis of anonymous authorship.

In other words, language data as single items – for example, words as static text products – are of little empirical analytic potential. They can never be used as solid proof of anything. In most cases, they do not reach beyond anecdotal data. The forensic linguist’s job in dealing with AA is to supply a systematic reconstruction of linguistic behavior of incriminating and comparison texts, and match these data to each other in a sound and convincing way: The linguist must identify (a) the system of features that is common to the incriminating anonymous letters; and (b) a sufficient amount of known comparison data from a suspected author. If sufficient data for both can be gained, one has to examine whether they are structurally compatible, whether they match in terms of common or partly common authorship and writership. One has to examine whether the products of linguistic behavior have a common denominator that suggests that the analysis is completed in a typical and analogous way. The next methodological step is to determine the

status of variation of *both* data sets of what one could call the general language standard. If the variation in both cases represents a recurrent proportion pattern, one can assume with some probability (different in each case, and to be measured thoroughly) that this is not just a speaker-specific set of features, but that these features taken together could possibly serve as specifying or identifying features. This is particularly appropriate in a systematic analysis of the co-variation and configuration between sets of dynamic data in different areas that, in principle, vary independently of each other. This is a process similar to that of sociolinguistic analysis.

Of particular importance, although not the only data of interest, are mistakes, errors, deviations, non-idiomatic versions and variations of language behavior. This does not refer to errors as some kind of erratic frozen product of written texts, such as in spelling, punctuation, morphological form, lexical choice, syntactic construction, or word order. Reference is made here to mistakes and errors as positively defined deviations in ways of handling communicative situations. Errors have to be defined with regard to their status of occurrence (for example, consistency in their use within and across text types) and in lieu of their general deviation status in relation to norms in the general language system.

One could simply say that it is not of particular interest how unusual or exceptional an expression is *per se*. Rather, it is of interest how consistent and typical it is for the language usage of particular speakers in their generally unusual way of managing the text: not merely *that* a particular item is an error, but *how* it works and behaves as an error if used in a particular linguistic and situational context.

The anonymous author in case II quoted above shows several dynamic characteristics of language use that suggest that they be analyzed more closely: there is no inflection marker for the German dative plural ending in a prepositional phrase, such as *macht zu Saubermänner* instead of *macht zu Saubermännern*. There are mistakes in agreement of number and gender of a specific pattern and frequency of occurrence in the texts. The anonymous letters also show a rather unusual letterhead and an idiosyncratic punctuation of relative clauses. All these features taken by themselves, and partly even taken together, can be characteristic of a population of German speakers/letter writers numbering in the thousands. Such an analysis usually does not yield much in an AA case. The decisive data are additional data describable as “dynamic configurations” of features in various areas of grammar, textuality and graphic and phenotypic make-up. If configuration patterns can be shown to exist between all the classes of features (grammatical, textual, make-up and so on), and if they match in the incriminating and the comparison

texts, these are data with a noteworthy descriptive and explanatory potential. If they can be marked as originating from a usage pattern that non-native speakers/authors of German (at least in this case) show, this may represent, with a high probability, a result relevant for the identification of the author.

To state the salient points of the argumentation once more: the argumentative potential of these features in combination, the configuration of features as a whole, is relatively strong, if it can be shown that in a particular text it is used consistently, and if neither the arrangement as a whole nor the particular features independent of each other can be explained in a different way. If it can be shown that special syntactic constructions are not conditioned by dialectal influences, and if mistakes in number agreement cannot be explained by the psychological state of the author's editing of a letter, this needs further attention. One always has to take into consideration the totality of configuration patterns applicable to alternative explanations to describe the etiology of errors. If they are of equal explanatory value, the solution envisaged has to be refuted. Deviations and configuration patterns and all other potentially significant data have to be matched in a complex multifold process to alternative explanations on various levels. Particular attention has to be given to contrary findings and their explanations. The greatest potential for explanation concerning speaker-identifying features of linguistic behavior lies in identifying those that are complex in form and remarkable in frequency of occurrence, and are configurations of syndromes in several respects atypical of a given standard of verbal behavior. The data of speakers of the Cologne dialect ("Kölsch") as their mother tongue, whose high (Standard) German shows the well-known influences in syntax, inflection of the verb and phonology, are no doubt of great linguistic interest but not of much help in forensic AA. One can find a great many of them in each text and case. Of particular explanatory value are those features, text type regularities, errors and mistakes – for example, in syntactic construction, government of the verb and word order – in the Standard German realization of a native speaker of Kölsch that are *not* explicable by influences of dialect.

Mistakes in number agreement are extremely frequent in spoken and written texts of almost every native speaker of German in everyday communication, if sufficiently long sentences are used, and if an informal speech situation is given, leading to superficiality of articulation. Such data are not really interesting for AA, except as some kind of a standard that can be (a) matched against the particular usage in a text or if just one particular type of error in agreement occurs, and not any others;

(b) if an individual speaker shows exactly the opposite distribution, for example, of features of formal language in text types that the rest of the speech community shows; or (c) if a particularly marked mixed distribution of textual and text type-specific features exists. As potentially speaker-specific data, behavioral patterns of one single type can be taken – such as agreement errors and concomitant non-occurrence of all other (in principle, equally possible) types of errors.

It is clear that FL expert opinions have to be written in closest cooperation and consultation with other fields of applied linguistics, such as linguistic error analysis, second language acquisition, aphasia research, stochastic linguistics and statistics and, above all, corpus linguistics. Linguists may almost seem to be more deserted than actually helped by their own science. If one concentrates intensely on the study of grammatical competence – that is, the ability to produce and understand an unlimited amount of grammatical sentences – one does not pay much attention to the fact that a creative linguistic competence, even concerning facts of grammar proper, is developed in different ways and to different degrees by different speakers in real life. In addition, it is necessary to analyze and define structures and systems of this diversity in a scientific way. A brief hint concerning random or characteristic errors, accidental or systematic mistakes, is not enough. Some non-linguists devoted to empirical science appear to have difficulty understanding these facts.

Summing up, one can define three general tasks of critical importance for FL expert testimony on AA that the field does not seem to have taken into consideration sufficiently to date:

(1) In AA, an FL analysis of features in combination is necessary, which, in principle, may vary independently of each other. This can best be termed as a “dynamic configuration of features”. It means that a consistent gradual addition of single (static) features is not the main goal. Rather, it is necessary that a multifactorial analysis of the connection between configurations of features is undertaken in a manner established in the social sciences. Mainly collecting and describing isolated material features does not really provide the answers in AA. The mere number of similarities and differences is hardly ever significant for questions of authorship.

(2) One should realize that texts are frozen products of linguistic behavior and have to be described as such. They do not show a total and in-depth representation of the linguistic behavior involved. The prime focus here is to describe (to “reconstruct” in the terminology of S.P. Corder (1971)) dynamic structures and units of linguistic behavior.

All definitions and empirical results should be conceived of and oriented towards such entities.

(3) Forensic linguists need to continue to discover significantly more information on repertoires of idiolectal behavior. It seems that there is a much larger number of speaker-specific and potentially speaker-identifying forms and styles of linguistic behavior than is assumed at present. This simply means that speakers cannot be identified in terms of particular words, orthographic errors or syntactic constructions. There is always a high chance that a large number of other speakers share these same static features. It is not of importance whether speakers have a certain item in their repertoire, but how they use it and also the repertoire as a whole in a specific way to distinguish them from other speakers.

“Dynamic features” of linguistic behavior cannot be manipulated by any speaker (including forensic linguists) for any larger amount of data, such as two or three A4 pages of a text. This is, in fact, some consolation for criminologists and forensic linguists giving expert testimony: No author, even if paying attention to this type of language use, could manage to oversee a larger amount of features simultaneously over a longer period of text production. Even if authors wish to disguise their authorships, fake a text or imitate someone else’s particular style of writing – for example, by using a lexical repertoire purposely not identical to their own – careful analysis can find this out eventually. The very structure of the manipulated linguistic behavior as a whole would be marked in a specific way. This, in itself, could help to reveal the fake. There will always be some concomitant or co-varying significant features of which the authors did not think. In other words, the very pattern of an intended simulation of another authorship can itself be indicative of the particular disguise or fake.

It is important to state that there are no simple “lucky cases”, accidental constellations of anonymous letters and comparison data of a suspect, in which FL authorship analysis can be applied successfully. Such material coincidences or parallels are much less frequent than is generally assumed. They never can supply a conclusive result or proof for the question of authorship of a text. There is a much higher number and a much more salient set of dynamic indicators of linguistic behavior, mainly “unconscious” usage. Of critical importance is that the dynamic data of linguistic behavior – of an anonymous extortion letter as well as of any other written product of linguistic behavior – are not analyzed by intuition and free projection by the linguist, but by a systematic, empirically safe method of investigation, which could be called “linguistic

differential diagnosis". To identify the author of an anonymous incriminating text is not merely down to luck, but the result of a systematic FL analysis of texts *lege artis*.

This also implies that (1) it is not possible to give quick answers regarding authorship; and (2) linguistic laymen normally cannot do this type of work. It appears that linguistic laypersons' assessments of identity of authorship of two classes of texts are even less reliable and less valid than laypersons' judgments in phonetic speaker identification. FL authorship attribution is a highly time-consuming, theoretically demanding, and practically and methodologically complex endeavor. This should be noted also by linguists that try to do FL authorship attribution "on the side".

As a final point, it should be stated explicitly that FL expert opinions concerning AA can, in general, only afford a lesser degree of evidence than, for example, forensic medicine or forensic chemistry analyses. In other words, FL expert opinions can only supply indirect degrees of probability for the questions of interest to the investigating authorities. This is a general limit that should always be kept in mind by people giving linguistic expert opinions on AA and by people that request them. There are no "idiolectal fingerprints", and, if anything like this were to exist, linguistics would not have the instruments necessary to measure them with the degree of accuracy needed (see Kniffka 1981: 598). This results from the very nature of language and language data. It is important to mention this here, since, frequently, linguistic expert opinions concerning AA are requested when there is a general state of emergency of proof in a case, or more precisely, when (all) other possibilities of empirical analysis and investigation (notably, by more exact sciences) have been unsuccessful. This is unfortunate and certainly somewhat counterproductive as far as the general assessment of the value of FL analysis of AA is concerned. Certainly, forensic linguists cannot come up with results that other (more exact) sciences cannot provide. So, the main value and salience of a forensic expert opinion on AA is that linguistic data can be supplied as data of an auxiliary science to support and supplement the results of the other sciences involved. Very frequently, a combination of the results of different sciences – which are as such homogeneous and lead in the same direction – is a valuable aid for what the investigating authorities seek. It would be inappropriate to attribute an argumentative potential and function to the field of linguistics that cannot be secured by other court experts and auxiliary sciences.

3

Status and Tasks of Forensic Linguistic Authorship Analysis

Introduction

This chapter is the English version of a paper read at the Symposium “Forensisch-Linguistischer Textvergleich”, which took place on 8/9 December, 1988, in the Bundeskriminalamt (BKA), Wiesbaden, and which first appeared in Bundeskriminalamt 1989: 205–36. It also appeared under the title “Nochmals: Verrät der Text den Verfasser?” (‘Once again: Does the text reveal the author?’) in the November 1990 edition of the journal *Kriminalistik*: 604–10 (Kniffka 1990e), which addresses criminalistic practitioners of all levels, mainly higher levels of German investigating authorities, such as the BKA, LKA and so on. The editor of the journal asked the author to rewrite the paper, addressed to a non-linguistic audience. In spite of the time that has since passed, it seems that this very perspective is of interest to forensic linguists some 20 years after it was written (describing, as it also does, the early stages of forensic linguistics in Germany, which are not always given proper attention).

It is interesting to see how the editor of the journal *Kriminalistik*, an experienced former director of an LKA, summarized the work of forensic linguists in criminal court cases to illustrate why he had invited the author to write yet another report for his journal. Summarized, his main points are as follows:

- (1) Linguistic terminology is very heterogeneous, confusing and, to a large extent, asking a bit much from non-linguists.
- (2) The scientists (linguists) involved do not seem to agree much, even in theoretical matters, on the possibilities of analyzing and identifying “author-specific linguistic features” and there seems to be a necessity

for clarification. Is it possible to summarize the *état d'affaires* yet again in a way that is also intelligible to non-linguistic laymen?

(3) Along these lines there are some doubts as to which function can, in general, be assessed to the field of linguistics as a criminalistic auxiliary science in practical cases. In how many cases were suspects successfully traced by linguistic expert opinions? In how many cases were sentences based on evidence supplied by forensic linguists, and so on?

(4) In particular, questions of practical concern are to be discussed. How large a data reservoir of incriminating texts and comparison texts is necessary upon which to base a linguistic expert opinion? Are there any general rules for supplying and interpreting comparison data?

These introductory remarks by the editor of *Kriminalistik* are stated here because they indicate the situation of FL from the perspective of practitioners in 1988/1990. One can ask, what is the situation today? I think the questions raised by the criminalistic practitioner then are probably the same as can be stated for FL today. This is worthwhile noting from the point of view of 2007 even if, due to restrictions of space, several linguistic aspects proper had to be neglected and can only be referred to in broad terms. The gist of the 1988 paper (Kniffka 1989) was stated in 14 theses, which, in an even more condensed form, are listed below.

Thesis 1: The notion of “idiolect” in linguistics

Little or almost nothing has been clarified in a scientifically sound and satisfactory manner about what can be defined as speaker- or author-specific linguistic features. Linguists who have worked in this area for some time know that they can be certain of very little with precision. Criminalist practitioners should know that most competent linguists know very little about this. One should take a rather skeptical view of all swift *ad hoc* recipes.

Linguistics as an academic discipline has always been interested, and probably will always be interested, primarily in the inter- and supra-individual data of language, in the uniting rather than the dividing lines of the language system, the idiosyncratic differences, and the speaker-specific features distinguishing speakers from other speakers in a language community. However, it is not justifiable and sensible to conclude from this that linguistics and the neighboring sciences do not have anything important to contribute to the analysis of speaker-specific behavior.

There is massive everyday evidence for the existence of speaker-specific characteristics, such as the language behavior in a written text,

in a telephone conversation, in other text types and, for example, the fact that one can recognize the voice and the speaking habitus of a family member or a long-time friend even after decades. There are other points; for example, the immense amount of proverbial data suggesting that there is something such as speaker-specific language behavior.

There also is sufficient scientific evidence that the field of linguistics can and must describe this type of language behavior. This holds in particular for several areas of applied linguistics. What is not generally agreed upon, and is still in a somewhat preliminary state, are the questions concerning which methods to use and by which conceptual instrumentarium the analysis and the description of such speaker-specific behavioral data should be accomplished.

Thesis 2: The term “forensic linguistics”

The question concerning the contribution linguistics can make as an auxiliary criminalistic science has not been applied sufficiently systematically to date, let alone been answered. The very fact that there is a term “forensic linguistics” should not lead laymen from other areas to the assumption that everything is settled.

Thesis 3: Forensic linguistics as an auxiliary science

A new interdisciplinarity in terms of quantity and quality requires an interdependence and reciprocity of information between the sciences involved. Linguists have to inform criminalists, and also vice versa. This applies expansively to the theory of the field and particular applications, and also to the theory of application. I have been unable to find precise information offered to linguists by the judiciary as far as theoretical implications or general criminalistic investigative work is concerned, or also concerning the expectations that originate from the judiciary with regard to the field of linguistics. Linguistic experts should also be informed, or acquire their own information, on experts and the problems related to expert opinions that have arisen in other established forensic sciences with a much longer tradition than FL; such as forensic psychology and psychiatry, forensic medicine, and forensic physics and chemistry. As an example, I cite the problem of standardization arising in psychological expert opinions and the so-called “scales of probability”, which are being worked on intensely in forensic psychology, and the double translation process of a forensic expert opinion (see Kniffka 1981). Also, a linguist should be informed about the literature and the work done

concerning the general status, function and problems of evidence (“Sachbeweis”) as discussed in Dippel (1986), Geerds (1983), Haller and Klein (1986) and others, and about the position of the expert in a criminal court procedure. This literature seems particularly revealing for linguists in lieu of special questions; such as the explanation of methods of investigation, the discussion of admission to original documents, and the court transcripts by the expert. In each case, however, linguists would have to check thoroughly whether it is useful and sensible to publish data from linguistic expert opinions that were helpful (conclusive) in particular cases. In fact, this does not seem to be a significant problem; that is, forensic linguists should be given ever more detailed information, in quantity and quality, by criminalists and other people working in the legal arena. It seems that this would, in turn, increase the contribution that forensic linguists can make to these fields.

More recently, linguists have worked thoroughly on the analysis and systematic discussion of corpora and data processing, which should be included in FL work. It is important to make it clear to non-linguistic people working in the law that language data are not just language data, irrespective of how, where, when, why and for what purpose they were recorded. Significant differences for the interpretation of a text can result, for example, from the exterior conditions in which a text originated – the localities, the circumstances and so on. Linguists should be able to supply the police with techniques and hints on how to collect and supply textual material as comparison data relating to a suspect. Law enforcement officers would probably benefit from linguistic background information as to which data sources to look for at the suspect’s home and how to select relevant data. If there are guidelines for sampling data for handwriting analysis (“Richtlinien für die Beschaffung von Schriftproben für die Handschriftenvergleichung”) and some well-defined standards for it, there is no reason to object to such general guidelines being set up in a parallel fashion for sampling textual data of incriminating and comparison texts in linguistics (“Richtlinien für die Beschaffung von Textproben für die Textvergleichung”). It seems beyond doubt that such standardized information and guidelines would be helpful for FL experts as well as for investigators and criminal police.

Thesis 4: “Features” as a linguistic term

Thesis 4 mainly concerns self-criticism within the field of linguistics. As an empirical linguist, I have collected many different types of

expressions for covering the denotatum of “variants of expression” (“Ausdrucksvarianten”) and “speaker-specific linguistic features”. Table 3.1a, which is self-explanatory, gives a (non-exhaustive) list of German expressions. It is easily intelligible that non-linguists from other areas are receiving the impression that linguistics must be a terrible science if such a terminological variation or confusion exists. So, it seems that the comment made by the editor of the journal quoted above on the difficult and variable terminology of linguistics may indeed be justified. Table 3.1b shows a variation of a larger terminological caliber.

These examples illustrate that the terminological variation can be as confusing to linguists as it is to non-linguists. But it is also true that a certain amount of terminological differentiation is necessary in linguistics, probably more than laymen are inclined to realize. The overwhelming impression is, however, that expansive and systematic linguistic basic research is needed to provide a proper conceptual and terminological instrumentarium. Also, precision in definitions must be developed in

Table 3.1a Terminological variants for the description of the variation of forms¹

-
- fakultative Merkmale (‘facultative features’)
 - intrapersonale Vorlieben (‘intra-personal preferences’)
 - stilistische Varianten (‘stylistic variants’)
 - stilistische Vorlieben (‘stylistic preferences’)
 - stilistische Alternativen (‘stylistic alternatives’)
 - kontrastive Elemente (‘contrastive elements’)
 - distinktive Merkmale (‘distinctive features’)
 - konkurrierende Elemente (‘competing elements’)
 - stilistische Neigungen (‘stylistic inclinations’)
 - sprachliche Optionen (‘language options’)
 - markierte Elemente (‘marked elements’)
 - mehrere Möglichkeiten des Ausdrucks (‘several possibilities of expression’)
 - Stilfiguren (‘style figurae et topoi’)
 - Stilmerkmale (‘style features’)
-

Note: The expressions refer partly to the same and partly to different linguistic matters.

Table 3.1b Terminological variation for speaker-specific linguistic features

-
- linguistische Merkmale (‘linguistic features’)²
 - sprachliche Merkmale (‘linguistic features’)
 - sprachliche Indikatoren (‘linguistic indicators/language indicators’)
 - Indizien (‘circumstantial evidence’)
 - markierte Elemente (‘marked elements’)
 - Parameter (‘parameters’)
 - Marker (‘markers’)
-

order to meet the requirements of the theory of science in terms of validity, reliability, intersubjectivity and economy. If necessary, even a certain set of norms should be defined by FL fields. It is also necessary to look into adjacent areas of linguistics; for example, sociolinguistics, in which much has been done to clarify the distinction between alternation and variation – such as Labov’s distinction between “indicators”, “markers” and “stereotypes”. No excuse can be seen for forensic linguists not to take this into consideration systematically.

There are still more problems to face. For example, what is meant by “stylistic variants” depends on the general conception one has of the field of stylistics. Sometimes even, one author uses various concepts for the same phenomena and also for different ones. A drastic change regarding clarification is needed if FL is to be taken as a serious forensic discipline in court. Some “linguistic” expert opinions suggest an inadequate and unjust picture of the field. The field of linguistics as a whole is better than some expert opinions appear to indicate.

Thesis 5: Methodology of authorship analysis

This thesis concerns some methodological postulates of FL authorship analysis. The question as to how many sentences or pages of an incriminating text are needed in order to administer a linguistic expert opinion *lege artis* cannot be answered in terms of absolute numbers or in quantitative terms at all. It is an open question whether this can be resolved in the future. It depends on the aim of the investigation, the size and nature of the available data (both the incriminating texts and the comparison data), the quantity and quality of certain notable features in the data investigated, which not only concern mistakes, errors and such, but also all other kinds of data including those that are correct. According to my experience in real life cases, it may be that three A4 pages of text for incriminating and comparison data are sufficient in order to arrive at a solid conclusion. It may also be that 30 pages of incriminating texts and twice the amount of comparison data do *not* allow significant results. It would be a great advantage if the pools of data available allowed detailed stochastic information on the size of incriminating texts necessary to base empirical generalizations thereon for specific text types, for a particular genre and so on. To date there is not much done in this direction.³ However, not everything is a matter of size of data (see below).

It should be a maxim of FL work to formulate specified weak hypotheses rather than strong hypotheses in lieu of method, area, range of generalization and other factors. Strong hypotheses, in particular in this area,

are easily falsified; for example, a generalization of an author's expression as a whole based on just a few pages of comparison texts.

It has also been particularly detrimental to the German FL scene that certain spectacular cases and certain "spectacular linguists" (semi-linguists and non-linguists) entered public discussion, thereby having a bad influence on people's (laymen's) thinking (including that of some linguists).

It can be assumed that the FL work on non-spectacular cases has, scientifically, at least the same methodological impact. There are no spectacular results *per se*. FL authorship attribution is a demanding, meticulous, time consuming, detailed and not at all spectacular endeavor.

Thesis 6: The need for theoretical reflection

This thesis is addressed in particular to criminalistic practitioners working or having to work with linguistic evidence. There are no quick and ready-made all-inclusive ways of solving cases in FL. It is not true that linguistics is superfluous because all that is needed is a layman's common sense. Quite the opposite has been shown in Kniffka (1981) (particularly in the area of expert opinions concerning defamation and libel). It is shown there that the common sense of the layman is not sufficient and cannot replace an empirical linguistic analysis. "Life experience" cannot replace theoretical reflection, systematic analysis and a strict methodology of linguistic data analysis.

I tend to agree with criminalists who postulate that linguistic theoretical basic research and applied basic research should be undertaken with a higher degree of reference to practical work. I also agree with criminalists that they cannot wait 10, 20 or 30 years until linguistics has developed a theory of the idiolect or idiosyncratic language data.

All this is no valid argument, however, against the necessity of a systematic linguistic analysis and theoretical reflection. The relevance of the field of FL, as in authorship attribution, has to be developed step by step. Results of other areas of linguistics, in particular those of applied linguistics and sociolinguistics – such as Labov's "cumulative principle" – have to be taken into consideration. Linguistic expert opinions as an auxiliary science for criminalistic investigation by their very nature have a cumulative principle as a maxim. The same seems to apply to expert opinions of many fields that have a similar amount of professional routine and experience. It is questionable, however, whether this quantity of "criteria of authority" can be defined for linguistic experts only and not also for experts' opinions in other areas.

Linguists differ from laymen not by the fact that they are better or more competent speakers (as many judges kept telling me in the old days of FL). The significant difference is, rather, that they have a training, knowledge and experience in the scientific analysis of texts and in the production of expert opinions. There is no question that a reservoir of frozen knowledge of linguistic experts and some kind of a reservoir of linguistic diagnostic know-how should be pooled and made available to young linguists working in the field. There is a huge deficit in research and teaching in this area that has not even begun to be accounted for systematically.

Thesis 7: Forensic linguistics as a branch of linguistics

To speak of forensic linguistics (“die forensische Linguistik”) is undoubtedly not a fully justified and appropriate way of designating the field. It suggests that FL is an established forensic auxiliary science as other forensic auxiliary sciences. This is not so, however, since several aspects of the theory of science, and also practical methodological postulates, have not been fulfilled to date. From my own perspective, it is not necessary to have a particular area of real life and/or neighboring sciences to constitute an own branch of linguistics called forensic linguistics. One could end up talking about “gastronomical linguistics” for the analysis of the language of restaurant guides, “nautical linguistics” for the language used in naval communication, or “communal linguistics” (“Kommunale Linguistik”) for the language used in offices of public administration. FL is a branch of linguistics applied to a particular section of everyday real life. One cannot, and should not, assume that FL is the sum of all expert opinions concerning authorship attribution, defamation and so on. At present, FL is a branch of linguistics that just happens to concern matters that are dealt with in the legal context.

At a later time it certainly will be necessary (1) to give a satisfactory extensional definition of the area(s) of FL; (2) to give an intensional definition of the features that distinguish this field from other fields of applied linguistics; (3) to establish a full-fledged science of FL, which would entail a systematic clarification and classification of the methods used; (4) to create a heuristic typology of the fields of real life work; and (5) to establish the position of FL in the framework of other forensic auxiliary sciences based on the theory of science. If the postulates (1) – (5) cannot be established, the term forensic linguistics and a branch of (applied) linguistics with this name are somewhat redundant. Linguistics is linguistics, no matter what the subject matter data happen to be.

Thesis 8: Sub-disciplines of forensic linguistics

There is no rationale to assume that FL is a linguistic auxiliary science merely for authorship attribution, though this is the prime concern of criminalistic investigators in Germany at present. It may even be necessary and useful, at a later time, to coin a new term “criminalistic linguistics” which would characterize the linguistic work in authorship attribution in criminal cases versus other fields such as FL analysis used in defamation and libel cases, trademark names analysis and press corrections, as well as analyses of legal documents, law texts, courtroom language and other fields. No matter what the particular point of view is, it seems clear that one cannot define a new branch of a science by only a very small section of it (that is, authorship attribution). So, in short, FL cannot and should not be identified with linguistic authorship attribution only. It should be used best as a cover term for linguistic expert research and testimony of any kind in/for a court. FL would be a hyponym of a larger field – language and law – which itself would be a branch of applied linguistics.

Thesis 9: Qualitative and quantitative analysis

This thesis concerns a very general and basic research claim, and strongly reflects the overall position that a linguist may have within the field as a whole. The main concern here is the question of the extent that FL should employ: in simplified terms, quantitative or qualitative analysis – or, more exactly, the extent to which FL must use quantificational measures and methods and to what extent it should use qualitative analyses.

In brief, my own conviction is: Everything that can be measured in formal (statistic, stochastic, quantitative) terms should be measured. By any means, yes. However, not everything in texts relevant for the analysis of texts can be measured: not everything is quantifiable, statistically analyzable. There are other relevant and important sections and domains of data that are not subject to formal measuring; for example, data that only gain their particular argumentative status and explanatory potential in connection with text-external evidence – as exemplified, for example, in the *Ethnography of Communication* originated by Dell Hymes and others (see Hymes 1962; 1966; see also Gumperz and Hymes 1972).

Thesis 10: FL anonymous authorship analysis within a sociolinguistic framework

It would seem that FL would benefit from using the results and methods of social sciences to a much greater degree and in a much more meaningful

way for the description and explanation of data. If linguistics can be defined as the science of how language works (a) as a system in itself; and (b) as a system of human communication in everyday speech situations, it is without question that (a) has been accounted for much more thoroughly and intensively in FL. It is also clear that many more data have to be described for (b) in a much more solid and intensive way than has been done so far, and that, in addition, the data have to be analyzed with the necessary methodological rigor, width, depth and detail.

The main potential and the essence of FL authorship attribution of anonymous texts are located in the sociolinguistic domain of the analysis. This is not just an impression, but an assumption based on observational data of linguistic expert opinions given in many real life cases. Simply put, one can say that sociolinguistics is “where it’s at” for the description and explanation of cases of authorship attribution. Among the postulates that linguistic expert opinions should have to take into account one way or another are the following:

- The so-called “observer paradox” (Labov 1972a) will have to be taken into account in a methodologically sound way. It means, in simplified terms, that linguistic behavior can only be described and explained adequately by systematic observation, but that this very observation changes the object of the description. This is of particular importance for the supply and the systematic analysis of comparison data.
- It may be necessary to distinguish “self-assessed”, “other-assessed”, and “factually observable” linguistic behavior as used, for example, in language attitudes and stereotypes research. This may apply to the linguistic object data described in the expert opinion as well as to the meta-linguistic data of the language that the linguistic expert uses in an expert opinion.
- The categories “assumptions about language”, “linguistic folk taxonomy”, and “language attitudes” should be taken into consideration as important areas of the evaluation of linguistic variants. These do not exist in empty space. Their values, which are considered to be speaker-specific in this context, and their status as speaker-specific features cannot be seen overtly. They do not have name tags attached to them. They exist in the language system and language use of speakers, including the specific evaluational system and understanding they have of their own and of other people’s language use. This applies, of course, to defendants, suspects and linguistic experts alike.

In other and even more simplified words: a particular linguistic variant is not simply describable as, for example, somewhat out of date or somewhat

exceptional. Different speakers may have very different assessments for one particular form. So, these subjective factors have to be analyzed systematically and have to be taken into consideration (also as possible distortion factors creating artefacts).

Thesis 11: Dynamic data of speaker-specific behavior

It is not sufficient to compare static linguistic products of texts, fossilized data – for example of incriminating letters – with those of comparison data secured by the police. It is of critical importance to work with dynamic entities, including strategies and routines of language behavior. Linguistic error analysis as created by S.P. Corder (1971) and others is a very convincing model here. A particular set of data that appears in a frozen text product does not mean anything if it is not used as a basis for the reconstruction of the speaker's behavior. This is all the more necessary, since the data that can be found in an incriminating text are but a tiny section of a person's linguistic repertoire and competence altogether. This also implies that it is not possible to make any generalizations or claims concerning the authors of anonymous texts as a whole, rather than on a particular set of writings they have produced. The comparison of single static linguistic items, isolated mistakes, errors and similarities – such as in two pictures in which certain mistakes can be found by comparison (“Suchbilder”) – is certainly a necessary step in the description in FL authorship analysis. But it is insufficient and incomplete. On the basis of a large section of data as provided by the total text products, it is necessary to make a “Ganzheit” analysis by means of a methodologically sound reconstruction. It has to take care of the interrelations and the general tendencies of configuration that exist for the various partial strategies and verbal activities responsible for a text product.

To stay in the same metaphorical domain: rather than describing a static picture or two static pictures and their deviations, the text products have to be seen as parts of a film that is to be reconstructed from a sequence of pictures of various states of motion. In each particular case, one has to use a different mutual background and standard for the description of dynamic language data. On this basis, an evaluation of differences and similarities seems more adequate than an evaluation made on a static basis, having the general view on a linguistic behavioral syndrome as a whole.

It is well known that certain text types are produced by a rather slovenly and/or superficial attitude as compared to others; that, for example, adolescents use more care in writing a dictation test at school

or an essay than in writing a letter to a friend. One seems (or one claims) to know what an accidental mistake (“Flüchtigkeitsfehler”) is and also one pretends to know where, when and how they occur. In reality, linguists seem to know little about the etiology of the occurrence and non-occurrence of certain mistakes, errors and deviations in a particular text type.

This is a huge area with much work to be done before these methods can be successfully applied in FL analysis of anonymous texts. If one accepts the general need for a processualization and dynamization of sociolinguistic data for FL analysis in a criminalistic forensic context of authorship attribution, the following specific questions can be addressed: Which speaker-specific invariants of linguistic behavior can be found in anonymous text X? Which can be found in text X and in the comparison data Y? Which individual range of variation exists? Which measure of deviation can be assumed?

Some linguistic data, including data of orthographic behavior, seem more sensitive to variation than others. The use of the apostrophe or of the hyphen in German shows, for example, a relatively higher degree of variation than, for example, capitalization (in German). The use of the full form of the second person singular imperative form in German as in *gehe, mache, höre* as opposed to the shorter form without *-e* (*geh, mach, hör*), and the general systematic linguistic background, (that is, not all German verbs allow the omission of *-e* in the forms of the second person singular imperative) suggest a rather complex descriptive pattern and potential for argument of the occurrence of such forms in particular text types. An even more demanding and complex endeavor is the identification of idiolectal linguistic behavioral syndromes, which show a specific mixture of various tendencies and strategies of usage in this case. Speakers can be classified on a continuum according to their use(s) of a configuration of such features (see Kniffka 2000b; 2003a; 2003b). A first impression of such a range of variation is given in the Duden, considered by laypeople to be the “bible” of German orthography.

Thesis 12: Linguistic “differential diagnosis”

The complexity of the data of linguistic behavior requires the development of complex methods of analysis and operationalization that secure the validity and reliability of the data and the methods used. “Method” is the key term in FL. The focus is the comparison of data that are also the focus of many other areas of linguistics; such as historical linguistics and sociolinguistics. An older school of descriptive linguistics, the so-called

“taxonomic structuralism”, has developed a method called “contrasting minimal pairs” according to the principle *ceteris paribus*. This means that the two members of a pair contrast on the same grammatical level – for example, in phonology – in one single segment only, such as *Wind* (‘wind’) and *Wand* (‘wall’) in German. By setting up sufficient minimal pairs, one can discover the phonological system of a language successively; that is, the smallest distinctive elements used to signal differences in meaning.

It would seem sensible for forensic linguists to develop a structurally similar method, one which would successfully allow corrections, successive precisions, and limitations in order to obtain reliable and valid results step by step, instead of trying to reach the results in one single global step. This would mean that texts, particular styles of linguistic behavior, and strategies would be contrasted in a sequence of specific perspectives. This would have to be done on all levels of linguistic analysis, phonology, morphology, syntax and semantics, and also in text linguistics, sociolinguistics and other fields.

For this method, adapted from other areas of linguistics, I have introduced the term “linguistic differential diagnosis” for FL, by analogy to the medical term “differential diagnosis”. The basis of the analogy here is formed by the methodological implications. The medical procedure means the successive delimitation and exclusion of symptoms of different diseases in order to reach a final diagnosis on a particular disease in question. This method seems to reflect the essence of FL authorship attribution. The same holds for FL, where an exclusion of alternative authorship has to be achieved in a case where there are possibly several authors of an anonymous text. It goes without saying that author-identification, which is also employed in this activity, is a somewhat ambitious term. FL can hardly ever produce solid proof in the sense of the natural sciences for the authorship of an anonymous text. A final result that reaches the highest point on the probability scale – in German, “mit an Sicherheit grenzender Wahrscheinlichkeit” (‘with probability bordering on proof’), is extremely rare indeed (see Kniffka 2003b). I have never used it in an expert opinion in a case of authorship attribution in more than 30 years of expert testimony.

Thesis 13: Interdisciplinary perspectives

As has been mentioned previously, new dimensions of interdisciplinarity are urgently needed. It seems that this could be operationalized on three

different levels:

- (1) Interdisciplinarity between an area of practical work A and practical work B (German: “Praxisfeld A und Praxisfeld B”);
- (2) Interdisciplinary exchange between *applied science* A and applied science B; for example, between applied/forensic linguistics and applied criminology;
- (3) Interdisciplinarity between *basic research* A and basic research B; for example, between general linguistic research (for example, on questions of idiolect) and basic research in criminology.

There has been a rather fruitful exchange on level (2) in the past, in particular between FL authorship attribution, forensic phonetic speaker identification and forensic handwriting analysis. This results, in part, from the fact that in some cases the nature of the data is such that each of the three neighboring sciences has its share and can contribute to a solution. Beyond that there has been little systematic interdisciplinary exchange, which is an urgent desideratum for future research.

Thesis 14: Documentation of research and casework

One of the most important practical desiderata of research at present is a global documentation of the various activities of expert testimony that has been given to date in/for German courts by forensic linguists. It should be documented in a thesaurus of FL work. Even in the BKA, which is technically very well equipped for this, there is no expansive documentation of this kind. Basic questions, such as the following, are not dealt with: Which cases of extortion and blackmail with written texts have occurred to date? Who gave expert testimony when, where and in which case? What was the outcome? Which linguist is working on which cases concerning authorship attribution, libel, slander and defamation cases and so on? Which cases actually made it to court? Which have never made it to court on the basis of linguistic expert testimony for the parties involved?

Another consideration is that if there is to be improvement on the professional level of FL – meaning the actual performance of FL experts – one would have to introduce “expert testimony on expert testimony” in the form of quality control and peer review.

In addition, an empirical inquiry should be made concerning the success of linguistic expert opinions in courts. It is very rare indeed that a

court will notify a linguistic or other expert on what the actual outcome of the expert opinion was. It would be very important to have this type of feedback, together with the percentage of "hits" or successful expert opinions. On this basis, it would perhaps be possible to arrive at generalizations on stringency and argumentative potential, even of particular linguistic characteristics of certain methodological concerns and precautions of linguistics experts, and also the impression that linguistic experts made in court. The reactions of the various parties (including the courts and legal practitioners) would also have to be analyzed and empirically documented. This analysis and documentation should apply to all linguistic cases, whether or not the courts had requested linguistic expert testimony. The documentation also should account for cases in which a linguistic expert opinion was satisfactory for the purpose or the general purposes it was expected to serve from the perspective of the judiciary, even though perhaps unsatisfactory from the linguistic perspective proper. Also, any expert opinion that has been well executed and drafted from a linguistic perspective proper, but has failed to serve its purpose in a particular case in real life should be noted. Cases should be documented where linguistic expert opinions were *not* able to supply sound proof on data that allowed the court to resolve a particular case in a satisfactory manner. Also, cases where a criminal investigation was brought to a successful conclusion by data supplied by FL as an auxiliary science should be thoroughly documented. These data collected for the German situation could be compared to those of other countries and legal systems.

4

“Shibboleths” as Data of Linguistic Behavior

This chapter is based on, and is partly a restated version of, an article entitled “SCHIBBOLETHS. Philologische Bestandsaufnahme und Gesichtspunkte zu ihrer soziolinguistischen Analyse”, which appeared in the journal *Deutsche Sprache* 1991: 160–77 and in which a general linguistic and a sociolinguistic perspective on the analysis of shibboleths as everyday verbal behavior is discussed.

This chapter extracts, restates and adds some systematic and methodological points from the 1991 paper. It deals with the applied linguistics perspective of the relevance of the concept of “shibboleth behavior” and elaborates on what it may mean in forensic linguistic authorship attribution and related areas.

The basic underlying hypothesis of what follows is that it is worthwhile rephrasing the main systematic and methodological issues of the analysis of shibboleth as a particular form of behavior, and that the gist of the 1991 paper in German is not at all outdated from the perspective of today’s applied linguistics. A critical reflection of some of the basic concepts of shibboleths and shibboleth behavior discussed here suggests that they may be of use in several areas of applied linguistics, but that a much more sophisticated and subtle methodology is yet to be developed to capture real life shibboleth behavior adequately, with the precautions and general considerations that have been outlined in mind.

Introduction

Judges, 12: 5–6: And the Gileadites took the passages of Jordan before the Ephraimites: and it was so, that when those Ephraimites which were escaped said, Let me go over; that the men of Gilead said unto him, Art thou an Ephraimite? If he

said, Nay; Then said they unto him, Say now Shibboleth: and he said Sibboleth: for he could not frame to pronounce it right. Then they took him, and slew him at the passages of Jordan: and there fell at that time of the Ephraimites forty and two thousand.

The text represents an entire type or genre of examples which seem to have a considerably high frequency of occurrence in everyday linguistic behavior – and which matter in a most sophisticated, sometimes tragic way. A few other examples well known in the literary tradition can be mentioned in passing: The *ciceri* of "I vespri Siciliani" (30 March, 1282), which were pronounced improperly (and to their disadvantage) by the French as /'siseri/ or /'ʃiseri/ rather than /'tʃitʃeri/. The Dutch words *schilde* and *vriend*, again articulated improperly by the French, are similarly well-known examples.

The German counter-espionage in the Second World War is said to have used the German word *Eichhörnchen* ('squirrel') in the native German articulation. Remarkably enough, the American counter-espionage is said to have used the word with the same denotative meaning ('squirrel') in the one-syllable US east coast articulation (equally difficult for non-natives) as a code word.

In a slightly more sophisticated experimental design during the Palestinian conflict in Beirut, the following – even more fatal – usage of shibboleths has been reported: military posts showed a tomato to people trying to cross the border in order to find out their linguistic and ethnic identity, a tomato being an everyday object that has a variety of expressions, articulations and shades of meaning in Arabic dialects. At stake was whether the "in-group" or the enemy articulation was being presented. On the other side of the fence, this test had of course been known for a long time, and the articulation required was presented properly. Then the military watchman continued the examination: "So, this is a tomato, right?" – "Yes" answered the testee, this time in his authentic vernacular way of speaking – and suffered the same fatal result. The word *yes* shows a similarly complex regional variation of local Arabic dialects.

It seems plausible to assume that the examples quoted above have been handed down to us by literary tradition because they concern the question of life or death. It also seems plausible that examples that do not pertain to such serious, potentially fatal decisions have not been handed down in equal number and quality. There is ample evidence in everyday linguistic behavior today that the processes of verification

and/or identification of speakers of a less serious and life threatening type have not been handed down, because they are trivial everyday acts of linguistic behavior. Verification and identification of (1) individual speakers, and (2) of members of groups of speakers in speech communities of various natures, are such trivial and extremely frequent everyday facts of verbal behavior that they do not need mentioning specifically. One presses the buzzer to open a door if one recognizes the voice of the person that rang the bell. It is of critical importance whether one speaks with an accent and which particular accent one is using, if one wants to rent a flat, obtain a bank loan, apply for a scholarship, receive an official invitation and so on. All this shows that not only academics – or linguists, for that matter – but also ordinary people are extremely interested in and make use of identifying features in their everyday communicative behavior. It seems surprising that this has received very limited scientific interest in linguistics itself, although the matter of shibboleths as such has found some reception (cf. Coulmas 1985; Oksaar 1967; Sornig 1985). A detailed and systematic linguistic account of shibboleths and similar communication behavior has not, however, been given to date.

The phenomenon of shibboleths, as with those of idiolect, disguised language behavior and others, belongs to areas of research frequently stated as very important but hardly ever analyzed in systematic depth. It seems hardly justifiable that other sciences, such as psychology, criminology and law, have shown a greater interest in shibboleths than has been demonstrated by linguistics. This is a remarkable deficit in our field, and one that should be remedied. It may have contributed indirectly to the fact that practitioners in various authorities and institutions, such as immigration offices in various countries, have used shibboleths and linguistic behavior similar to shibboleths in a way that can hardly be called scientific when they try to determine whether an unknown and non-identifiable person seeking asylum is a speaker of a language X (and hence a citizen of country X). These procedures are entirely non-scientific, non-legitimate and not very efficient. Being a native speaker of language X is not at all identical with citizenship in country X (for a critical evaluation of German cases see Kastenholz 1998; Australian cases are discussed in Eades 2005; Eades and Arends 2004).

This chapter supplies some preliminary theoretical and empirical explanations of shibboleths and similar linguistic behavior. First, some brief philological comments will be made about the origin of the term, a few methodological ingredients of the way it has been analyzed, and some structural characteristics of the shibboleth examples handed down to us, which can be of relevance for a (new) *linguistic* description. The

chapter then discusses some descriptive features of shibboleths and illustrates the problem as such. The following section outlines some of the theoretical and empirical implications that the concept may have for the analysis of verbal behavior, a linguistic working definition and elaboration on some heuristic domains of occurrence of shibboleths. The chapter closes with a summary of the main findings and gives an outlook on potential future research and applications in an exemplaric fashion.

Origin and history of the concept

In the text Judges 12: 5–6 quoted above, the wording "for he could not frame to pronounce it right" is of linguistic interest. Martin Luther, in his German translation of the bible, renders this wording as (Judges 12: 6) "... weil sie es nicht richtig aussprechen konnten" ('because they could not pronounce it right').

This wording is indicative of some of the problems represented by shibboleths and their methodological status as verbal behavior. Instead of making a (limited) statement about the data of verbal behavior – which are actually there – a general statement concerning the speaker's *ability* is being made: the speaker "was not able to pronounce it right." This is a classic mistake, so to speak, occurring in forensic linguistic expert opinions of authorship attribution even today: overgeneralization of a few given data for a speaker's total behavior (and ability, see below).

It is possible that some translators assigned the Hebrew verb form JKJN (3. sg. imperf.) 'he put' – which belongs to the root KWN, originally meaning a causative verb 'to make stand' – to the root JKL 'to be in a position, to be able to'. The translation of the Septuaginta is somewhat more precise in the next illustration: in the Greek verbal expression of Judges 12: 6 "καὶ εἶπαν αὐτοῖς 'εἶπατε δὲ σύνθημα' καὶ οὐ κατέϋθυναν τοῦ λαλεῖσαι ἠούτως", a verb is chosen which is glossed in LSJ as 'to succeed in doing'. The Greek verb is derived from the adverb *euthýs* ('straight ahead'; 'immediately'); the notion may be comparable to the modern American expression 'to get something *straight*'. The point here is that the translations "he did not succeed to pronounce it right" and "he was not able to pronounce it right" make an important difference.

The signs <šīn> and <sīn> in the word "shibboleth" merged in writing. The Ephraimites have solely the pronunciation /sīn/; the other tribes have both pronunciations. In classical Greek, there is no such sign (and no phoneme /š/). The Greek translators tried to solve the problem by translating the word meaning(s) for "shibboleth" – for example,

stachýs ('ear' as in 'ear of rye'), or by choosing a translation equivalent for the (function of the) concept: *sýnthēma*, 'anything agreed upon, pre-concerted signal, password, token or sign' (LSJ).

The English and the German translations have yet another sense: 'for he could not frame to pronounce it right' and 'weil er es nicht richtig aussprechen konnte' both suggest that some *material substance* of linguistic data rather than a *way of speaking* is at stake here, which the classical Greek translation of the Septuaginta clearly indicates by the adverb *hoútōs* ('in this way').

This illustrates yet another aspect of shibboleth behavior: we act as if peculiarities of pronunciation and language use can be described for individual linguistic items rather than the way in which they are used as a whole. The translators give the impression (as do some linguistic experts giving expert testimony on authorship attribution, even today) that some material characteristics of linguistic items exist in relation to speakers (users of a language). There are no static or permanent material values that could be ascribed to particular linguistic items – for example, words – or which would have to be assigned to these in a static fashion. Rather, they represent dynamic entities of language behavior.

Descriptive characteristics of shibboleths

A closer look at the old records of shibboleths permits a first step in a proposed revision of the notion of shibboleths, and also of the notion of "speaker-specific features" used in forensic linguistic authorship attribution. In addition to the statements made above, the following characteristics concerning the status and the evaluative potential of shibboleths by linguists can be made:

(1) The standard set of examples of shibboleths, historically recorded or actually still used today, suggests their validity in determining the group a speaker belongs to. These records do not make any generalizations about the significance of the procedure as such, undecided cases, intersubjectivity of examples, and definitely nothing about the discovery procedures of the testers. The overall behavioral syndrome seems to have been *in dubio pro morte*, at least in the cases quoted at the beginning of the chapter. Also, nothing is said about how these shibboleths were selected by testers. One could get the impression that testers think there are unlimited shibboleths in every language. By this very fact, the concept of an instrument for determining somebody's "linguistic identity" by a token or a password would lead itself *ad absurdum*. If all words of a

language were shibboleths, there would be no need for a concept of this kind.

(2) It seems as if there was a necessary, inseparable relation between a (native) speaker and the particular linguistic product/word uttered in any given usage of a shibboleth. This assumption would also imply that shibboleths are of a static nature and do not allow any variation.

(3) All examples of shibboleths recorded to date seem to refer to spoken rather than written language. It would seem logical that there would, in principle, be as many salient items and areas of shibboleth usage in written forms of language as there are in spoken language. This holds even more firmly if one assumes a continuum of two dimensions of written and spoken varieties, as suggested by Koch and Österreicher (1994). Features of written language may refer to author-specific textual behavior and also to features of orthographic (Kniffka 2003a; 2003b) and handwriting behavior (Michel 1982).

(4) The examples of shibboleths mentioned above almost exclusively relate to the fields of phonetics/phonology and lexicon. It is almost a *topos* in forensic linguistics that morphology and syntax do not show shibboleths and idiolectal features to the same extent. Some people say that morphology by its very nature does not allow any, although there is scant evidence for such a hypothesis. In fact, it is quite likely that morphology and syntax do have a sizeable proportion of shibboleths and idiolectal features. They may be fewer in number than in phonetics/phonology and the lexicon, but this should be thoroughly proven empirically rather than dismissed as a mere possibility.

(5) The most important descriptive characteristic of previously described shibboleths is that they pertain to *groups* of speakers and speech communities rather than to individual speakers. In other words, shibboleths most generally are related to sociolects and dialects, rather than idiolects.

As a concept, shibboleths run the risk of being overstated or used as a *petitio principii* in linguistic theory, as with the usage of the concept of idiolect which has also been overstated.

As remarked by Edward Sapir in his seminal article "Speech as a Personality Trait" in each case the social norm(s) in existence would have to be proven in "a painstaking study of the social norms of which the individual phenomena are variants" (Sapir 1927: 54). This holds for speaker-specific features of all types as well as for shibboleths.

This implies that shibboleth is not some kind of "passepartout notion" for all kinds of features of utterances or forms of communication behavior. The descriptive characteristics stated above are (in general) of

a restrictive and/or negative type, suggesting the very opposite; that is, there is a necessity to have a strong empirical indication in each case and further terminological differentiation. Sapir (1927) hints at the reasons why idiolect is a somewhat neglected concept in linguistic research. It does not, however, indicate that the phenomena, the area and the concepts are, as such, unnecessary, unimportant or exotic as objects of linguistic inquiry.

Salient structural features of shibboleths

The catalogue of descriptive characteristics given in the previous section contains several empirical and theoretical elements that need to be taken into account when giving a meaningful description and explanation of how shibboleths work. Obviously, the catalogue is incomplete and somewhat arbitrary concerning the selection of data and the size. It also is theoretically incomplete in that it neglects, for example, the data of pragmatic ingredients of communication. Below, a selection of elements is given as an intensional definition of the pragmatics of shibboleths.

It should be mentioned in advance that, in addition to the list of features given on p. 92f., the “attitudinal component” of such speech events (see pp. 94–7) has to be taken into consideration, which, unsurprisingly, has escaped philological and linguistic attention.

A description of shibboleths necessarily contains attitudinal data relating to the sender and the receiver (including the differentiations in addressee and audience) that are viewed as components and introduced by Hymes (1962). In short, it is not sufficient to talk about shibboleths as linguistic entities (individual material items). One has to include the attitudinal features and values held by the sender and the receiver (or the victim in some cases).

The shibboleth speech situation

The components of a shibboleth speech event are basically defined using Hymes’ and Gumperz’ concept of “components” of the Ethnography of Communication (Gumperz and Hymes 1972; Hymes 1962).

In a heuristic taxonomy, the following features of shibboleth speech events can be found:

(1) The shibboleth speech situation is determined by a “natural experimental design”. This means that it is itself a natural experiment performed by lay people as examiners and examinees, or testers and test takers. It is natural also in the sense that the examinees do *not* know it is

an examination situation, its impact and the elements of shibboleth testing as a whole.

(2) The interlocutors are thus essentially characterized by an asymmetry of knowledge: the examiners have a goal and use a method (or several methods) which the examinees do not know. This is a constitutive condition of the test. It could not be performed without this condition being fulfilled.

(3) For the components "manifest message form" and "manifest message content", it is a constitutive condition that there is at least one element of the utterance that has a *prima facie* or overt meaning as well as a hidden or covert meaning. One could call this the "synthema status" of the verbal behavior that is designed as a shibboleth.

(4) It is also a constitutive condition of this speech event that the interaction is receiver/examinee oriented in terms of the hidden or covert goal, and, at the same time, controlled by the examiners/senders.

(5) Grice's (1975) sincerity condition is violated by shibboleths *ipsa natura*.

(6) The most salient constitutive condition for the description of shibboleths as a speech event is a situation of *intercultural contact* where there is a minimum of cultural and linguistic diversity between the interlocutors. It is necessary, at least, that the examiners are under the impression that such a diversity (of a higher or lower degree) exists.

This condition can also be seen in other domains of shibboleth behavior in which the superordinate aim is not to find out who the partner is but to find out "What do you know about me? What do you think?", as in spontaneous or institutionalized forms of cultural contact, for example in foreign language instruction.

(7) As pointed out above, shibboleth behavior not only means idiosyncrasies of linguistic behavior, but also the attitudes of the senders/examiners and the receivers/examinees towards the shibboleth. Without such an attitudinal dimension, the speech event shibboleth does not take place. The attitudinal data were not mentioned as descriptive characteristics in the preceding section of this chapter because evaluation data such as attitudes, opinions, stereotypes relating to language and linguistic behavior can only be described on the basis of actual utterances and the linguistic data they contain. Undoubtedly, this is even more difficult with examples handed down to us in the long course of history. One has to determine in detail what constitutes the "psychological reality" of the attitudinal data of the record, and what is the "psychological reality" of the attitudes of interlocutors in the actual event.

Postulates of a more adequate theoretical description and empirical analysis of shibboleths

As previously stated, shibboleths are an important and frequent element of everyday linguistic behavior of speakers and speech communities. It is also undisputed that not much is said about the theoretical and empirical status of shibboleths in linguistic analyses. To describe shibboleths as “features of linguistic behavior” is not sufficient. Sornig (1985) has merits as a first attempt of a brief linguistic discussion of the phenomenon, which does not merely repeat the old set of examples. He introduces new data, the validity of which will not be discussed here.

It makes little sense, however, to simply list “phonematische Eigenheiten, die Sprechergemeinschaften voneinander unterscheiden” (‘phonemic features that distinguish speech communities from one another’) (Sornig 1985: 78) as shibboleths. This would not even yield a working definition.

The list of examples of shibboleths given by Sornig (1985: 178ff.) illustrates the horizon of the general problem proficiently. All the items mentioned – “das bolognesische /θ/, das Leipziger /g̊a/, das steirische /ou/ und als ‘lexikalische Elemente in dieser Funktion’ ... Verben des Arabischen für ‚sehen‘ (*raʿa/šaf*); DDR-deutschen *Volkswald* und *Eier-Soll*” – can be understood and “processed” as shibboleths *only* by someone who knows the particular languages or varieties of languages and also the speech communities reasonably well, and is able to judge the range of deviation and status of each variant from the particular social norm in relation to Sapir (1927).

If one does not know Italian or German and their local variants well enough, the “Bolognese /θ/” and the “Leipzig /g̊a/” do not mean anything, let alone stand as potential shibboleths. The /θ/ could be interpreted like the <th> in English, for example, and someone who is not acquainted with the details of the variety of German spoken in some parts of Austria would not know what to think of the “Styrian /ou/”. There are additional complications, however. The word *Volkswald* (in the former GDR) as an isolated item, similar to the names *Volksgarten* (in the city of Cologne) or to *Volkspark* (in various other cities in the former FRG), can and (equally reasonably) cannot be taken as a shibboleth. Maybe the expression *Volkseigener Betrieb* (VEB) would be a better example for a GDR-marked shibboleth.

The list of examples in Sornig (1985) also contains heterogeneous data. This could be an argument that there are in fact very different kinds and classes of phenomena to be distinguished. At any rate, it would have to be determined which empirical predicates can be stated

in which quantitative and qualitative terms concerning the group of speakers envisaged. Are the examples *Volkswald* and *Eier-Soll* examples of shibboleths in or for the speech community GDR (whatever a solid empirical definition of this would be)? Are they some kind of indication of a variety that might be called "GDR German"? Would such a usage of words describing a variety of language be feasible and justifiable – since it would be yet another dimension of shibboleth behavior? Is it empirically safe to state that a particular usage of a lexical unit or word can be considered a shibboleth *per se*? Are the two Arabic verbs that Sornig (1985) refers to really comparable here, and to which varieties and speaker populations do they refer?

There is one more point in the definition of shibboleth behavior that comes to mind. The examples quoted in Sornig (1985) are by no means the only ones that can be stated about the particulars of the Bolognese, Leipzig and other speech communities' ways of speaking. Obviously, if shibboleths are defined as dynamic entities of language behavior, one would have to take into account *all* the concomitant features and factors that make a particular feature just one part of a holistic picture of covariation.

As has been mentioned before, sociolinguistics seems to be the branch of linguistics from which data for an adequate description of shibboleths can be drawn. To name a few examples: first, the Labovian variationist paradigm (Labov 1966; 1970; 1972c; and others) and the Ethnography of Communication paradigm developed by Gumperz and Hymes (1972), Bauman and Sherzer (1974), Tannen and Saviile-Troike (1985) and others, and most of all Language Attitude studies as outlined by Agheysi and Fishman (1970), Shuy and Fasold (1973), Cooper and Fishman (1975), and, last but not least, research in social psychology on "social markers in speech" (cf. Scherer and Giles (1979), and others).

In Labov's terms, the distinction of indicators, markers and stereotypes seems to be of particular relevance:

Changes from below begin as *indicators*, stratified by age group, region, and social class. At this stage, they show zero degrees of social awareness, and are difficult to detect for both linguists and naive speakers. As they proceed to completion, such changes usually acquire social recognition as linguistic *markers*, usually in the form of social stigma, which is reflected in sharp social stratification of speech production, a steep slope of style shifting, and negative responses on subjective reaction tests. Ultimately, they may become *stereotypes*, the subject of overt comment, with a descriptive tag that may be distinct enough from actual production that speakers do not realize that they use the form themselves. (Labov 2001: 196 f.)

It seems that shibboleth behavior has to be classified according to these distinctions. Stereotypes, which are established, socially real widespread patterns of evaluation applied by the speakers of a speech community, can act as shibboleths if the asymmetry of the examiners and examinees described above is taken into account.

It was mentioned that Language Attitude research is of critical importance for the description, evaluation and identification of shibboleths and similar linguistic behavior. It is the key notion that takes the own socio-cultural situation of the speaker, the attitude towards the speaker's own language and usage of language, and that of the addressee(s), and several entities of culture, historical tradition, into account.

The importance of Language Attitude research for the description and explanation of shibboleth behavior results in particular from certain complex facts. First, there is an acute realization of an attitude – namely, that of the examiners in a shibboleth situation. Second, there is the predisposed attitude of the examiners towards the (revealing) usage of the examinee(s). Third, there is the self-assessment of their own usage by the examiners and the examinees. As a fourth attitudinal category, one could perhaps include the basic assumption that only native speakers are able to produce an authentic way of speaking, and that the shibboleth situation actually has some validity for determining native or non-native articulation.

In addition to this complex set of shibboleths as language attitudes, there are several other elements of attitudinal categories that have to be accounted for. One is that, between different attitudes, stereotypes and evaluations, a considerable number of relations and overlapping connections seems to exist. Attitudes towards language(s) and linguistic behavior are intermingled and closely associated with attitudes towards ethnic, social, religious and other entities. Another consideration is that various interrelations between factual, observable linguistic behavior and self-assessment of it exist. Shibboleth behavior has to be described from the perspective that behavior assessed by people themselves and by other people and factual, observable behavior (whatever the instruments for measuring this exactly may be) are not identical.

The following questions point out a horizon that has to be taken into consideration (though they cannot be used as a check-list):

- Which “carriers”, potentials and features of recognition and identification of a speaker as a member of a speech community are accepted by whom in what way and are actually recognizable by whom?
- What is the consensus within a speech community in this respect for a particular item of shibboleth behavior?

- Which types of linguistic behavior are considered significant for which group(s) of speakers in the sense of established and internalized stereotypes (independent of their status in actual language behavior)?
- Which degree of linguistic exposure and background knowledge is necessary to recognize "Brooklynese" or "Lower East Side Talk"? Does one have to be from there, or have lived in the neighborhood, in order to be able to give a proper evaluation? Or is it sufficient just to have heard this accent once, maybe on the radio only?
- Is it sufficient, to quote some examples from German, to identify speakers from Westphalia just by their use of interjections such as *woll!* and *hömma!* (standard: *hör mal!*)? Can people from the area called Sauerland be identified solely on grounds of the fact that they articulate just about every word initial phoneme /g/ as [x] and pronounce the phoneme /x/ in word internal and word final position exclusively as [x] rather than as its allophones [ç] and [χ] as would speakers of (more or less) standard German? What and how much time does it take to produce such items on occasion? Which frequency of occurrence justifies the designation of a speaker's shibboleth behavior?

This brief list of questions, most of which have not been answered to date, indicates the amount of work that has to be done in the future to arrive at plausible hypotheses concerning shibboleths and speaker behavior.

A new perspective on shibboleths as linguistic behavior

Social-psychological research, such as Scherer and Giles (1979), may have an even larger impact on the empirical linguistic analysis and theoretical definition of shibboleth behavior. The papers by Brown and Fraser (1979) ("Speech as a marker of situation"), Giles (1979) ("Ethnicity markers in speech"), Giles *et al.* (1979) ("Speech markers in social interaction"), Laver and Trudgill (1979) ("Phonetic and linguistic markers in speech"), Scherer (1979) ("Personality markers in speech"), and all of which appear in Scherer and Giles (1979), seem of particular relevance here and will now be outlined. It seems that the practical empirical analysis and the theoretical concept of shibboleth behavior need to incorporate some of the findings gained by this research.

(1) A definition of shibboleth behavior has to take the semiotic status expressed in the so-called "Second Peirce's Trichotomy" into account

(cf. Feibleman (1946: 90) as quoted in Laver and Trudgill (1979: 2):

the icon, a sign which refers to an object by virtue of characters of its own which it possesses whether the object exists or not; the index, a sign which refers to the object that it denotes by virtue of being really affected by that object; and the symbol, a sign which refers to the object that it denotes by virtue of a law, usually an association of general ideas, which operates to cause the symbol to be interpreted as referring to that object.

The indexical function as stated by Laver and Trudgill (1979: 3) is the most interesting method to apply to “how speech identifies the speaker”. This holds for shibboleth behavior as well as for features of speaker-specific behavior discussed, for example, in forensic linguistic authorship attribution. All shibboleths, of any form, configuration, level and mode of representation, have an indexical function of one kind or another. The weathercock on a steeple is an index of the direction of the wind, the height of the mercury level in a thermometer is an index of the temperature. Phonetic and linguistic features *can* be indices for the characteristics of the way(s) in which a particular person speaks. As mentioned before, there is no direct parallel between these three classes of data and, consequently, they do not provide an exact parallel in the kind of measuring involved.

For the linguistic entities and indices, which undoubtedly have to be of a more dynamic nature, several distinctions have to be made, which is entirely an empirical question. First, phonetic (iconic and symbolic) and linguistic (purely symbolic) features will have to be differentiated. Second, within the realm of linguistic entities and features, there may be differences between layers and levels of communication, the code and the media involved. The various layers and forms of spoken language may need a different set of descriptive criteria than, for example, forms and genres of written language and also non-verbal communication.

(2) As has been repeatedly stated, shibboleths cannot be conceived as material items that can be defined as language items themselves. They can only be defined as a specific usage of linguistic items of one kind or another. This implies, among other things, that a linguistic description has to be widened into a holistic sociolinguistic description including ethnographic data of the situation, the interlocutors and the other components involved. In addition to these ethnographic data, language attitudinal data concerning the recipient/examinee and the sender/examiner of a shibboleth must also be described.

(3) The ethnographic components cannot be taken for granted in this analysis. They have to be specified empirically in terms of the modalities involved. This implies, for example, the fact that a continuous variation of language use is not perceived as such but, very frequently, as a dichotomy, as pointed out by Laver and Trudgill (1979: 23):

Sociolinguistic research also indicates ... that while linguistic variation is often continuous, perception of it is often dichotomous: listeners may thus not be sufficiently consciously aware of linguistic differences to be able to report on them.

We saw earlier, in the case of linguistic variability, where percentage of use of a given linguistic form constituted a social marker, that listeners' perception of a continuum of this sort tended to be dichotomous. Such a phenomenon is one basis for the stereotyping process in attribution.¹ (*Ibid.*, 28)

For example, one does not normally hear how strong an accent somebody has, but only that one speaks with an accent or not. Shibboleths can only be described proficiently by including this fact concerning the perception of the recipient. The question is not: "Is form or pronunciation x used at all?" but rather "What percentage of use of the form x exists for a given variable?"

(4) One of the reasons why shibboleths can only be defined successfully if the receiver (audience and addressee) component is included, is the need for certain distinctions of social markers as discussed by Laver and Trudgill (1979: 26 ff.). They distinguish "actual markers" and "apparent markers" and the latter category introduces yet another distinction of "misleading markers" and "misinterpreted markers".

"Actual markers" indicate a factual observable feature of speech. "Misleading markers" are given when a speaker pretends or fakes an accent (one that he does not have authentically, mostly in a social or sociolectal sense). A "misinterpreted marker" is given if, for example, a hearer wrongly interprets whispering conditioned by acute laryngitis as motivated by a conspiratorial situation and acts accordingly.

(5) Even these distinctions are insufficient. It is easily conceivable that speakers fake an accent that they do not normally have in their authentic repertoire – for example, an upper class or a New England accent – and that the hearer does not understand it as such but misinterprets it as a *British* upper class accent. This may be for several different reasons; for example, because the speaker is a non-native speaker of English and does not know how to distinguish between various types of

accent. In other words, disguised language behavior, including an accent, can be detected or decoded correctly as such by a hearer, but it can also be decoded in a factually inaccurate sense or be misinterpreted as an incorrect accent.

(6) Brown and Fraser (1979: 54) point to yet another distinction; the fact that there are more ambiguous than unambiguous markers. Ambiguous markers seem to be particularly frequent for the components, interlocutors and setting. Labov's /r/ in New York English may be a marker for socio-economic status and, at the same time, for the formality of the setting.

(7) As mentioned before, shibboleths and other speaker-specific features of various kinds do not, generally speaking, apply to the verbal behavior of speakers *per se*, but rather to the verbal behavior of speakers in specific communication situations, settings, text types and genres (in Hymes' (1962) definition). This seems to hold for written communication as well. Speakers can be said to show a specific or typical covariation or configuration of features only in regard to particular text types and genres (see Kniffka 1981; 1990a; 1992; 2000a; 2003b). This indicates that there is still an enormous amount of work needed if we are to provide a sound empirical description of typical, or perhaps prototypical, feature configurations for particular text types and genres.

What can be summarized for an empirical analysis of the salient features of shibboleths? If shibboleths can be defined only in regard to setting, interlocutors and genres, if one has to distinguish whether "actual markers" and "apparent markers", and, within the latter, "misleading" and "misinterpreted markers" are given, if interlocutors go by dichotomized perceptions rather than by different shades and grades thereof, three basic alternatives can be assumed:

(1) The concept of shibboleths and of speaker-specific features of verbal behavior is much more complicated, differentiated and multifold than has been assumed.

(2) It is much more difficult to supply a methodologically proficient empirical description.

(3) As a criterion for determining to which group a speaker belongs, shibboleths are much less significant and valid than has been assumed to date. It also seems that shibboleths should be distinguished clearly by speaker-specific features of verbal behavior because these require a much more solid empirical method and experimentation. There are no clear cases in the empirical analysis of real life cases in forensic linguistic

authorship attribution. It appears that Scherer (1979: 169) is absolutely right in stating that linguistic entities have to be described for their individual differences in terms of (1) relative frequency of occurrence; (2) variability; (3) periodicity of occurrence.

Scherer also appears to be heading in the right direction in postulating 4 classes of "personality markers in speech":

- Formal characteristics (word and sentence length)
- Semantic functions (types of reference)
- Syntactic functions (parts of speech, transformation types)
- Pragmatic functions ("sender states", for example "avoidant verbalization", "hostility", "immediacy", and others).

It seems that pragmatic functions are of particular interest for the description of shibboleths and speaker-specific features of verbal behavior, but they have hardly ever been analyzed systematically by linguists. It can be argued that this list can, and must, be supplemented by several different types and dimensions of linguistic features and parameters. It is an empirical question how long this list should be for which analytical scientific context. In other words, the social markers defined by Scherer (1979: l.c.), by Laver and Trudgill (1979: 3) and others – such as group markers, individuating markers, affective markers, social markers, physical markers, psychological markers and the linguistic entities relating to them in shibboleths and speaker-specific features – should be regarded as a question rather than an answer. The empirical predicates by which an actual behavior of a speaker can be described, and the general background knowledge of markers and features by a speaker, are still different categories to be analyzed. Speakers may know a variant to be more or less atypical for their own standard usage, but they may nevertheless use it to an extensive degree.

Working definition of shibboleths and domains of occurrence

The following preliminary working definition can be given: By "shibboleth" is meant a verbal (written, spoken and possibly also a non-verbal) interaction between interlocutors P and G in which P is interested and, subjectively, able to recognize, describe and classify G concerning her/his mother tongue background and status as a speaker of a certain speech community. P is the examiner/tester, G is the examinee/test

taker. P states a hypothesis about G's native tongue by intuitively comparing her/his own (knowledge of the) way of speaking with that of G. To claim to be able to find out anything else (other than the linguistic question of native tongue) is impossible from the very onset. There can be no guarantee that P can determine G's native tongue accurately or precisely.²

Basically, the hypothesis states that G's way of speaking is not identical with P's, which, in extent and quality, may correspond to several different levels and layers empirically. It may mean that G has a different mother tongue, that G is "not from here", that G is a "true", real, prototypical speaker of the local dialect, and many other possibilities.

It is also likely that G does not know the aim and method of P's endeavor and is not able to detect it in the interaction. If G notices that she/he is being tested, the shibboleth speech event may not necessarily be unsuccessful or a failure, but its validity and significance is highly limited (G could manipulate the language behavior accordingly, might use disguises and several other options).

For the definition and classification of shibboleth speech events, several other terminological and conceptual distinctions are relevant, not so much concerning the actual internal success of the endeavor but in relation to the linguistic peculiarities in question. Factual implications of shibboleth events that go beyond the scope of linguistics cannot be discussed here.

Of central linguistic relevance are the dynamic particularities of shibboleths as empirical predicates. A threefold distinction between (1) "shibboleths", (2) "pseudo-shibboleths" and (3) "semi-shibboleths" can be suggested.

(1) If P's assessment of G's way of speaking can be shown to be empirically valid, or if G actually uses a way of speaking that P perceives correctly as such, this could be called a shibboleth. It may be necessary to define a shibboleth by finding more data of the language usage of G than just the actual fractions of data tested.

(2) If G's way of speaking is misinterpreted, misunderstood and incorrectly evaluated by P, if P assigns peculiarities and characteristics to G's way of speaking that do not actually exist but are merely projections (artifacts) by P, one could call this a pseudo-shibboleth.

(3) A third type exists where P has a sufficiently clear impression of G's way of speaking and membership of a speech community, although the description that P gives is not valid and is insufficient in an empirical sense, be it that P assigns non-pertinent, inaccurate features (for

example, because she/he does not know the correct and salient features due to inadequate linguistic training). For this type of data constellation, the term semi-shibboleth might be appropriate, although the terminological distinction, as such, is not of critical importance here. The theoretical distinction and the empirical validation of the three different types of "facts of verbal behavior" are what matter most.

There are many examples of semi-shibboleths in everyday communication. A lay person may make an absolutely correct assessment that a speaker is from a certain geographical area: for example, the Sauerland in Germany, such as the late President of the Federal Republic of Germany, Heinrich Lübke, based on his regionally identifiable speech. Nevertheless, that person may not be able to identify the salient features correctly and sufficiently. It could also be that the assessment is subjectively and objectively correct, but that the lay examiner is not able to give a detailed description of the general facts (of the "Sauerlandian" way of speaking), let alone give a linguistically sound description and definition of the features involved.

The distinction of shibboleths versus semi-shibboleths does not correspond to that of linguistic (scientific) analysis versus folk taxonomy, however. It is a distinction that goes beyond assessment by laymen and linguists. The phenomenal ability of lay people to identify speakers from their neighborhood with astounding precision, first described systematically by Labov, along with the general fact that laymen do have such a remarkable sensitivity towards the perception of linguistic variation, would not be explicable in this fashion. Assessment of linguistic behavior by lay people and by professional linguists may both be valid and invalid as far as the "facts" are concerned. There is, at best, a gradual difference in validity, reliability and precision. The field of linguistics, to date, has developed no clues that can explain this amazing competence of lay people, also called "the dialectology of dialect speakers" (Klaus Mattheier) and the "competence of natural experts" ("natürliches Expertentum"), referring to lay people's assessment of anonymous authorship of texts (Kniffka 1990a).

The three-fold distinction of shibboleths, pseudo-shibboleths and semi-shibboleths is certainly not meant as a full-fledged satisfactory definition of the whole range of phenomena described by shibboleth events. It is a minimum differentiation that is necessary theoretically and empirically, and is also useful for practical considerations. It is a first step in the right direction and a heuristic means to arrive at a more comprehensive account of the phenomena in empirical and theoretical terms.

This also applies to the following non-exhaustive and accidental list of phenomena of shibboleth behavior. It is not meant to represent a clean-cut taxonomy, let alone a definite statement of whether a certain example is to be called a particular type of shibboleth or not. It is incomplete in terms of types and tokens, and is highly subjective in that the examples are almost all taken from the area of intercultural contact of German and Arabic.

The main purpose of the list of examples is to create sensitivity to the fact that such shibboleth behavior is widespread and exists, very probably, across languages and cultures.

(1) The phoneme system of modern Standard Arabic, the most important phonological processes and basic regional (dialect) variants are sufficiently well documented in semitic linguistics. What is the general shibboleth-type assessment that Arabs themselves use for their own language (Standard or Classical Arabic)? Native speakers of Arabic would say “Arabic is the language of the *ḍād*”. In a phonological sense, Arabs do not consider the glottal or pharyngeal fricatives, the emphatic consonants or any other special set of phonemes as the most prevalent and distinguishing characteristic of their language. Neither do they assess the velar /ʔ/, uniquely occurring in the word *Allah*, as a shibboleth of their own language.

Needless to say, they do not believe that their language sounds “guttural” or accords to whatever other impressions Europeans have about Arabic. For an Arab, Arabic is just “the language of the *ḍād*”. This sound is the prominent phonetic feature that makes Arabic what it is as a language, in the shibboleth-type view of the natives, in particular those of the Arabic peninsula.

(2) There are several phonetic features of German (alien and hard to learn for Arab speakers) which they could, but do not call shibboleth-type behavior, rather referring to them as particularities of German; for example, the difference between /ç/ and /x/.

Neither is the use and the extent of nominal compounding in German – which is, again, hard to grasp for Arabs learning German – anything they consider to be a language attitudinal shibboleth-like behavior. However, the sound sequence and syllable structure in the German word *fünfundfünfzig* (‘fifty-five’), and the mere existence of a word such as *Lotto Toto* in German, are real shibboleths of Standard German as viewed by Arabs.

(3) However, to the contrary, it is remarkable that there are certain phonetic features of German that have a shibboleth behavioral status

where not expected to, whereas others that are much more likely to be assessed as shibboleths from a scientific linguistic point of view cannot be described in this way. The "ich-Laut" in words such as *Licht*, or the affricate /ts/, as in *Zeit*, *Zug* and so on, do not seem to have achieved shibboleth behavioral status for native Arabic speakers who speak German as a foreign language. On the other hand, there are examples such as the articulation of velar nasal + voiced stop found in the German words *Hunger*, *Zunge*, *lange*, by which a native speaker of Arabic, even if a fluent speaker of German for decades, can be recognized as a non-native speaker quite easily. This is probably because there is no corresponding phonological rule in Arabic (that the stop is not articulated at all after the nasal /ŋ/) and also because the instruction of German as a foreign language does not commonly deal with such cases.

(4) A certain type of pseudo-shibboleths is exemplified by the non-native ways of speaking and articulation of one's native tongue in many quotations and proverbs. Customarily, such evaluations of non-native articulation – or, articulation of "the people across the fence", (see Kniffka 2002) – are pejoratively evaluated by natives. The examples here again show a continuum of various types and perspectives, such as the well-known *Il parla italiano com'una vacca spagnola* (which may be less pejorative than the supposedly original *comme un Basque*), or the German variants of the example *Wie ein Polak sprechen* ('to talk like a Pole').

A similar, yet different type of pseudo-shibboleth is represented by "speaking examples", which are frequently used as caricatures, also pejoratively characterizing another dialect of the same language, though in a phonetically rather inadequate, if not plainly incorrect, and exaggerated fashion. German people from the Rhine area caricature people from the Sauerland by mispronouncing town names: for example, *Meskede* instead of *Meschede*, and *Lüdenskeid* instead of *Lüdenscheid* (that is, with an articulation of /sk/ instead of /f/). However, in the local articulation of Standard German in the first example, there is actually no articulation /sk/, but a combination of /s/ and the voiced velar fricative /ɣ/. Thus, it is *Mesgede* rather than *Meskede* (the /sk/ being reserved for a mocking dialectal articulation).

(5) There are several other variants of shibboleth-type behavior, such as letting somebody say the name of a local town or illustrating by way of a joke an articulation in a particular dialect. Well-known in German is the example of *Gänsefleisch* ('goose meat'), a joke in the Cologne area for the dialectal articulation of standard German *können Sie vielleicht* ('could you possibly') in Saxonian. Several linguistic stereotypes are intermingled with ethnic, regional, social and other stereotypes. Each

country and each speech community seems to have some dialect variants of different levels of sophistication, which may be a true universal. The shibboleth types of behavior – for example, by speakers from Cologne, Berlin and elsewhere – towards the varieties of German spoken by the Ostfriesen (people in Northwest Germany), people from the Sauerland or from Saxony – are paralleled by those of the people of the Hejaz about the people from the Nejd in central Saudi-Arabia, or by all Saudis towards the variety spoken by the people of Hadramaut in Yemen and the south coast of the Arabian peninsula. In China, the people south of the Yangtse River share the same (mostly pejorative) assessment of the dialectal variant of the people of the northern shore of the Yangtse. What is important here linguistically is the fact that all these types of shibboleth behavior (concerning other dialects of the same language) seem to describe just one or two features at the expense of others, so to speak.

(6) An interesting example of a pseudo-shibboleth from levels other than phonetics and phonology is the use of the word *insha'allah* by Arabs, in particular in Saudi Arabia and the Arabian peninsula, and its evaluation by (mostly non-Arabic speaking) Europeans (Kniffka 1991; 1995). *Insha'allah* ('God willing') is a word with high text frequency in spoken and written language, in particular in everyday situations referring to future events. This use is quite similar to the way of speaking of older people in Germany, and the use of the formula *So wahr mir Gott helfe* ('So help me God') in an oath of office in Germany. Linguistically, *insha'allah* is used as a future marker or as some kind of an assertion of what has been said. On no account can it be interpreted as a "fatalistic restriction of what is said", which is a frequent misinterpretation by Westerners trying to interpret a levantinic attitude into this usage.

This is a substantial cultural misunderstanding. It led to the fact that American companies employed Arabic-English speakers who searched the texts of Arabic and English contracts for the word *insha'allah*, assuming that the texts where it occurred with higher frequency had a more "iffy" status than those where it did not occur (which caused the Americans to make the contracts at less favorable conditions).

(7) Finally, there is a shibboleth behavior referring to individual or idiolectal language use rather than that of groups or an entire speech community. For example: If someone in the US in the early 1970s said *Let me make one thing perfectly clear*, every adult native or even non-native speaker immediately knew that this was associated with the late President Richard Nixon. Similarly, if someone in Germany in the 1990s heard the phrase *In diesem unserem Lande* ('In this our country'), which

is not really very specific, incorrect or colorful an expression (merely somewhat peculiar), everybody in Germany knew immediately that it referred to former German chancellor Helmut Kohl.

Personal ways of speaking, idiosyncrasies, mannerisms and other types of peculiarisms, such as the two phrases quoted, establish relations to one particular speaker, and the speech community as a whole shares this evaluation of linguistic behavior. They are unambiguous in the sense that it is clear who the original author of this particular phrase was and that they used it with very high frequency. Nevertheless, such phrases are not helpful ways of identifying speaker-specific features, in laymen’s speaker recognition or even in professional linguistic authorship attribution. It is clear who originated the particular phrase. It is absolutely unclear, however, who the author of a particular usage of the phrase is after it has been used for a while and spread in a speech community. One could say that almost everybody can use it in this particular way.

Yet a different kind of semi-shibboleth may be represented by the reverse situation. Native German children in a 6th grade English class in Jeddah, Saudi Arabia, practiced a rather peculiar two-syllable articulation of forms like *looked* – /'lʊkɪd/ – and *cracked* – /'krækɪd/ – which could not stem from usage by the present teacher, the parents, other children or the English of the Arabic-speaking environment. What at first appeared to be an individual mistake, turned out to be the standard articulation of a former teacher from South Africa, where this articulation is quite widespread.

Summary

The definition and classification of shibboleths, pseudo-shibboleths and semi-shibboleths given above are intended as a basis for a future heuristic study of this neglected but fascinating area of linguistic research. Although the main purpose of the paper is to illustrate the size and nature of the phenomena that have to be taken into consideration, no claim is made that this is a theoretically satisfactory final classification of the phenomena envisaged.

It seems practical to use the term “shibboleth” (and also the terms “pseudo-shibboleth” and “semi-shibboleth”) (1) to refer to situations of language behavior that include an evaluation of linguistic behavior of others in intercultural contact situations, and (2) as speaker-group-specific or speech-community-specific language behavior. Shibboleth behavior is applied, produced and used by groups of speakers, and it also

refers to language behavior and evaluation of language behavior used by a group or groups of speakers (described above as examiners/testers and examinees/test takers).

Even if the empirical crystallization is that of a single utterance by one specific person in the “classical” test situation of a shibboleth, it refers to prevalent group-specific behavior data. Thus, the concept of shibboleth (behavior) primarily addresses and refers to a group of speakers – rather than an individual speaker. This distinguishes it from truly speaker-specific idiolectal features as it is often used in forensic linguistic authorship attribution.

As Sapir (1927) noted, group-specific features can sometimes be used to identify a particular speaker, not, however, as an individual, but as a representative of one group or speech community as opposed to another. This means that, even in theory, shibboleths cannot simply be used for the scientific purpose of recognizing or identifying an individual speaker or author in a definitive empirical manner. Undoubtedly, several illuminating aspects and heuristic indices can be created by shibboleths. However, they can never be used as satisfactorily empirically based speaker-specific features that are valid in forensic linguistic authorship attribution.

In an even more basic sense, shibboleths and similar language behavior are not just a theoretical construct of the armchair linguist, or mere anecdotal evidence that philologists might enjoy quoting, if for no other reason than for their “anciennité”. Shibboleths are a widespread, psychologically real and text frequent form or variant of linguistic behavior; an everyday routine of language behavior. It is not an exaggeration to state their importance for intra-cultural – and, more particularly, inter-cultural – language contacts. This has not been adequately taken into account by linguistics and adjacent sciences.

It is certainly legitimate to speak of shibboleths even in the traditional (item oriented) way. One should always keep in mind, however, that this term denotes only the tip of the iceberg or, more precisely, just the shade of the tip of the iceberg as perceived by a particular group of speakers. A scientifically sound description is an important task that still lies ahead.

There is no question that a “theory of shibboleth behavior” is an important desideratum, to which the field of linguistics can make a substantial contribution. Several basic theoretical implications, and even the notation, have remained somewhat unclear to date though. There also is a need for more expansive and thoroughly empirical analyses of particular data, which would allow an extensional definition of the

domains of occurrence. A solid catalogue of the most important and customary types of shibboleth behavior is a research desideratum.

Having stated the difference between shibboleths and true speaker-specific or speaker-identifying features of behavior, it should be stressed that research on shibboleths may indirectly be of great relevance for forensic linguistic and other empirical analyses of speaker-specific data. This is one of the desired outcomes of this chapter.

This page intentionally left blank

Part III

Ongoing Research in Forensic Linguistics in Germany

This page intentionally left blank

5

Libel, Linguists, and Litigation in Germany¹

This chapter reports on defamation and the role the forensic linguist may play as an expert in such cases. In particular, the chapter concerns an insult case at a German Superior Court in 1974, in which the author gave linguistic expert testimony. It discusses some methodological and interdisciplinary implications of the linguist's work. In this context, some classical misconceptions will be noted that (German) linguists and people in the legal field have (had) about each other's work. Finally, the development of the field of forensic linguistics in Germany will be sketched briefly and certain needs for future work will be outlined.

Orientation

As about six- to ten-year-old children in post-war Germany, we had a large repertoire of insults, mainly for other kids. One such insult – in fact, the most severe, the most insulting and heaviest – was *Deine Mutter ist ein Scheißochse!* ('Your mother is a shit-ox!'). When I cite this insult to my students, 20–28 years old, they smile, understand it, know it is used by kids, but that's it. They don't know anything more about it; for example, that uttering this as the final point, the climax, in a longer exchange of insults means asking for, and getting, a nonverbal response from the "offensee", usually a boy of the same age. That was understood and agreed upon by all parties involved. I have not heard it used by adults, although compounds with *Scheiß(e)* as first constituent are quite frequent in colloquial German spoken by adults (such as *Scheißkerl*, *Scheißhund*, *Scheißtyp*).

This age-graded usage is, in a nutshell, the problem linguists deal with when working on insult and defamation. In fact, the overall pattern of insulting behavior is a much more complicated matter: In such cases

people often do *not* agree on the status, the meaning, the illocutionary potential of an insult.

What can we learn from examples such as this? First, that every speech community, every age and peer group, every culture and every sub-culture has its own value system, range of verbal behavior to go by or not to go by, and quite a creative repertoire of phrases and words to get it across. What is important about the exchange cited above is that the speaker as well as the receiver, the audience and the addressee, are all absolutely sure that it is an insult, and that the intention is to make an insult and to provoke some nonverbal interaction.

We can also learn that the insulting party is, in fact, aiming at the most valued object of affection, the mother, although of course the target of the insult is a boy who is not attacked directly. Labov's (1966) research on verbal dueling and "sounding" illustrates the same point.

But there are a few other important points about this exchange of insults.

(1) One point is that the exact denotation and connotations of an insult change quite rapidly over time and age. Speakers who are five years older than these boys may not use it anymore; those 20 years older may not even understand it.

(2) The second point is that the subsequent age groups do not even remember why such an expression is/was so highly insulting at an earlier age. The phrases or words so insulting just do not "sound" that way anymore.

It is perhaps surprising that most people seem to assume that, throughout their entire so-called "adult" life, the denotative and connotative meaning of words, including insults, do *not* change and are fixed and/or invariant. The denotations and connotations of a word change considerably by adult age, however, and this makes the job of the forensic linguist dealing with defamation more difficult.

Linguists have not really looked into this nearly as carefully as they have into the first four years of language acquisition.

This marks a first desideratum of linguistic research on insults. We act as if there is one clear understanding of a sentence, an expression, a word used by an adult "average unbiased person", the "ordinary reader" or similar constructs. This is not so. In fact, there is a large variation of different readings, understandings and connotations. This first desideratum of research is some kind of a categorical imperative for linguistics, particularly forensic linguistics: "Start doing substantial research on all

kinds of age-, age group-, life cycle-specific language variation and change, including semantic change." In addition, a sociolinguistic account has to be given for social, racial, ethnic and religious variation. But even this is not the whole story. Different age groups have entirely different views about whether or not an expression is an insult. Speakers of the same age but from different social, racial or religious backgrounds show considerable differences in this respect. Even the same speakers show semantic variation in their language on different occasions. With this, we are already in the middle of what linguists must deal with in libel, slander and defamation cases. It can be seen that linguists have a substantial contribution to make to the analysis of such cases. In fact, linguistic work is of critical importance here, as one of the most prominent linguists in this field, Roger Shuy, noted (personal communication): "Defamation law is not at all that clear. If ever there was an area of law that needs linguistic help, it is this one." This chapter will give an illustration of this based on real life empirical data in a libel case.

While working on the expert opinion for the case on which this chapter is based, the question came up whether there are any expressions at all that *always* must be taken as an insult, and which would have defamatory meanings and status *per se*. An entire linguistics department looked for examples of defamations/insults *per se*, but did not really find any, except for some expressions with names such as *Sie altes Nazi-Schwein* ('You old Nazi-Pig') (which *could* be used, macabre as it would be, "jokingly" as an allusion). Most insults, even heavy insults, and most defamatory verbal items can be used with a sarcastic connotation, jokingly and in similar ways. It is very difficult, if not impossible, to exclude such a situational or linguistic context for an expression *a priori*.

This may suffice as an illustration of where we stand in the analysis of meaning and language use in forensic linguistics, and what the major tasks are when we describe defamation and insults.

The most basic postulate is that language (behavior) is to be described in sociolinguistic terms: it depends on who says what to whom, when, in which context with which intention and to which end.

The example given above is also an illustration of the fact that an insult is not an insult to everybody and all people in the same way. It may be so to some, but not to others. The specific weight of an insult also varies. In other words, an insult is received differently by people, depending on their age, race, sex, culture, experience or education, among other variables.

The example also shows that there are several elements that are difficult to describe and even more difficult to evaluate on a scale of

“insultingness”. It is widely understood among linguists that the semanto-pragmatic status of an utterance is of equal critical importance for the linguistic analysis as the wording itself. One implication of this is that if it is so difficult to deal with so-called non-forensic “clear cases”, how difficult should it be to deal with questions of meaning in forensic cases in which there is total disagreement about the basic facts of language usage and meaning by the parties involved?

One way to conclude this introduction would be to say that (1) it is extremely difficult, as forensic linguists, to deal with the meaning of an insult; (2) in the areas where we have the tools to analyze the insult, we still need much more research about insulting behavior in order to offer useful and complete results to people in the legal field.

German defamation laws

There are basically three sections² in the German Penal Code (“Paragrafen des Strafgesetzbuchs”, StGB) (see Table 5.1) that deal with insults, defamation and similar offenses.

Table 5.1 Defamation laws (excerpt) in the German Penal Code³

§ 185 Beleidigung

Die Beleidigung wird mit Freiheitsstrafe bis zu einem Jahr oder mit Geldstrafe und, wenn die Beleidigung mittels einer Tätlichkeit begangen wird, mit Freiheitsstrafe bis zu zwei Jahren oder mit Geldstrafe bestraft.

‘Section 185 Insult (libel and slander)

Insult shall be punished with imprisonment for not more than one year or a fine and, if the insult is committed by means of violence, with imprisonment for not more than two years or a fine.’

§ 186 Üble Nachrede

Wer in Beziehung auf einen anderen eine Tatsache behauptet oder verbreitet, welche denselben verächtlich zu machen oder in der öffentlichen Meinung herabzuwürdigen geeignet ist, wird, wenn nicht diese Tatsache erweislich wahr ist, mit Freiheitsstrafe bis zu einem Jahr oder mit Geldstrafe und, wenn die Tat öffentlich oder durch Verbreiten von Schriften (§ 11 Abs. 3) begangen ist, mit Freiheitsstrafe bis zu zwei Jahren oder mit Geldstrafe bestraft.

‘Section 186 Malicious Gossip

Whoever asserts or disseminates a fact in relation to another, which is capable of maligning him or disparaging him in the public opinion, shall, if this fact is not

Continued

Table 5.1 Continued

demonstrably true, be punished with imprisonment for not more than one year or a fine and, if the act was committed publicly or through the dissemination of writings (Section 11 subsection (3)), with imprisonment for not more than two years or a fine.'

§ 187 Verleumdung

Wer wider besseres Wissen in Beziehung auf einen anderen eine unwahre Tatsache behauptet oder verbreitet, welche denselben verächtlich zu machen oder in der öffentlichen Meinung herabzuwürdigen oder dessen Kredit zu gefährden geeignet ist, wird mit Freiheitsstrafe bis zu zwei Jahren oder mit Geldstrafe und, wenn die Tat öffentlich, in einer Versammlung oder durch Verbreiten von Schriften (§ 11 Abs. 3) begangen ist, mit Freiheitsstrafe bis zu fünf Jahren oder mit Geldstrafe bestraft.

'Section 187 Defamation (in the narrow, technical sense)

Whoever, against his better judgment, asserts or disseminates an untrue fact in relation to another, which maligns him or disparages him in the public opinion or is capable of endangering his credit, shall be punished with imprisonment for not more than two years or a fine, and, if the act was committed publicly, in a meeting or through dissemination of writings (Section 11 subsection (3)), with imprisonment for not more than five years or a fine.'

Considering the differences between the German and the US laws pertaining to defamation, there are five facts that should be taken notice of:

(1) "Defamation" is used (1) as a cover term for the offenses that are to be dealt with here (insult, libel, slander, malicious gossip and defamation), and (2) as a technical term for a specific offense (corresponding to German "Verleumdung", dealt with in § 187 StGB).

(2) All three sections mentioned, §185, §186, §187, apply to offenses irrespective of the mode or the media involved. In other words, it does not matter whether an insult is made in writing, orally or in non-verbal form. To all these forms, section 185 applies indiscriminately, which means that there is no equivalent in German law corresponding to the distinction between libel and slander in US and British law.

(3) It should also be noted that German defamation legislation is *federal law*, as opposed to the US system, where there are different laws in each state. Defamation legislation applies to the nation as a whole. There are a few press laws concerning the "Landespresse" ('state press'), which apply in addition to federal laws.

(4) In the German Penal Code there is a basic difference between "Verbrechen" ('crime') and "Vergehen" ('offense'). "Vergehen" means that the sentence is up to one year in prison as a maximum.

“Verbrechen” means that the sentence is one year in prison as a minimum.

(5) All defamation laws are so-called “Antragsdelikte” that is, they are not “automatically” prosecuted by the state’s prosecutor (Staatsanwalt), but only upon special request by the offended party. There are many other legal aspects that would be of interest in comparing the German and the US systems, which are outside the linguistic concern of this chapter.

Figure 5.1 indicates, very roughly, the translation equivalents of the German and English terms. One terminological problem is that defamation is used as a cover-term, a hyperonym, of all the other terms. At the same time, it is used as the equivalent of section 187 (§ 187 StGB) “Verleumdung” (see above).

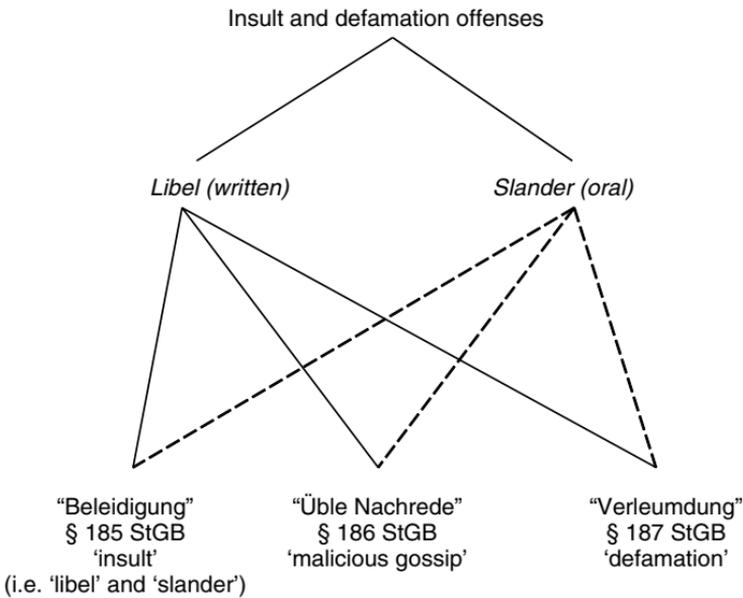


Figure 5.1 English translation equivalents of German insult and defamation offenses⁴

Basic statistics of insults and other offenses

Table 5.2 is an excerpt of the official Police Crime Statistics⁵ of 2003. It illustrates the rank of “insult” (as a cover term for “defamation”) as an offense in numbers of recorded cases compared to others like robbery or fraud. Insult is rank seven of the most frequently occurring offenses listed. There were 164,848 cases of insult recorded for 2003 in Germany.

Table 5.2 Share represented by attempts with regard to individual offenses (by offense group)

Key no.	Offense (categories) ⁶	Recorded cases	Including: attempts share		
			Number	%	2002
0100+	Murder and	2,541	1,721	67.7	65.7
0200	manslaughter				
1110	Rape and sexual coercion – Sect. 177 (2,3 and 4)	8,766	1,489	17.0	18.4
2100	Robbery	59,782	10,950	18.3	17.9
2220	Dangerous and serious bodily injury ⁷	132,615	10,141	7.6	7.6
2240	(Intentional slight) bodily injury	313,112	6,397	2.0	2.0
2300	Offenses against personal freedom	154,894	3,751	2.4	2.2
3***	Theft without aggravating circumstances	1,540,632	26,753	1.7	1.7
4***	Theft committed under aggravated circumstances	1,488,458	247,338	16.6	16.4
5100	Fraud	876,032	42,959	4.9	5.0
5200	Breaches of trust	50,897	0	–	–
5300	Embezzlement	102,565	702	0.7	0.8
5400	Document forgery	69,097	1,270	1.8	1.9
6200	Obstructing public authority and offenses against public order	122,079	442	0.4	0.4
6300	Aiding and abetting, obstructing criminal justice, receiving, and money laundering	28,459	1,815	6.4	6.5
6400	Arson and creating a fire hazard	30,308	2,267	7.5	8.8
6500	Competition offenses, corruption offenses, offenses committed in office	5,922	130	2.2	1.5
6710	Non-support	19,257	0	–	–
6730	Insult	164,848	0	–	–
6740	Damage of property	717,914	4,961	0.7	0.6
6760	Offenses against the environment (PC)	24,573	241	1.0	1.0
7100	Offenses against	31,301	240	0.8	0.8

Continued

Table 5.2 Continued

Key no.	Offense (categories) ⁶	Recorded cases	Including: attempts share		
			Number	%	2002
7250	supplementary criminal legislation in the economic sector Offenses against Aliens Act and the Asylum Procedures Act	153,882	5,823	3.8	5.1
7260	Offenses against the Weapons Act and the War Weapons Control Act	33,603	215	0.6	1.1
7300	Drug offenses	255,575	3,411	1.3	1.6
--	Total no. of offenses	6,572,135	379,563	5.8	5.9

It should be noted here that the authorities consider the number of unreported cases to be very high, since insult, malicious gossip and defamation are prosecuted upon request only (*"Antragsdelikte"*, see above). So the actual number of cases of insult, malicious gossip and defamation are probably much higher than around 164,000 per annum. In many instances, people refrain from going to court because of the risks of a case and for many other reasons, as will be explained.

In Table 5.3 the total number of offenses recorded for the years 1987–2003 is reported for cases of insult, malicious gossip and defamation.

As shown in Table 5.3, the number of cases increased from 79,552 in 1990 (within the borders of the old Federal Republic of Germany) to 164,848 in 2003. Most of these cases were verbal disputes and attacks in road traffic, in which instances of non-verbal communication of an insulting nature are of a particular high frequency. These include signs such as tapping the forefinger on the temple (meaning 'you are stupid'), and putting together the thumb and forefinger to form an O (meaning 'you are an asshole'), which are extremely frequent in German road traffic. A considerable increase is reported also for school scenarios, where students, or parents, disagreeing with the grade the teacher has awarded have resorted to all kinds of insults.

Taking all this together, one can say that there is an enormous increase in cases of insult, malicious gossip, and defamation in German courts, which confirms Roger Shuy's dictum (personal communication): "If linguists are needed anywhere, this is exactly the area where they are needed the most."

Table 5.3 Insults. Total number of offenses recorded for 1987–2003

Year	Total number of offenses	Cases solved		Number of suspects investigated
		Number	%	
1987	72,177	63,370	87.8	61,772
1988	78,227	68,492	87.6	66,529
1989	81,027	71,212	87.9	70,031
1990	79,552	69,765	87.7	68,258
1991	79,698	69,379	87.1	67,974
1992	83,737	73,144	87.3	70,303
1993	99,885	86,474	86.6	85,589
1994	103,771	90,736	87.4	89,555
1995	115,240	101,203	87.8	98,118
1996	117,629	104,774	89.1	101,792
1997	126,585	112,875	89.2	109,893
1998	130,051	116,585	89.6	113,189
1999	136,285	122,625	90.0	118,664
2000	152,282	136,486	89.6	132,489
2001	161,941	144,652	89.3	139,154
2002	162,884	147,116	90.3	142,319
2003	164,848	149,079	90.4	145,041

Source: Adapted and translated from Bundeskriminalamt Police Crime Statistics (2003)

Some differences between German defamation laws

The differences between the three offenses – Section 185, insult (§185 StGB “Beleidigung”), Section 186, malicious gossip (§186 StGB “Üble Nachrede”), and Section 187, defamation (§ 187 StGB “Verleumdung”) – are illustrated by the three drawings given in Figures 5.2a, 5.2b and 5.2c respectively.

Section 185 (insult)

There are three ways in which an insult sanctioned by this section can be committed:

- (1) By utterance of a “Werturteil” (‘value judgment’) of an insulting nature directly towards the offensee; for example, *You asshole!*
- (2) By utterance of a value judgment about an offensee who is not present) made in the presence of a third person, who is not the offensee; for example, *Mr. X is an asshole!*
- (3) By a “Tatsachenbehauptung” (‘assertion of facts’) damaging the honor and reputation of a person uttered in the presence of or towards the offensee; for example, *You have stolen my hat!*

Section 186 the issue is that ‘the assertion of facts cannot be proven to be true in court’ (“nicht erweislich wahr”).

The number of trials and sentencing according to Section 187 (defamation) is relatively low for two reasons. First, the fact asserted must be proven to be untrue. It has to be proven in court that the reverse of the assertion is true. This is difficult to achieve, and this is why in many cases the result is not a sentencing under Section 187. Second, Section 187 requires proof of the intention to make an untrue assertion. This requirement is even more difficult to be met in court. It is not easy to prove that somebody really intended to make that untrue assertion. In German lawyers’ jargon this would read “§187 tritt hinter §186 StGB im Wege der Gesetzeskonkurrenz zurück”, meaning something like ‘Section 187 is overruled by Section 186 by way of legislative conflict’. Cases are more often tried according to Section 186 than Section 187. Several kinds of complications and combinations are possible here. If an assertion of facts is, at the same time, an intended abusive characterization of an offensee who overhears the conversation, this is by the same token an offense according to Section 185 (insult), in this case insult and defamation both apply (“Beleidigung” and “Verleumdung” are committed “in Tateinheit”). These few hints must suffice for a preliminary sketch of the differences of the law from a layman’s perspective.

From yet another (ethnographic-sociolinguistic) perspective, the differences between the three offenses in the German Penal Code can be described, hopefully adding some clarification to the description given above. There is a cross-classification due to the distinction, (1) whether the utterance is made towards the offensee or to a person other than the offensee whether addressee and offensee are different or the same; (2) whether a “value judgment” or an “assertion of facts” damaging the reputation or the honor of the offensee is made. The offense to which Section 185 (“Beleidigung”, ‘defamation’, ‘insult’) of the German Penal Code refers implies that the offensee and the addressee are the same person in face-to-face communication or in non face-to-face communication (for example, by letter). The insult can be a value judgment (*You asshole!*), a gesture (the medium differences do not matter) or an assertion of facts. A value judgment, no matter if addressed to the offensee or to a third person, is always dealt with as an insult under Section 185. An assertion of facts is dealt with as an insult, if it is made directly to the offensee. If made to a person other than the offensee, it is dealt with under Section 186 (malicious gossip) or Section 187 (defamation). The specific feature of the offense described in Section 186 (malicious gossip) is that the assertion made is not proven to be true by the offender. If the

offender cannot prove that the assertion is true, he may be sentenced for malicious gossip. Section 187 (defamation) is a specialized case under Section 186. In this case, the assertion made is proven to be untrue, a fact that was known by the offender. To make an untrue assertion against one's better knowledge is dealt with according to Section 187 (defamation).

An exemplaric libel case: *Konkubine* in today's Standard German (1972)

In this section, a brief report about an exemplaric case of libel is given.

A tenant, Mr. A, lives in a house with several parties and keeps writing letters to the house owner complaining about his neighbor, Mr. X, and his girlfriend, Ms. Y. Both Ms. Y and Mr. X are personally well known to Mr. A. In evidence are dozens of letters from Mr. A such as those given in Table 5.4.

Ms. Y, the lady named in the letters, filed an action for an injunction against Mr. A concerning the use of the word *Konkubine*, which she considered a "libel" ("Beleidigung") and "verbal injury" ("Verbalinjurie"). The trial court followed her argumentation. The verdict stated that the defendant, Mr. A, must not call her or write using this term any more. The defendant, Mr. A, appealed to this in the next higher court on the basis of an expert opinion (not by a linguist, but by a philologist) that stated: "Konkubine ist keine Verbalinjurie, keine Beleidigung, sondern eine wertneutrale Bezeichnung." ('The word concubine in today's Standard German is not a verbal injury, but a neutral term.'). Ms. Y in turn ordered a linguistic expert opinion concerning whether the word *Konkubine* is an abusive, libelous and defamatory word in (1972) Standard German, or whether it is "neutral" (non-pejorative). The original question posed by the court was: "Ist

Table 5.4 Libel letters of Mr. A, June–August 1972

(1) *Herr X bzw. seine Konkubine haben wiederholt die Waschküche zu Zeiten benutzt, an denen es ihnen nicht zustand.*

'Mr. X or his concubine has used the laundry facilities at times when they were not allowed to.'

(2) *Herr X bzw. seine Konkubine haben wiederholt den Hund den Rasen vor dem Haus nassen lassen, so daß der Bewuchs auf Jahre hinaus gefährdet ist.*

'Mr. X or his concubine has let their dog wet the lawn in front of the house repeatedly so that its growth will be endangered for years to come.'

Konkubine ein Schimpfwort im heutigen Deutsch?" ('Is concubine an abusive word (swearword) in today's German?'). The linguistic expert testimony made some general remarks on linguistic expert testimony in libel cases, on language system and language use, and concluded that this use of language – *Konkubine* in this particular situation in the letters of the defendant – is by no means neutral, rather, it is a pejorative, derogatory usage intended to be libelous in the text in question. The superior court agreed with the argumentation given in this expert opinion and rejected the appeal (a detailed report on the case is given in Kniffka 1981).

Briefly summarized, there are at least two questions that the linguistic expert opinion had to clarify:

- (1) What does the word mean in a particular linguistic and situational context in terms of denotation and connotation?
- (2) A pragma-linguistic question: What does it mean if this utterance is used in a particular context? Furthermore: What does this particular use mean vis-à-vis the standard language use in 1972?

One general point to be made here is that it seemed, and still seems, necessary to explain to the German courts that we cannot simply refer to data of the language system but have to include data of language usage at a specific time. Linguistic expert testimony, particularly in cases concerning the meaning of a specific expression, has to take into account data of language system *and* language usage, which is not self-evident for many people in the legal field (and in linguistics, for that matter). I do not believe that we deal with a special kind or branch of linguistics here, but rather an area of applied linguistics that happens to refer to data of language system and language usage relating to court action and to matters *in foro*. It might be called "applied sociolinguistics" – which, in fact, it was called in the 1981 paper. The essential notion is that one deals with dynamic linguistic data of language behavior rather than with fixed notions of a word as a fossilized item in a dictionary.

What was done in the linguistic analysis of the disputed usage of the word *Konkubine* can be illustrated as follows:

- (1) First, we tried to give reasons that the terms "Schimpfwort" ('abusive term', 'insult'), "Verbalinjurie" ('verbal injury', 'defamation' and so on) are linguistically very difficult to define. This is not only because of inter-speaker variation; this is mainly because one and the same word can be used in very different senses. These terms imply norm definitions and value judgments that are not the domain of descriptive linguistics.

It is clear that, in the primary objective of the sociolinguistic analysis in such a case, one has to describe, in Labov's terms, "language/linguistic behavior in its social context" (Labov 1970). A set of operationalizable questions includes the following:

- Which denotative and connotative meaning(s) does *Konkubine* have in the letters from the defendant?
- Which features of the linguistic and the situational context are characteristic and decisive for this particular use?
- What is the current Standard modern German use of this word (1972)?
- What does this word imply (implicate) and what is the speaker intending to say if he uses it?
- With which other words does this word usually co-occur in German texts? With which can one combine the word *Konkubine* in an utterance? Which are the usual standard collocations?
- With which other word can this word be substituted in a given utterance?
- Which synonymic words exist, and which structural relations between them and the word in question can be described? What does the semantic field look like?
- In which particular communication situation can *Konkubine* be replaced by word X but not word Y, or vice versa?
- In which contexts, linguistic and situational, does *Konkubine* usually *not* occur?
- What general frequency of occurrence does the word *Konkubine* have in the spoken and written Standard German of today's adult speakers?
- Who uses the word; who understands the word (in relation to the entire speaker population of German 1972)?

It was interesting to see that even students of linguistics with a supposedly high sensitivity towards language and with an obligatory "Latinum" (nine years of Latin in school or a university examination) – more precisely, ten of the 120 students tested – were *not* able to decode the word. They had no idea what it meant.

This list of questions requires that a methodology that incorporates all the techniques and methods used in empirical linguistic analysis is necessary and applicable here; such as participant observation of language usage, document analysis (analysis of primary texts), concordances, corpora, dictionaries, and similar sources, systematic interviews and experiments with speakers of German. One important real life fact should be mentioned here (see also pp. 139–42 for the status of linguistic expert

opinions). Usually, the linguistic expert does not have enough time to administer any systematic inquiry (interviews, tests, experiments) when asked to give an expert opinion. In the case in question, only two tests were done with two classes of linguistics students, each with 60 participants serving more as general tests to gain a set of primary hypotheses. They were asked: "What do you associate with the word *Konkubine*?" This question concerned the inherent semantic features of the word and the relationships towards others. We also used questionnaires with closed questions of the type described in Tables 5.5a and 5.5b.

Table 5.5a Questionnaire 2a

Herr Meier hat Besuch von seiner Konkubine, Frau Müller.

Welche der unter (1) – (5) genannten Interpretationen trifft den Inhalt dieser Äußerung am genauesten? Kreuzen Sie bitte eine an.

- (1) Herr Meier hat Damenbesuch, nämlich Frau Müller.
- (2) Herr Meier hat die mit ihm nicht verheiratete Frau Müller zu Gast.
- (3) Herr Meier hat Frau Müller zu Gast, mit der er mal geschlafen hat.
- (4) Herr Meier hat Frau Müller zu Gast, mit der er schon seit längerer Zeit ein geschlechtliches Verhältnis hat, aber nicht verheiratet ist.
- (5) Herr Meier bekommt schon seit längerer Zeit Besuch von Frau Müller, mit der er nicht verheiratet, aber gut befreundet ist.

'Mr. Meier is visited by his concubine, Mrs. Müller.'

Which of the interpretations (1) – (5) matches the content of this utterance? Mark this X.'

(1), (2), (3), (5) are distractors, (4) is the correct interpretation:

'Mr. Meier is visited by Mrs. Müller, with whom he has an intimate relationship but is not married.'

Table 5.5b Questionnaire 2b

Gegeben sei der Satz:

'Herr Meier und seine Konkubine stehen vor dem Haus.'

Welches der unter (1) – (5) genannten Wörter entspricht dem Inhalt von Konkubine in diesem Satz am ehesten? Bitte kreuzen Sie dies an.

- (1) Bekannte
- (2) Geliebte
- (3) Freundin
- (4) Besucherin
- (5) Braut

'Mr. Meyer and his concubine are standing in front of the house.'

Which of the nouns (1) – (5) matches the content of the word *Konkubine* best in this sentence? Mark this X.'

(1), (3), (4), (5) are distractors, (2) (*Geliebte*, 'lover') is the correct answer.

The test represented in Table 5.5a does not allow for an evaluation of actual verbal behavior but instead asks for an attitudinal interpretation of a given utterance. Subjects were given an utterance and five different interpretations, four as distractors, which they had to match with the given utterance: The results can be found in Kniffka (1981: 624–31).

In the test represented in Table 5.5b, the semantically closest synonyms of *Konkubine* were looked for.

The linguistic expert testimony given by the author in this case can be briefly summarized as follows: first, we contested the scientific status of statements such as (in the previous expert opinion) “*Konkubine* is an abusive term/verbal injury” or “*Konkubine* is not an abusive term/verbal injury.” This was rejected since it represents a naive concept realism. It would not be the result of a linguistic analysis but a judiciary subsumption under a section of the Penal Code, which is not the linguistic expert’s job. In fact, in Kniffka (1981) it is stated explicitly that a linguistic expert opinion should be “clinically clean” of all evaluations, subsumptions, terms and reasoning of the judiciary. This seems a very important postulate indeed, since judges not infrequently react strongly to it.

Whether an actual use of a particular word is to be subsumed under “insult”, (Section 185 (§185 StGB)) or “defamation”, (Section 187 (§187 StGB)) of the German Penal Code is the judges’ business and theirs alone (see below). The other expert’s opinion also discussed the question of the quantity of linguistics to which the court and the judge can/should be exposed. This is difficult to answer, since courts and judges vary considerably concerning their background in and acceptance of linguistics.

Just two examples of this will suffice. Introducing the difference between “denotation” and “connotation” (denotative meaning and connotative meaning) in the expert opinion given was well accepted and received by the court. It was done by way of easily understandable examples, such as *Polizist*, *Bulle*, *Polyp*, *Greifer*, *Grünes Männchen* for “*Polizist*” (‘policeman’, ‘the fuzz’, ‘cops’ and so on, for ‘policeman’) each of which have different connotations, but the same denotation.

However, theoretical pragmatic terms such as *illocutionary force*, *speech act* and others, and linguistic argumentation were generally speaking not accepted by the court. We stated that connotations, as a rule, do not vary randomly, and explained by means of examples in colloquial German that we think that there are no words that are defamatory *per se*. We went on to explain that dictionaries and all kinds of collections that are used by

linguistic laymen as authorities for word meanings do not help much when a particular actual use of a word in a particular linguistic and extra-linguistic context is analyzed, and that a systematic empirical inquiry with native speakers is needed in any case. This was much more difficult to get across to judges. Probably from their own experience with dictionaries during their school days, dictionaries were generally considered highly authoritative in language matters. We also stated that etymological data, as a rule, are of little value for describing and explaining a particular modern language use. By way of examples of ongoing semantic change; for example, of spontaneous appreciation by German adjectives such as *stark*, *geil*, *krass* and so on – most judges were successfully persuaded that language (use) does change considerably. We went on to explain that Dell Hymes' (1966) definition of communicative competence – who says what to whom, when, why and how – may contain some more important concerns and data than the actual inherent semantic features of a word. Judges did not like this (they usually want more fixed quantifiable categories) but accepted it in the end, again on the basis of language data of a simple persuasive status such as: a form of address – say, *darling* or *I love you* – said by a man may mean different things depending on the addressee and the circumstances; if, for example, said to a policeman or to one's wife.

The specific question about the meaning of the word *Konkubine*, the systemic semantic state of the word, and the question whether it is neutral or not was the focus. On the basis of the empirical data (tests, interviews) outlined, we explained that the denotation of the word in the letters was 'a woman who has intercourse with a man without being married'. The empirical testing revealed that the denotation of *Konkubine* has a component 'sexual contact' as one of several inherent semantic features. In comparing the word *Konkubine* with words "Partnerin"/"Freundin"/"Bekannte" ('partner'/'girl-friend'/'acquaintance') and others, it can be seen from collocations that the word *Konkubine*, furthermore, has a clearly pejorative or derogatory connotation as opposed to the other words of the same semantic field in the German speech community of 1972 (for details, see Kniffka 1981).

Points of argument

What convinced the court eventually about the semantic analysis of utterances with the word *Konkubine*? More than anything else, it was that these were facts of language use in everyday German. Examples were much more convincing than any theoretical deliberations.

We argued in the expert opinion that, if the word *Konkubine* in the German contexts in question was entirely neutral, and if the word was

Table 5.6 Distinctive contexts for the word *Konkubine*

(1) Offizielle Einladungskarten eines Firmenchefs, der seine verheirateten und unverheirateten männlichen Mitarbeiter zu einer Festlichkeit mit Damenbegleitung einlädt:

Ich würde mich freuen, wenn Sie Ihre Frau/Partnerin mitbrächten.

**Ich würde mich freuen, wenn Sie Ihre Frau/Konkubine mitbrächten.*

[grammatically correct; text pragmatically unacceptable]

(2) Richter in einer Urteilsbegründung mit „wertneutraler“ Diktion, wenn es sich um eine mit einem Mann außerehelich verkehrende Frau handelt:

Herr X und seine Partnerin/Bekannte.

**Herr X und seine Konkubine.*

[grammatically correct; text pragmatically unacceptable]

(3) Nicht-akzeptabler Verwendungskontext:

Eine erwünschte künftige Schwiegertochter würde von ihren Schwiegereltern gegenüber dem eigenen Sohn, der erwiesenermaßen mit ihr schläft, mit großer Wahrscheinlichkeit nicht als *Konkubine* bezeichnet werden. In einer (ansonsten vergleichbaren) Situation mit umgekehrtem Vorzeichen – einer unerwünschten Schwiegertochter – ist Vorkommen des Wortes (in pejorativem Sinn) durchaus denkbar.

really a non-derogatory term, one should be able to substitute it for synonyms in contexts which indisputably are neutral or which are marked (at least) as non-pejorative. In Table 5.6, three examples discussed in the expert opinion are given to illustrate this.

Example (1): In official invitations, the president of a company inviting his married and unmarried male employees to a celebration with their partners would write without any problem, *I would be delighted if you could bring your wife/partner along*. He would hardly write something like *I would appreciate if you could bring your wife/concubine with you*. This would be unacceptable in a (festive, non-sarcastic) invitation, which seems to suggest that it has a non-appropriate pejorative and derogatory connotation.

Another linguistic context is referred to in example (2). In German court cases the judge (in cases with more than one judge, the Presiding Judge) reads the verdict at the end of the whole session. The “Urteilsverkündung”, a somewhat formal ceremonial act, would never refer to a man and a woman who are unmarried and live together as *Herr X und seine Konkubine* (‘Mr. X and his concubine’). It would read *Herr X und seine Bekannte* (‘Mr. X and his acquaintance’), *Herr X und seine Partnerin* (‘Mr. X and his partner’), or – this was the most decisive point for the court – *Herr X und Frau Y* (‘Mr. X and Ms. Y’) as the (most) neutral and unmarked form of reference to the two people.

Significant in this case is the striking *asymmetry* in the form of reference to the man and to the woman. Both were personally known to the defendant Mr. A (he had known them both by name for years), but he kept writing *Herr X bzw. seine Konkubine* ('Mr. X or his concubine respectively'). As noted above, the unmarked version would undoubtedly be "Mr. X and Ms. Y" in the contexts in question.

The defendant's use of the word *Konkubine* in this linguistic and extra-linguistic context conveys an irrelevant assertion of facts; namely that these two people have a sexual relationship without being married. The fact that the assertion is made repeatedly in writing to a third party suggests that the defendant (the writer of the letter) was not concerned with giving neutral descriptive information that there are two people living together next door. It seems instead that he wanted to use a speech act that one could easily identify in German as *jemandem eins auswischen* ('to disparage someone' or 'to get someone').

The gist of the expert opinion was: the particular language use cannot be termed as neutral or jokingly made. It is clear that the word *Konkubine* in the contexts in question has a pejorative or derogatory component of meaning. It refers to extra-marital intercourse, which, as the writer explains elsewhere in a letter, is not a decent thing in his view. The text suggests that the writer, the defendant Mr. A, apparently wants to make defamatory statements about the woman so described.

The lack of neutrality of the word *Konkubine* in this context can also be gathered from data of the situational context (see Table 5.6, example (3)). Prospective parents-in-law would hardly describe their daughter-in-law by using the word *Konkubine*, even if she had intercourse with their son. On the other hand, if she was an unwanted daughter-in-law, one could imagine the use of such a word.

In short, the following three points in combination convinced the court that *Konkubine* in these contexts was *not* used as a neutral, non-derogatory term: (1) the textual asymmetry in the reference to the two people involved ('Mr. X or his concubine, respectively' instead of 'Mr. X and Ms. Y', the name of the lady in question being known to the writer); (2) the text linguistic evidence by several examples of everyday contexts excluding an occurrence of the word *Konkubine*, as in invitations to a celebration and in court verdicts; (3) ingredients of situational contexts clearly neutral and non-pejorative, in which the word does not occur.

What can be learned from this type of linguistic expert opinion, concerning the "Verständnisnachweis" ('analysis of a particular semantic use of language')?

(1) Constructs such as “*der unvoreingenommene/unverbildete/unbeeinflusste Durchschnittsleser*” (“the ordinary reader, the average, uninfluenced reader”) can be abandoned without any real loss. Rather than working with such vague terminology, one should explain in detail how (and how many) data were gathered, what they mean for the particular case, which restrictions apply, and similar questions. The data have to be collected in a systematic empirical investigation from native speakers using all methods of analysis available in the social sciences.

(2) The most frequently named authority, the dictionary, can also be easily abandoned here since it is rather non-exhaustive, subjective – and is really a “history book” (not covering the actual stage).

(3) A forensic linguistic expert opinion should be “clinically clean” of all terms and routine argumentations of the judiciary, and, most definitely, of any subsumptions under a particular section of the law (of the German Penal Code). Linguists must not, and cannot, do the “*Beweiswürdigung*” (‘the evaluation of proof of the analysis’), which is exclusively the job of the judge. Judges are very sensitive, and rightly so, about linguists or any other expert witnesses trespassing beyond the limits of their work.

(4) A linguistic expert opinion should use as little linguistic jargon and terminology as possible. It should be written in such a way that the judge, as a linguistic layman, is able to understand it fully.

As a final note, one should also mention here the general development of changes in the social ethics of sexual behavior. What caused this case in the early 1970s in Germany would probably not lead to any legal action whatsoever nowadays. It is unlikely that anyone of the younger generation would feel insulted by somebody else stating that he or she is living together with someone else without being married (including reference to their sexual relationship). But, in the 1970s, a considerable number of the West-German population did not consider this as socially and morally acceptable.

What judges and linguistic experts need to learn about one another

It is important to note how the judiciary and linguists giving expert testimony view each other, and what linguistic experts should keep in mind when giving expert testimony. Some aspects are touched upon elsewhere in this book (see the Introduction and Chapter 2). The following remarks are not intended as accusations about either party. In most misunderstandings, both parties share the blame.

In the German system, judges have a remarkably powerful position concerning the evaluation and use of any expert testimony given. The German law says that the “*Würdigung des Sachverständigenbeweises*” (‘the evaluation and assessment of an expert opinion by the court’) is the judges’ job and theirs alone. This may be somewhat puzzling to outsiders to the legal system at first and one simply has to get accustomed to this. German judges are absolutely free to request a linguistic expert opinion, or not. They may feel “I am the linguistic expert and I don’t need anybody else”. If they order a linguistic expert opinion, they may interpret it as meaning the complete opposite to what was intended by that expert. This does not happen often, according to my own experience – but it does happen. Judges may reappraise the linguistic expert’s arguments at their discretion. In order to fully understand the judges’ role, the following questions should be answered:

(1) The first question one has to address is: “What are the canonical stages of analysis involving a linguistic expert opinion in a case like this?” There are three stages to be distinguished at least.

- (a) The “translation”, the rephrasing of a lay question into a linguistic scientific question that can be operationalized in linguistic terms;
- (b) The actual linguistic analysis and the statement of the results in linguistic scientific terms;
- (c) The “re-translation” of the results into a generally intelligible text and summary to be sent to or presented in court.

This multiple translation process is one of the most important matters one has to deal with. In giving linguistic expert testimony, even excellent expert opinions and experts may go astray in court by not observing the fact that German judges are not interested in linguistics, let alone in controversies in linguistic theory. Likewise, they do not care about things such as pragmatics. They are interested only in understanding that portion of linguistics specific to the particular case that will help them make a decision. Perhaps one could say that most German judges are interested only in the final evaluation, the summary of the linguist’s expert opinion. They most frequently start reading an opinion from the end rather than from the beginning.

In German courts, in the main, linguistic experts do not actually appear in court but send their written testimony to the courts/judges (see Introduction). It is more likely that judges will request a linguistic expert opinion in criminal cases (threatening and extortion letters of anonymous authorship) rather than in cases of defamation. In the latter, judges frequently feel that they themselves are the linguistic experts and

that they know the language well enough. Another interesting feature typical in the German judicial system is that the details and the argumentation of a linguistic expert opinion are rarely ever attacked by a lawyer without linguistic support, that is, by a non-linguist, in court. If there are objections and criticisms concerning an expert opinion, the opposing lawyer can order another linguistic expert opinion. With this, the former expert can be confronted and can respond to it, most frequently also in writing. There does not seem to be a real equivalent to cross-examination of experts in the German system.

(2) In the context of data and “data pedagogy”, it is also necessary to take into account maxims and professional routines to which the judiciary has been accustomed, has been exposed to, and that it has used in its own work to date. The notion of the ominous “uninfluenced ordinary reader”, the unbiased average reader, which in German is almost a liturgical stereotype in law texts (“*der unvoreingenommene/unverbildete/unbeeinflusste Durchschnittsleser*”), is a good example. It is important for linguists to start from there and to study such notions and basic concepts of the judiciary carefully to ensure that no misunderstandings arise from the very fact that such notions are received quite differently in both groups. Our own science, linguistics, is more to blame here than “the law” in terms of interdisciplinary awareness. There are culture-specific reasons, to be certain. There are German (mostly higher court) judges who are very interested in linguistic questions and who actually write papers in linguistic journals. They could act as a taskforce in bridging the gap between the different constituents of language and law.

(3) The majority of (German) judges, however, seem to be somewhat unaware of – and also somewhat uninterested in – linguistic matters. The point to be made here is that the number of people in the legal field that have dealt with linguistic questions in academic publications seems to be reasonably higher than the number of linguists that have legal training. In fact, in several decades of linguistic expert testimony, I have not met a single (German) linguistic colleague working as an expert who had a full training in linguistics and law – whereas, admirably, some (few) of the colleagues in the US have. What is needed here is that linguistic experts should not merely comment that this ominous notion of the “ordinary” reader does not exist. They should understand and explain to the judiciary that what is meant by this construction is a reasonable point which would have to be broken down into empirically operationalizable linguistic data.

The “ordinary German judge” would probably say: “The average uninfluenced reader of a text is me. How do I understand the text?”

Criticism of such a notion, which is an idealized construct, is not *that* it exists, but rather *how* to work with it and *how* to base strategies and theories on it. It would have to be explained to judges that language use and the understanding of an utterance as cited above for a word such as *Konkubine* are the object of an empirical science and that linguistics can come up with a systematic analysis and results that are valid and meet the basic criteria of science theory; such as reliability, validity, and intersubjectivity.

(4) It also should be explained to (German) judges that all expert testimony – including that of natural sciences, forensic chemistry, DNA analysis and others – requires interpretation. All data gained are not valid in themselves. An X-ray by itself is not identical to the verbal diagnosis of a radiologist. It should be explained to judges that linguistics is a science, not merely a field of people who have more knowledge of language(s). It must be made clear that language has its own rules, laws and postulates. This is said here because linguistics and handwriting analysis and all sciences that do not require a large technical apparatus are frequently erroneously considered to be directly accessible to laymen (and judges). A medical or chemical analysis is, in principle, no different than a linguistic analysis, or should not be. Different sciences have different measurements, and different needs and options of quantification.

(5) Judges would also have to realize that linguists are not chosen as experts on language use simply because they have some kind of personal authority as speakers of a language or because linguists are better speakers altogether. In the old days of linguistic expert testimony, one could often hear judges apologize for their own (command of) language(s) saying: “I am of course not as good at formulating texts as you”. Ironically enough, linguists seem, as a rule, not to be prototypical examples for rhetorical finesse and excellence, let alone for being more competent speakers than judges or professors of law. To be a competent speaker of a language and to be able to give a sound scientific analysis according to the methods and instruments of linguistics are not the same thing.

As mentioned before, one has to persuade the judiciary that one cannot just “see” or “feel” the semantic and communicative value of a word such as *Konkubine* in given contexts. It is not possible for the layman to know why that word has only these and no other denotations and connotations. Since judges are always very interested in precisely measurable and quantifiable scientific results, they need to learn to accept the laws, methods and routines of the science of linguistics as of every other

science. Certainly, the differences between the natural sciences and the humanities have to be accounted for.

(6) There is yet another area in which linguists need to exert some pedagogical influence on the judiciary. This is the general aversion or reservation that (German) judges had, in the old days, against “new” sciences, such as linguistics, which is not justifiable. One would have to convince judges that the linguistic expert is interested in enlarging their factual knowledge of language data and methods, not in influencing or restricting the judges’ competence in deciding a case or taking over the judges’ job. The worst thing linguistic experts can do is to take over the judges’ job and behave as if they have been asked to do a juridical evaluation of the linguistic data. It is important that linguists learn this.

(7) A common misunderstanding on the part of the judiciary is this famous sentence quoted with stereotypical frequency: “das Wort X bedeutet im Zusammenhang ...” (‘the word X means in the context in question ...’). The meaning in a particular context is what matters, not the meaning a word has in isolation. There is hardly any written statement of a judge in a defamation case in Germany, in which the phrase “im Zusammenhang bedeutet X ...” (‘the word X in the context given means ...’) does not occur. Also, the importance of the context is a commonplace in the legal literature. This is absolutely correct and necessary from a linguistic perspective too. It falls short, however, of three essential matters; namely, (1) of analyzing and explaining which linguistic data make up this context; (2) which particular proof it can or cannot provide. Sometimes it is used as a *deus ex machina* to evoke a special meaning with some kind of magic. So, it is the linguist’s job to explain the size, nature and function of a particular context, the fact that there is a linguistic and situational context that can be substantiated, why a particular context X and not Y is given, and what means and measures are used to describe them. The most salient deficit, however, is (3) that it is never explained *how* the context in question is to be analyzed and determined in an empirically sound manner. The judiciary has to be convinced by linguists that a systematic linguistic analysis of language system and language use is one of the preconditions *sine qua non* that judges need to know when they make an evaluation of linguistic analysis.

It was pointed out above that there are many – in fact, even more – misunderstandings on the part of linguists about the work that “law people” do. One concerns the context of a linguistic utterance. As stated above, the judiciary has long been aware of the fact that the “Zusammenhang” (‘context’) is of critical importance. Several so-called judicial “Praxiskommentare”, commentaries, which have been used for

about 50 or 100 years, read like introductions to pragmatics in this respect. To give just two examples (see also Appendix I):

Schönke and Schroeder (1976), "Praxiskommentar" to §185 StGB (on "Beleidigung" ('insult')):

ob eine Kundgebung der Nichtachtung oder Missachtung vorliegt, ist im wesentlichen die **Frage des Einzelfalls**. Handlungen oder Äußerungen von schlechthin beleidigendem Charakter gibt es nicht ... persönliche Eigenschaften oder Beziehungen des Angegriffenen, die Anschauungen bestimmter Kreise und die Gewöhnung der Beteiligten können den Angriff nicht als eine Missachtung, Verunglimpfung oder sonstige Herabwürdigung des Betreffenden erscheinen lassen

('... There are no acts or utterances of an insulting character *per se* ...')

Preisendanz (1975), "Praxiskommentar" to §185 StGB (on "Beleidigung" ('insult')):

Es gibt kaum eine Äußerung oder Handlung, die schlechthin als Beleidigung angesehen werden kann. Entscheidend sind immer die Umstände des Einzelfalls. Hierbei sind insbesondere Alter, Bildungsgrad und Stellung des Täters, die persönlichen Beziehungen zwischen den Beteiligten, das soziale Rangverhältnis, der Verkehrston in den betreffenden sozialen Schichten, sowie die Ortsüblichkeit bestimmter Ausdrücke zu berücksichtigen ...

('There is hardly any utterance or act that can be described as having an insulting or defamatory status *per se*. What matters are the circumstances of the particular situation of each case. For this, in particular the age, the educational level, the social position of the offender, the personal relationships between the people involved, the hierarchical social relations (the social ranks), the established ways of communication in the social classes involved, and the established local value of certain expressions have to be taken into account.')

It is a smart solution that the term "Beleidigung" ('insult') is *not* defined in the law and that the judiciary chooses to say that it depends on each particular case, situation and context how a language item is to be understood. It is remarkable that the judiciary, by itself, has elaborated such a formulation in line with what is now called the "pragmalinguistic" features of a text. It should be brought to attention here that it has only been some 50 years that sociolinguistics and pragmatics have received the proper attention they deserve in linguistics, mainly in the

US by scholars such as W. Labov, Ch. A. Ferguson, R. Shuy, S. Ervin-Tripp, D. Bolinger, D. Hymes and J. Searle, among others. Even today, not all linguists have adopted the view that a sociolinguistic and pragmatic frame of reference is of prime importance for the interpretation and empirical analysis of texts in legal contexts and elsewhere. From this perspective, the old scholars of law, as with the authors of the commentaries, have done a great job.⁸

To sum up: the notions and maxims of the judiciary, their nomenclature and routines of the analysis of texts have to be taken into consideration by linguists thoroughly and systematically. That a word has a specific meaning in a given linguistic and situational context and that one has somehow to refer to the invariants common to the majority of speakers is a fact that deserves proper linguistic attention and operationalization, more so than it has found to date. Many linguistic expert opinions would look different if linguists had had a look at the judicial literature in which such problems have been discussed for a long time. Even if not all linguists can achieve a law degree or work towards one, a more serious and more in-depth interdisciplinary perspective seems of critical importance.

The status of linguistic expert testimony *in/pro foro*

The status and a preliminary classification of linguistic expert testimony in Germany were given in a report some 25 years ago (Kniffka 1981: "Der Linguist als Gutachter bei Gericht" ('The linguist as expert witness in court')). Only a few remarks must do here.

As outlined in Figure 5.3, one can distinguish one branch of forensic linguistic expert opinions aimed at a "Verständnisnachweis" ('analysis of what is meant, including of what is said, in a text'). Another branch is called "Autorschaftsnachweis" ('analysis of who is the author of an anonymous text'). For expert opinions concerning the "Verständnisnachweis", four examples may suffice:

- (1) The analysis of the meaning of an utterance, the case of *Konkubine* as a libel (in the area of "Beleidigungsrecht"), as discussed above;
- (2) The analysis of whether an utterance or a part of an utterance – for example in advertising – can have only one meaning X and not another Y (in the area of §670 BGB, "Wettbewerbsrecht");
- (3) The analysis of text(sub)types such as "Widerruf", "Unterlassungsgebot", "Gegendarstellung", ('revocation', 'press correction');
- (4) Questions of trademark law, product names (in the area of "Warenzeichenrecht").

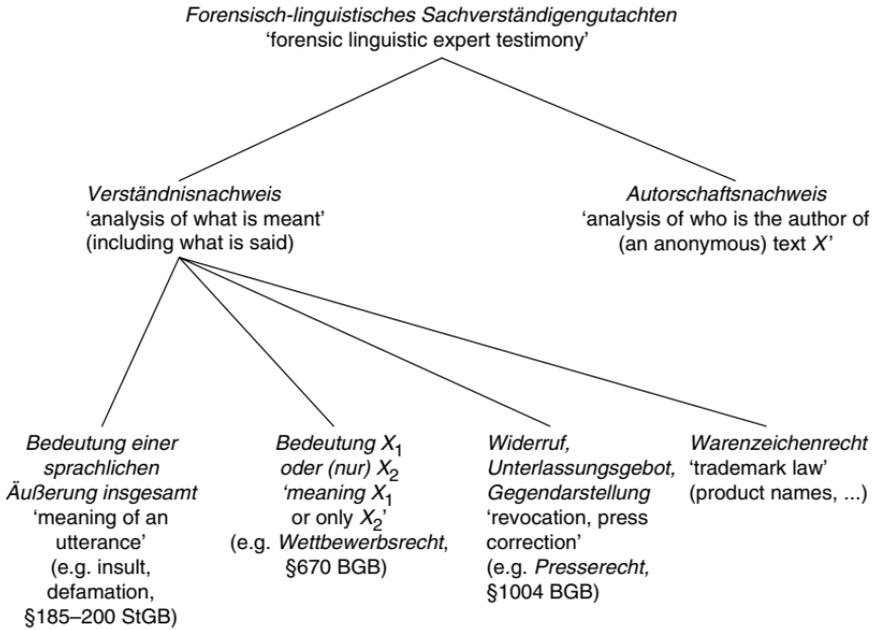


Figure 5.3 Areas of forensic linguistic expert testimony

Source: Kniffka (1981).

Several of the postulates that apply to a linguistic expert opinion have been described above. The most important practical ones can be briefly summarized as follows:

- (1) A linguistic expert opinion should be clinically clean from judiciary terms, jargon, and routine argumentation.
- (2) A linguistic expert opinion should never do a judiciary subsumption under sections of the law.
- (3) A linguistic expert opinion should also abstain from (excessive) special linguistic terminology, in-group jargon, and theoretical debates that are not understandable to laypeople; for example, judges. The twofold translation process (a layman question – into linguistic operationalized question – back to a layman answer) has to be handled with great care.
- (4) A linguistic expert opinion, contrary to older established routines, should not only describe *what* the results are, but also *how* they were found. It should also specify which hypotheses were tested on which samples. The results should, in principle, be quantifiable.

- (5) A linguistic expert opinion should clearly be focused on the particular purpose it is to serve with the people that request an expert opinion; for example, the judiciary. It has to meet the expectations and supply the information the people who order it wish to obtain. It should give precise response to the question(s) asked by the judiciary. If it fails to do so, the linguistic expert opinion is unsatisfactory – if not worthless. In some cases, the purpose for which the judiciary requires the linguistic expert opinion is served in a completely satisfactory manner. In other cases, the judiciary's aim cannot be met by the linguistic expert's work. In yet other (rare) cases, the linguistic expert opinion surpasses expectations.

Outlook and future tasks

It is difficult to foretell any general tendencies for the overall future development of linguistic expert testimony in Germany since there are no detailed documentations of expert testimony, which in itself is a problem of prime importance. Much has changed since we started giving expert testimony in the early 1970s. Some things we have learned, others we have not. Several other things the judiciary has learned, some it has not. Undoubtedly, the number of linguistic expert opinions given in German courts has increased rapidly over the last 20–30 years. There are no specific figures available to date, although work is being done in this direction.

Seemingly quite a large number of (German) colleagues from linguistics and the philologies have given forensic linguistic expert testimony on occasion, but, as a rule, they seem reluctant to talk about it. After one gets to know them better, they reveal that they did give an expert testimony some years ago in one case or even two or three. So, greater transparency and lucidity on expert opinions given in German courts would be an important desideratum for the future. Very few people indeed have focused their work on this field in Germany, as is the case in the US.

Most German judges seem to have heard or read about forensic linguistics. A few have published articles in the judicial or even the linguistic literature dealing with a linguistic question.

The development and general acceptance of forensic linguistics in and by the German judiciary can be very briefly summarized as follows:

- (1) Incidence of forensic linguistic expert testimony has increased considerably. The word was passed around that there is such a thing as forensic linguistic expert testimony, and there was a particular genre of court cases in Germany some decades ago in which linguists were

consulted, mainly linguists of the Bundeskriminalamt ('German Federal Criminal Police Office'): cases of terrorists with anonymous "Bekennerschreiben" ('confession letters'), which were usually left at the scene of crime by RAF (Red Army Fraction) terrorists. There was also an increasing number of linguistic expert testimonies in criminal and civil cases in the 1970s.

(2) German judges tend to believe much more in, or seem to be more impressed by, numerical quantifiable empirical evidence supplied by natural sciences – such as chemistry or DNA-analysis – than by humanities' argumentations and data. Judges prefer "countables" in statistics and are, *ipso facto*, impressed by any kind of "exact" measurement and quantification, no matter how sensible it is. It is almost a stereotypical question: "Can you give us the percentage with which someone is to be considered the author of this anonymous threatening letter?" or "Can you give us the percentage of how many people of the German population would consider this an insult or not?" To be sure, in most such cases the time frame is so short that the linguistic expert cannot undertake this type of empirical investigation. In some such cases, it just does not make much sense to count and state probabilities.

(3) Any kind of theoretical linguistic elements and discussion – for example, of semantics and pragmatics, which seem pertinent in many respects for defamation cases – are met with very reserved and critical skepticism rather than by positive cooperation by the majority of (German) judges. This is a somewhat strange situation. Many judges and law people actually employ pragmatics to a considerable extent when dealing with the meaning of a word in an utterance in a particular linguistic and situational context. The occurrence of a word in a particular context is a pragmatic category, after all. But many do not seem to appreciate linguists' expertise in sociolinguistic and pragmatic concepts and methods.

(4) What did convince judges frequently, according to my own 35 years' experience as a forensic linguistic expert, were references to and examples of real life items of everyday conversation. In the case of the *Konkubine* analysis, judges did accept the distinction between connotation and denotation whenever examples of different connotations were given. As for theoretical concepts, they were skeptical, sometimes even hostile, regarding any theoretical elaboration. However, in the *Konkubine* analysis for the non-neutral (pejorative) meaning they were swayed by text examples of everyday conversation and text type evidence, as given above. On this basis, they eventually could be persuaded that *Konkubine* could not be considered a neutral term but rather had a derogatory connotation in the contexts in question.

Appendix I: “Praxiskommentar” (Herdegen 1975)

As an illustration that the German judiciary has dealt with *pragmalinguistic* questions for some time, and also of the problems that defamation cases entail in Germany, the comments on some paragraphs of the “Praxiskommentar” (Herdegen 1975: 108ff) are given below. “Praxiskommentar” is a term in opposition to “Lehrkommentar”. In oversimplified (laymen’s) terms one could say that “Lehrkommentar” discusses interpretations of legal texts from a theoretical perspective, also mainly for students of law. “Praxiskommentar” is intended to give practical guidelines and illustrations to the judiciary and lawyers applying law.

The term “Beleidigung” (‘insult’) is (purposefully) *not* defined in German law. The excerpts of the “Praxiskommentar” quoted below in an approximate abbreviated translation illustrate why this is so.

Herdegen 1975, 108 (translation HK):

The action as facts of the case is called insult (“Beleidigung”) without any further explanation in the law ... The utterance must be objectively injurious to the honor of a person and thus violate respect for that person. The SENSE [of an insult] primarily is defined by the recipient’s understanding of the utterance. Certain information is addressed to him ... The general OBJECTIVE SENSE [of the utterance], however, the objective meaning of the content of thoughts, is determined by the general communication standards and interpretation of the larger (speech) community in which the interaction takes place, unless the special communication situation of the recipient of the information prevails

The sense of an utterance is to be analyzed and to be determined by the trial court (judge) ... Whether the sense of the utterance implies a violation of respect is entirely dependent upon the PARTICULAR CIRCUMSTANCES and the context. Age, gender and the clean record of the offensee as well as the customary communication habits (ways of interaction and the general attitude of the interlocutors involved and their traditions) are to be taken into account as well as the particular context in the intention of the expedient ...

In order to evaluate the text of an interview, e.g., not only the particular verbal constituents but also the entire contexts of the article are to be taken into account, the aim that the author wants to achieve, and the behavior of the offensee in the interview ...

The same action, depending on the circumstances under which it was made, can be an expression of disrespect in some situations and

not in others. An utterance that in itself is entirely harmless can turn into a violation of honor (e.g. by derogatory intonation ...). Even an act that is generally regarded as an expression of respect can take on the status of disrespect ... On the other hand, an utterance violating somebody's honor can be used in special circumstances that take away any disrespectful characteristics of it.

ABSOLUTE insults *per se*, behaviors that under all circumstances and in all contexts are a violation of honor, do not exist ...

From this it follows that everything is a matter of evaluation of the particular case in a particular context situation ...

Abusive terms and swearwords are insulting if they claim the absence of the moral integrity of the offensee in his own way of life or in his behavior towards others (such as rotter, bastard, dog, swine) or attest an elementary human deficit ...

Appendix II: Original text of a forensic linguistic expert opinion (English translation)

The actual expert opinion given in the *Konkubine* case described above (see Kniffka 1990c) is given below in an informal translation (some notes and parts of the text are left out here).

“Gutachten zum Inhalt des Wortes *Konkubine* in den Briefen des Beklagten”

(‘Linguistic expert opinion on the content of the word *concubine* in the letters of the defendant’)

0. Preliminaries

0.1 The linguistic expert opinion wants to analyze which content and meaning(s) the word *Konkubine* has in the letters of the defendant, Mr. Y, which are attached. In particular, it wants to determine whether it is used as a neutral (non-pejorative) term or has to be understood in a pejorative/derogatory sense.

0.2 It would mean applying a naive (and linguistically not justifiable) concept realism, if one made a statement “*concubine* is an abusive term/a verbal injury” or “*concubine* is not an abusive term/a verbal injury” (as is done in the preceding expert opinion by ...). This would not report the results of a linguistic analysis but give a judicial evaluation of a linguistic fact, to which we as linguists do not feel entitled and competent. Terms such as “Schimpfwort” (‘abusive term’) or “Verbalinjurie” (‘verbal injury’) cannot be defined linguistically in any reliable and sensible way. The

evaluation of the extent to which a pejorative term in the perspective of the judiciary manifests an offense of “Beleidigung” (‘insult’) as defined in the *Großer Brockhaus* (“jede vorsätzliche, die Ehre eines anderen kränkende rechtswidrige Kundgebung” (‘every utterance intentionally violating the honor of another person in an unlawful way’)) is not the linguist’s realm but that of the judiciary.

1. General linguistic criteria of analysis and evaluation

1.1 It is useful and necessary to distinguish between the denotative meaning (“conceptual content”) and connotative meaning (“emotional sense”) of a word. It should be noted that connotations do not randomly vary from person to person. Rather, connotations (as with denotations of a word) are governed by rules that are in part super-individual. A word such as *Nazischwein* in an everyday speech situation (for example, in an utterance such as *Sie altes Nazischwein!*), can hardly have positive connotations.

1.2 The entries in an encyclopedia or in a dictionary concerning the meaning of a particular word, which are always quoted as an authority, are totally inefficient and of little value for the analysis of the actual meaning of a word in a given context. They may be useful as a very first orientation only. For the actual specific meaning within a certain speech act, the *a verbo* meanings supplied in a dictionary give only minimal clues, much like etymological data given in an etymological dictionary. In any case, the analysis of the particular linguistic context in which a word is used and, possibly, the analysis of judgments and attitudes of native speakers of the language, are of much more avail for the analysis of a special meaning of a word than the consultation of dictionaries. Dictionaries, generally speaking, are very heterogeneous in style and method and also have many missing data. Also, they are “history books”, not covering the actual data of language use.

1.3 In addition to the linguistic context in a narrower sense of the word, also the extra-linguistic (situational) context, the specific ingredients of the particular speech situation, have to be taken into consideration. Modern linguistics, particularly sociolinguistics and speech act pragmatics, has defined these extra-linguistic contexts as being as important and decisive for the status and meaning of an utterance as the data of the linguistic code itself. *Who says what to whom, when, and how* may contain more relevant information on the holistic status and value of an utterance than the content proper, the verbal elements uttered. The same linguistic elements (words) can express different speech acts. The form of address *Schätzchen* (‘dear’, ‘love’), depending on the interlocutor and

the situation, can represent an act of verbal endearment, but also a provocation or threat. *I love you* can be a marriage proposition or the final sequence of a telephone conversation with one's wife.

1.4 The necessity to include the data of the extra-linguistic speech situation is not meant to imply that a word according to the linguistic and extra-linguistic context can mean all or nothing, i.e. that its semantic interpretation is absolutely random. Rather, the use of linguistic elements in different situations underlies a complex set of rules that are not at the disposal of particular speakers, but are of super-individual status. To analyze the linguistic and extra-linguistic context in order to describe the particular meaning of a word does *not* mean that the interpretation can be made randomly in one way or the other. On the contrary, it means that a more exact and specific analysis of the content and the meaning is possible, which is certainly progress in terms of the judicial evaluation of the linguistic facts.

2. Analysis of the content and meaning of *Konkubine* in the letters of the defendant

2.1 Even independent of contextual arguments, which, in this case, are quite unambiguous in our opinion (see 2.4 below), one can show that the word *Konkubine* in today's standard German (1972) is not a neutral term. By its history, the word denotes a woman, who, for a substantial period of time, has extra-marital sex with a man. *Konkubine* in today's German does not only denote a woman who is living together with a man, but the gist of the semantic component is that it is a woman who has sexual intercourse with a man without being married. This very semantic component carries the violation of honor because such an activity (without being married) according to the general moral standards in this country at the present time is still considered somewhat despicable, improper, socially unacceptable, which the defendant clearly states in other passages of his letter. The comparison with synonymic words such as "Partner" ('partner'), "Freundin" ('girlfriend'), "Bekannte" ('acquaintance') shows that *Konkubine* definitely has a pejorative semantic component. A usage of one of the other three synonymic words in the particular context would have entirely different semantic implications.

2.2 The fact that the word *Konkubine* is not entirely neutral can also be illustrated by an analysis of the possible textual occurrences of the word. If it were a neutral (non-pejorative) term, it should be possible to replace it by synonymic words in linguistic contexts that are marked as

neutral (or at least non-pejorative, non-derogatory). Two examples to illustrate this:

Official letters of invitation from the head of a company who wants to invite his married and unmarried male employees could read *Ich würde mich freuen, wenn Sie Ihre Frau/Partnerin mitbrächten* ('I would be very happy if you could bring your wife/partner along'). This would be absolutely OK. An alternative *Ich würde mich freuen, wenn Sie Ihre Frau/Konkubine mitbrächten* ('I would be very happy if you could bring your wife/concubine along') would not be acceptable, however.

Another example from a different context:

A judge in a verdict ("Urteilsbegründung"), which undoubtedly requires a precise, factual and neutral way of speaking, would *not* designate a woman who has extra-marital sex with a man by *Herr Y und seine Konkubine* ('Mr. Y and his concubine'), but only by *Herr Y und seine Partnerin/Bekannte* ('Mr. Y and his partner/acquaintance'). The most appropriate and neutral reference would undoubtedly be *Herr Y und Frau X* ('Mr. Y and Ms. X').

2.3 Concerning the extra-linguistic context, the pejorative (derogatory) connotation of the word *Konkubine* could be demonstrated easily by experiments in which the partners of interaction, the situation, and the style would vary. Without having been able to do such experimental analyses it seems justified to conclude with high probability that, for example, a prospective daughter-in-law would not be addressed as *Konkubine* by her parents-in-law vis-à-vis the own son even if they have sexual intercourse.

2.4 An even more decisive indicator for a pejorative (derogatory) connotation of the word *Konkubine* is the particular linguistic and situational context in the letters of the defendant. It only and repeatedly contains syntactic forms such as *Herr X bzw. seine Konkubine* ('Mr. X or his concubine respectively'). There is a striking *asymmetry* of this formula – because Ms. Y's name was known to the defendant. The unmarked form would have been *Herr X und Frau Y* ('Mr. X and Ms. Y'). Additionally, using the word *Konkubine* repeatedly in a large number of letters contains information (that is extra-marital intercourse), which is irrelevant for the general context. All this indicates that the author of the letters was not intending to make a neutral statement that two people were living together

without being married, but that he wanted to perform a speech act that a native speaker of German would easily be able to decode as 'to get someone' or 'to offend someone'.

3. Summary of the linguistic evaluation

3.1 By its status and function as a speech act, and by the linguistic and extra-linguistic context, the use of the word *Konkubine* in the letters of the defendant cannot be explained as a neutral (non-derogatory, non-pejorative) or even joking use of a word.

3.2 No matter what the etymological history of the word *Konkubine* and the entries in dictionaries may say, it is clear that the word *Konkubine* in the context in question contains a pejorative (derogatory) semantic component, that it wants to express continuant extra-marital intercourse of the offensee, that the speaker considers this fact to be non-integer and immoral, and that it is the intention of the speaker to insult the person so named.

6

A Heuristic Author and Writer/Typist Taxonomy

Introduction

It is a well-known fact in forensic linguistics and other sciences dealing with written texts that the author of a text, the person who conceived, invented, and formulated the text as it appears as a frozen product on paper, and the person who wrote it, typed it on a typewriter or a computer or wrote it by hand, are not necessarily one and the same person. Frequently they are one and the same (frequently, also, they are not, depending on the text type, the function, the status of the text and so on). For some reason people, in particular laymen, seem to assume that author and writer are identical and also that there is, unless otherwise stated, just one author, and one only, of a written text.

These assumptions are not justified, taking into account the vast amount of texts of a non-criminal denomination appearing on paper, let alone a written text of an incriminating nature. The most salient methodological postulate deriving from this implies that a one-to-one relationship of author and writer/typist of a written text cannot be presupposed or even taken for granted, as has been expressed time and again in the literature (cf. Kniffka 1981; 1990b; 1996b; 2003b; Shuy 1993a; 1998a). The intricate problems of the description and evaluation of written incriminating texts and the matching process with comparison data has also been discussed in greater detail elsewhere (Kniffka 1990d). Almost every real life expert opinion makes some kind of reference to the nature and status of the written data of the incriminating and the comparison texts. This is of critical importance since, as a rule, the linguistic expert has no way of examining and supplying solid evidence that the comparison data have actually been written by the defendant and nobody else. Linguistic experts, usually, can only accept and work

with the data that they have been given by the people that have requested the expert opinion. The practical problems of dealing with orthographic data of incriminating and comparison texts are beyond the realm of this chapter.

In focus, here, are the theoretical possibilities and the more or less “standard” constellations and combinations of the authorship and writership of a text that have to be accounted for.

It is assumed that a more detailed empirically based heuristic taxonomy may help to recognize a certain author (henceforth AU) and writer (henceforth W) constellation more easily in practical real life work in the future. This chapter can only describe a very few of the relevant aspects in an exemplaric and abbreviated fashion, rather than giving a full-fledged description and an exhaustive discussion of mathematical models accounting for each and every possible variant. The focus is on the relevance to the practical work of the forensic linguistic expert.

Dimensions of variation

To account for the most important variants for the authorship and writership of a text, it seems necessary to assume at least the following *dimensions of variation*:

(1) *Number of AU and W*

The numbers of AU and W each can be different. Not only can there be only one AU and one W, be they identical or not. For AU and W a number of persons >1 can apply. In addition, there may be some overlap as far as the functions (see below) are concerned. It may be that the writer or writers are simultaneously the authors. There may also be authors who simply perform the role of writers at the same time.

(2) *Relationship of AU(s) and W(s)*

The relationship of AU(s) and W(s) can vary considerably. The AU and W may be friends, family members, colleagues, collaborators, competitors, accomplices, lovers, enemies. The W(s) can be forced by the AU(s) or by other persons at gunpoint to write what is being dictated, they may have been taken hostage and have to write what they are told. There is an enormous range of possibilities here.

(3) *Text entity and entities and “segmentability” of a text*

A text may consist of or represent one total indivisible entity as far as the authoring and the writing process are concerned. Or it may consist of

several different parts/sections/chapters/fractions. It may be that persons A and B authored a section of a text that person C is writing/typing, whereas sections 2 and 3 are authored by persons D, E and F and are written/typed by persons H and I. I use the terms “multiple authorship” and “multiple writership” for these constellations. It has also been stated before (Kniffka 1981; 1990b) that, unless there is a special constellation of different languages, alphabets, ways of writing and other more general differences, it is very difficult to determine who authored what and who wrote what. In most cases, only the people directly involved know how to answer these questions (this applies also to many other cases, such as “the fathers of the German constitution”).

(4) *Modalities* of production of written texts

The fourth, and perhaps most salient, dimension of the variation of the outward conditions of a written text may be represented by the different verbs that designate it. For the time being, it is sufficient to use terms of everyday conversation for the various activities and modalities to create a heuristic taxonomy of the types of writing process, as follows:

- (a) A text is written by one person but dictated by another;
- (b) A text is written by one person while being dictated by another by force or at gunpoint;
- (c) Someone is writing out by hand or typing (for example, on a typewriter or a computer) a text that has been pre-written (for example, in handwriting (“Abschreiben/Reinschrift eines vorge-schriebenen Textes”));
- (d) A text is written/spelled out on the basis of (abbreviated) notes (which involves some authoring at the same time);
- (e) A text is copied from another (more or less identical) text.

(5) *Media*

As a fifth dimension, one can assume the media through/by which the written text is produced. This mainly concerns the distinction of handwritten versus typed texts, stencil-written texts and others. The data of the technical, material and physical aspects of the production of a written text are addressed by sciences other than forensic linguistics. Handwriting analysis (the official German term being “Gerichtliche Schriftvergleichung”) has its own experts, who work together with linguists and other linguistic experts in many cases and frequently consult forensic linguistic experts when questions of orthography or other functional categories are concerned. In addition, there are forensic experts for typewriting and typewriting

machines and computers who can give expert testimony in case there are typewritten texts amongst the incriminating and comparison data.

A systematic analysis in these five dimensions of variation places the forensic linguistic expert analysis on safer and more systematic grounds. In particular, this holds for the analysis of the writing process of written texts, which has been somewhat neglected, as opposed to linguistic questions proper in forensic linguistics.

A heuristic taxonomy of features and variants of authorship and writership of incriminating texts

Combinations of dimensions

Taking the five dimensions of variation outlined above into consideration, one could set up, somewhat from a layman's perspective, the following (highly simplified) heuristic taxonomy for the constellations of authorship and writership of a written text. Table 6.1 illustrates a first heuristic description to state some working hypotheses. They are of theoretical value only, since the real life proposition of the variants needs ample empirical study in its own right.

Table 6.1 Author and writer constellations from a layman's perspective

(1)	1 AU = 1 W	W and AU identical, authorship/writership is not disguised
	1 AU = 1 W	
	1 AU = 1 *W	
(2)	1 AU ≠ 1 W	W and AU different, W writes/types text by AU voluntarily, undisguised writership
	1 AU & 1 W	
	1 AU & 1 *W	
(3)	1 AU ⇒ 1 W	W and AU different, W writes/types text by AU involuntarily/by force
	1 AU & 2 Ws	
	1 AU & 2 *Ws	
(3)	1 AU & 2 Ws	2 Ws write/type text by AU voluntarily, undisguised writership
	1 AU & 2 *Ws	
	1 AU ⇒ 2 Ws	

Continued

Table 6.1 Continued

(4)	2 AUs & 1 W	
	2 AUs & 1 W	2 AUs design text as “joint venture”, 1 W writes/types whole text voluntarily, undisguised writership
	2 AUs & 1 *W	2 AUs design text as “joint venture”, 1 W writes/types whole text voluntarily, disguised writership
	2 pAUs & 1 W	Each AU designs a part of the text, 1 W writes/types the whole text voluntarily, undisguised writership
	2 pAUs & 1 *W	Each AU designs a part of the text, 1 W writes/types text voluntarily, disguised writership
	2 AUs \Rightarrow 1 W	2 AUs design text as “joint venture”, 1 W writes/types whole text involuntarily/by force
	2 pAUs \Rightarrow 1 W	Each AU designs a part of the text, 1 W writes/types the whole text involuntarily/by force
(5)	2 AUs & 2 Ws	
	2 AUs & 2 Ws	2 AUs design a text as a “joint venture”, each W writes/types one part of the text voluntarily, undisguised writership
	2 AUs & 2 *Ws	2 AUs design a text as a “joint venture”, each W writes/types one part of the text voluntarily, disguised writership
	2 pAUs & 2 Ws	Each AU designs one part of the text, each W writes/types one part of the text voluntarily, undisguised writership
	2 pAUs & 2 *Ws	Each AU designs one part of the text, each W writes/types one part of the text voluntarily, disguised writership
	2 AUs \Rightarrow 2 Ws	2 AUs design a text as a “joint venture”, each W writes/types one part of the text involuntarily/by force
	2 pAUs \Rightarrow 2 Ws	Each AU designs one part of the text, each W writes/types one part of the text involuntarily/by force

Key:

AU:	author
W:	writer
1, >1:	number of W/AU
T:	text
pAU:	partial text author
pW:	partial text writer
=:	AU and W are identical
≠:	AU and W are non-identical
*:	disguised authorship or writership
&:	text is “joint venture” of AU and W
\Rightarrow :	AU forces W to write text

Set of binary features

A more systematic and precise account of the affiliation and of the types of variants occurring can be given by defining a set of binary features. Table 6.2 gives a heuristic taxonomy of text products defined by a set of binary features, which seem to be relevant for the classification of incriminating texts here. In a practical preliminary fashion, the features can be defined as listed below. They are given in a decreasing hierarchical order; that is, the feature [\pm crim] is on a higher level than [\pm disguised], [\pm forced] and others.

The main objective of this heuristic taxonomy is to allow for a more systematic and detailed analysis and classification of incriminating text types and texts. It is an empirical question whether a different hierarchical order of features might be feasible for non-incriminating texts.

Table 6.2 Heuristic taxonomy of text products

(i) [\pm crim]	<i>Text is or is not of incriminating nature.</i> [+ crim] means incriminating texts – such as extortion, blackmail, threatening, libel, defamation letters; [– crim] means non-incriminating letters of various kinds – such as love, business, condolence, letters of application.
(ii) [\pm anon]	<i>The (name of the) author is or is not revealed.</i> [+ anon] means that the author did not give his/her real name/did not sign a text; [– anon] means that the author gives his/her real name, which is somewhat rare in incriminating texts (see pp. 159–62).
(iii) [\pm id]	<i>Identity of author(s) and writer(s).</i> [+ id] means that the author and the writer are one and the same person or that the authors are, at the same time, the writers; plurality of authors matches plurality of writers; [– id] means that author(s) and writer(s) are two (or more) different persons.
(iv) [\pm >1AU]	<i>Number of authors of a text.</i> [+ >1AU] means that the text is authored by more than one person; [– >1AU] means that the text is authored by one person.
(v) [\pm >1W]	<i>Number of writers of a text.</i> [+ >1W] means that the text is written/typed by more than one person; [– >1W] means that the text is written/typed by one person.
(vi) [\pm dis AU]	<i>Authorship is or is not disguised.</i> [+ dis AU] means that the author of the incriminating letter(s) tries

Continued

Table 6.2 Continued

	to disguise her/his authorship (e.g. by imitating the known style of a different person using jargon, other regional variants); [- dis AU] means that the author does not try to disguise his/her authorship.
(vii) [\pm dis W]	<i>Writership is or is not disguised.</i> [+ dis W] means that the writer/typist tries to disguise her/his writership (e.g. by writing with the left hand, by imitating another person's handwriting, by writing in capital letters); [- dis W] means that the writer/typist does not disguise his/her writership.
(viii) [\pm for AU]	<i>Author formulates a text voluntarily/at her/his free will or is forced by other person(s).</i> [+ for AU] means that the author is forced (e.g. threatened at gunpoint) to formulate a text against her/his will; [- for AU] means that the author formulates a text not forced by anybody but voluntarily/at his/her free will.
(ix) [\pm for W]	<i>The writer/typist writes/types a dictated text voluntarily or forced by someone.</i> [+ for W] means that the writer/typist is forced (e.g. threatened at gunpoint) to write a text as dictated; [- for W] means that the writer/typist writes/types a text as dictated voluntarily/at her/his free will.
(x) [\pm cop]	<i>The text is copied from another (more or less identical) text.</i> [+ cop] means that the writer(s)/typist(s) is/are given a text already existing in (finalized) written form to copy; [- cop] means that the writer(s)/typist(s) write(s)/type(s) a text given not in (finalized) written form, but dictated on tape or live (by another person) using short written notes and so on.

It seems that this list of ten binary features is sufficient to account for a description of the main variants of incriminating texts. The ten features concern major theoretical dimensions of variation (see pp. 156–62). Many more could be set up to account for all the possible aspects of variation, but this is beyond the concern of this chapter.

Nothing is said here about their real life occurrence and distribution as cluster-combinations defining text-(sub)types (see pp. 159–62). For this, there are no reliable data available at present. It is extremely difficult to collect quantifiable data, such as which variants/clusters of features are the most and the least frequently occurring ones. In many cases, this cannot be determined on safe empirical grounds. For example,

in cases of multiple authorship (cf. Kniffka 1981; 1990b; 2000b), only the people involved know for certain who wrote what under what conditions. In addition, there seems to be a large range of variation in the definition of what a shared or multiple authorship actually means. This being so, a systematic explanation seems even more critically important.

Hierarchical structure of features of authorship and writership

Using the set of binary features defined above, one can illustrate the hierarchical affiliation of features (see Figure 6.1). It is designed to illustrate and illuminate the variation of texts and text types that are found amongst incriminating texts.

The aim of Figure 6.1 is not to give an exhaustive list of all features and text types possible but to give an exemplaric set of features to illustrate symptomatically some selective features that play a major role in the definition of the everyday, most frequently occurring text types of incriminating letters.

The numbers in parentheses at the bottom refer to the particular type or subtype of the incriminating letter defined by the “Distinctive Text-Profile-Matrices” (henceforth DTPM; see pp. 159–62).

As said before, a much larger number of features, and also variants, may have to be taken into consideration in order to give an exhaustive list and definition of all incriminating text types that occur, but this is beyond the scope of this chapter.

To what extent this set of features applies to incriminating and non-incriminating texts is also beyond the scope of this chapter. It is an empirical question. Some of the features may be of equal importance for the description of non-incriminating texts. It also is possible that a considerable number of additional features will have to be taken into consideration.

There seems to be some kind of “predisposition” of features for text types, which could even be stated as correlations. The value [+ anon], for example, seems to occur in a strong correlation with types that constitute an offense (a crime or a tort) by themselves, by their very nature. A person who writes an extortion, blackmail or threatening letter is committing a crime by this activity. This, in turn, is reason enough for an offender to stay anonymous. This is not necessarily the case with some libel and defamation letters. If the fact claimed in the letter can be proven in court to be true, there is no offense according to German law and the author will not be charged. From my own experience as a forensic linguistic expert, I would estimate that a very

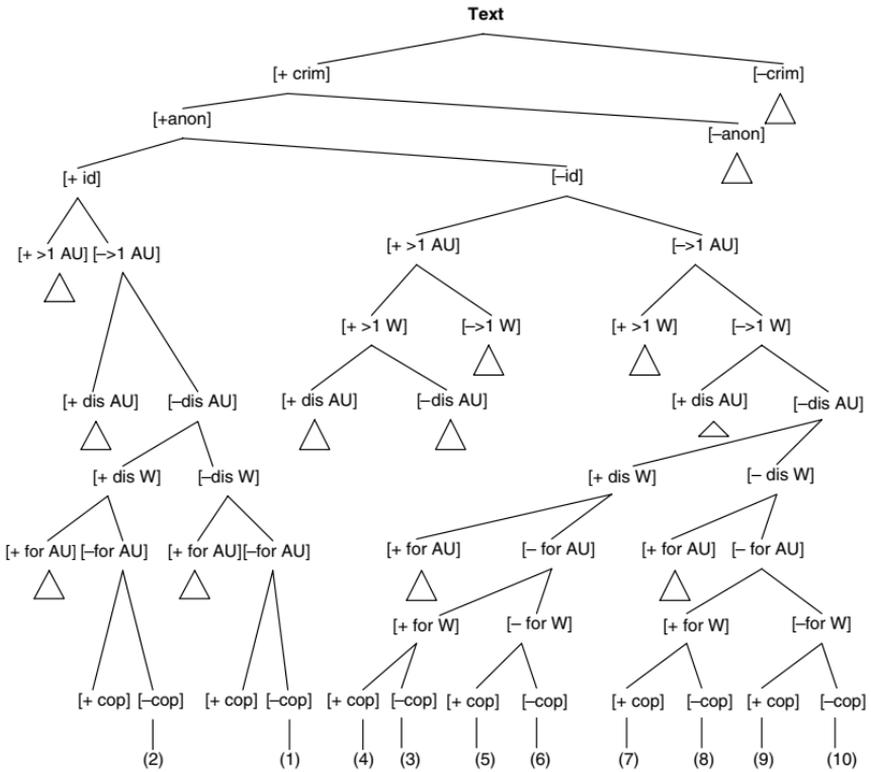


Figure 6.1 Hierarchical affiliation of features of authorship and writership

large percentage of defamation and libel letters are sent anonymously, under a false or a non-specific name. Since there are no reliable data available at this time, a detailed analysis of the number of cases and their characterization will have to wait for an empirical validation to be made.

This is yet another point illustrating the urgent need for a full-fledged empirical documentation of the nature and structure of incriminating texts in real life cases. Since such data are not available, all assumptions about the statistics of the real life occurrence and distribution of the features and the variants discussed cannot be stated in exact terms. This is a huge desideratum for future work.

It may also be necessary to consider the definition of binary features in some cases. There has been little success in the analysis of anonymous multiple authorship cases (at least, in my own forensic work). When there is a group of authors, only the people directly involved know for sure who has written what, where, how and to what extent in the text. A large range of variation is possible if several authors are involved. One

author could be more dominant or more experienced in formulating a text than another. This might necessitate using scalar rather than binary features or assuming yet an additional set of binary features. At any rate, considering the fact that so little empirically safe work has been done, giving a theoretical and systematic discussion of the general hierarchical structure of features seems all the more necessary.

It also would need further empirical and theoretical study to examine whether it is really justified and necessary to assume binary features [\pm dis] both for authorship and writership and [\pm for] for author(s) and writer(s). There may be several empirical elements that work as constraints here. Can one disguise one's *authorship* when creating an incriminating text to the same extent, in the same way and in the same manner in which one can disguise one's *writership*? Intuitive evidence gathered from working as a forensic expert would suggest that disguise in both authorship and writership may need more sophisticated differentiation, and may show more differences than similarities than have been discussed so far. The use of the same terms may actually run the risk of neglecting some salient differences. Disguise of writership uses well-established identical variations (see above). Disguise of authorship does not seem to have such a well-defined and established set of possible variations at its disposal. In simple terms, hiding one's personal identity – difficult as it may be over a longer stretch of text – and (also at the same time) imitating somebody else's writing, may need many more sophisticated differentiations than are known so far. There may be two large repertoires of activities involved, denotable by the Latin verbs *dissimilare* and *similare*, which may concur but are quite different structurally. This is another important area requiring further forensic linguistic study. To date, there has been no empirical study of the co-occurrence of disguise of authorship and disguise of writership. Disguise does not only mean anonymity/ hiding/non-revealing of the author's and/or writer's full name or the use of a pseudonym. "Disguise of authorship" can be defined as the "deliberate choice of a language behavior other than one's customary and usual way of speaking/writing in communicative situations and texts of a comparable nature for the purpose of concealing one's authorship." "Disguise of writership (handwriting)" can be defined as "the deliberate alteration of one's customary handwriting in order to conceal one's writership of a text."

The features [\pm for AU] and [\pm for W] seem to be a rather simplistic notion concerning certain empirical elements that also ask for a much more detailed differentiation. Can one actually *force* another person to formulate/author a (new) text in the same way that one can force another person to write a text given in dictation? In forcing someone to

author a text, it seems that the person who forces the author would, in one way or another, have to “prefabricate” the text and get it across to the second author. This needs a lot of further in-depth discussion and empirical analysis.

The feature [\pm dis W] needs further specification as well. It applies to handwritten texts only. For example, one can only physically disguise one’s handwriting, not one’s typing. One may disguise one’s true orthographic and spelling behavior, however, suggesting/pretending indirectly that the writer/typist is somebody else (someone who would “write that way”).

As mentioned previously, the overall hierarchy of features given in Figure 6.1 may have to be modified for texts other than incriminating texts, and possibly even for incriminating texts, depending on the particular text type and subtype. The hierarchy given in Figure 6.1 is thus one possibility out of several others.

Taxonomy of variants of incriminating texts

Using an adapted version of the technique established many decades ago in the “distinctive features” framework of phonology (cf. Jakobson *et al.* 1951) one can set up combinations or clusters of features that can be combined for the description of text products. Below, an exemplaric set of 10 “Distinctive Text-Profile-Matrices” for types of incriminating texts that occur in real life is given. The selection itself is very abbreviated, fragmentary, and non-exhaustive. Its purpose is to give a first-hand impression of what the variation of text products looks like and how it can be described and classified in a more systematic and adequate way. As previously stated, a list of all theoretically possible constellations of features as shown on p. 155 is of no relevance for the description of the text types of incriminating letters. They are of considerable theoretical concern, but will not help much in describing text types of incriminating letters occurring in real life.

The matrices listed below are chosen because they refer to text types of incriminating letters that have been established on empirical grounds as actually occurring. Relying on my personal experience as a forensic linguistic expert, it would appear that the matrices describe the more frequently occurring types rather than the rarely or extremely rarely occurring ones. However, as stated above, there are no reliable data available at this time by which one could determine the number of each particular text type or subtype. Even if exact and quantifiable data were available, it would be difficult to determine the actual number: In constellations such as (1) and (2) it is not possible to determine, generally speaking, whether a disguise can be assumed or not. Several cases have to be left open in this respect.

Table 6.3 Distinctive text-profile matrices

(1)	$\begin{bmatrix} + & \text{ crim} \\ + & \text{ anon} \\ + & \text{ id} \\ - & >1 \text{ AU} \\ - & >1 \text{ W} \\ - & \text{ dis AU} \\ - & \text{ dis W} \\ - & \text{ for AU} \\ - & \text{ for W} \\ - & \text{ cop} \end{bmatrix}$	<p>One person formulates and writes/types an incriminating anonymous text at his/her own discretion without disguising his/her writership.</p> <p>It looks as if this is the “ordinary” and most common type of an <i>anonymous</i> incriminating text, in text types such as extortion, blackmail, threatening, defamation letters, and others. In these, [+ anon] seems to be an inherent feature.</p>
(2)	$\begin{bmatrix} + & \text{ crim} \\ + & \text{ anon} \\ + & \text{ id} \\ - & >1 \text{ AU} \\ - & >1 \text{ W} \\ - & \text{ dis AU} \\ + & \text{ dis W} \\ - & \text{ for AU} \\ - & \text{ for W} \\ - & \text{ cop} \end{bmatrix}$	<p>One person formulates and writes/types an incriminating anonymous text at his/her discretion and free will and disguises his/her writership (for example, by writing in capital letters, with a stencil).</p>
(3)	$\begin{bmatrix} + & \text{ crim} \\ + & \text{ anon} \\ - & \text{ id} \\ - & >1 \text{ AU} \\ - & >1 \text{ W} \\ - & \text{ dis AU} \\ + & \text{ dis W} \\ - & \text{ for AU} \\ + & \text{ for W} \\ + & \text{ cop} \end{bmatrix}$	<p>One person formulates and pre-writes an incriminating anonymous text and forces another person to copy it with disguised writership.</p> <p>As stated on pp. 156–9, only a disguise of handwriting, not of typing seems possible. In orthography and spelling an indirect disguise is possible (see pp. 156–9) in that a person can create errors and mistakes that suggest a different person as writer.</p>
(4)	$\begin{bmatrix} + & \text{ crim} \\ + & \text{ anon} \\ - & \text{ id} \\ - & >1 \text{ AU} \\ - & >1 \text{ W} \\ - & \text{ dis AU} \\ + & \text{ dis W} \\ - & \text{ for AU} \\ + & \text{ for W} \\ - & \text{ cop} \end{bmatrix}$	<p>One person formulates an incriminating anonymous text and forces another person to write/type it with disguised writership by dictation (rather than by copying).</p>

Continued

Table 6.3 Continued

(5)	$\left[\begin{array}{l} + \text{ crim} \\ + \text{ anon} \\ - \text{ id} \\ - >1 \text{ AU} \\ - >1 \text{ W} \\ - \text{ dis AU} \\ + \text{ dis W} \\ - \text{ for AU} \\ - \text{ for W} \\ + \text{ cop} \end{array} \right]$	<p>One person formulates and pre-writes an incriminating anonymous text. Another person copies it voluntarily with disguised writership.</p> <p>This variant represents, for example, an extortion letter in a crime in which two criminals are involved.</p>
(6)	$\left[\begin{array}{l} + \text{ crim} \\ + \text{ anon} \\ - \text{ id} \\ - >1 \text{ AU} \\ - >1 \text{ W} \\ - \text{ dis AU} \\ + \text{ dis W} \\ - \text{ for AU} \\ - \text{ for W} \\ - \text{ cop} \end{array} \right]$	<p>One person formulates an incriminating anonymous text. Another person writes it voluntarily by dictation with disguised handwriting.</p> <p>(Variant very similar to (5)).</p>
(7)	$\left[\begin{array}{l} + \text{ crim} \\ + \text{ anon} \\ - \text{ id} \\ - >1 \text{ AU} \\ - >1 \text{ W} \\ - \text{ dis AU} \\ - \text{ dis W} \\ - \text{ for AU} \\ + \text{ for W} \\ + \text{ cop} \end{array} \right]$	<p>One person formulates and pre-writes an incriminating anonymous text and forces another person to copy it with non-disguised writership (handwriting).</p> <p>A variant in which no disguise of writership occurs is represented, for example, by ransom notes, which the victim (a hostage) is forced to write by dictation</p>
(8)	$\left[\begin{array}{l} + \text{ crim} \\ + \text{ anon} \\ - \text{ id} \\ - >1 \text{ AU} \\ - >1 \text{ W} \\ - \text{ dis AU} \\ - \text{ dis W} \\ - \text{ for AU} \\ + \text{ for W} \\ - \text{ cop} \end{array} \right]$	<p>One person formulates an incriminating anonymous text and forces another person to write it by dictation with undisguised writership (handwriting).</p> <p>(cf. type (7))</p>

Continued

Table 6.3 Continued

(9)	<table border="1"> <tr><td>+</td><td>crim</td></tr> <tr><td>+</td><td>anon</td></tr> <tr><td>-</td><td>id</td></tr> <tr><td>-</td><td>>1 AU</td></tr> <tr><td>-</td><td>>1 W</td></tr> <tr><td>-</td><td>dis AU</td></tr> <tr><td>-</td><td>dis W</td></tr> <tr><td>-</td><td>for AU</td></tr> <tr><td>-</td><td>for W</td></tr> <tr><td>+</td><td>cop</td></tr> </table>	+	crim	+	anon	-	id	-	>1 AU	-	>1 W	-	dis AU	-	dis W	-	for AU	-	for W	+	cop	<p>One person formulates and pre-writes an incriminating anonymous text. Another person copies it voluntarily with undisguised writership (handwriting).</p>
+	crim																					
+	anon																					
-	id																					
-	>1 AU																					
-	>1 W																					
-	dis AU																					
-	dis W																					
-	for AU																					
-	for W																					
+	cop																					
(10)	<table border="1"> <tr><td>+</td><td>crim</td></tr> <tr><td>+</td><td>anon</td></tr> <tr><td>-</td><td>id</td></tr> <tr><td>-</td><td>>1 AU</td></tr> <tr><td>-</td><td>>1 W</td></tr> <tr><td>-</td><td>dis AU</td></tr> <tr><td>-</td><td>dis W</td></tr> <tr><td>-</td><td>for AU</td></tr> <tr><td>-</td><td>for W</td></tr> <tr><td>-</td><td>cop</td></tr> </table>	+	crim	+	anon	-	id	-	>1 AU	-	>1 W	-	dis AU	-	dis W	-	for AU	-	for W	-	cop	<p>One person formulates an incriminating anonymous text. Another person writes/types it voluntarily by dictation with undisguised writership.</p>
+	crim																					
+	anon																					
-	id																					
-	>1 AU																					
-	>1 W																					
-	dis AU																					
-	dis W																					
-	for AU																					
-	for W																					
-	cop																					

Relevance for applied linguistic and criminalistic classification of text types

The taxonomy of incriminating texts according to bundles of distinctive features described above is, hopefully, of scientific value in itself. It allows a more detailed and systematic classification of a large variety of incriminating texts. The main issue is, however, for linguistics as an auxiliary science of criminology to help in the analysis of real life cases of authorship attribution of incriminating letters. The prime objective of the taxonomy therefore is its application as a practical analytical tool.

Criminalistic and forensic linguistic experts may find this taxonomy an incentive to think about alternative explanations for author and writer constellations in a particular case. It may help them look at the data in more differentiated and systematic ways, and thereby allow for a more detailed and systematic examination, which in turn may help to solve the case. Enriching the analytical horizon of the forensic linguistic expert and the criminalist also is a value in itself.

The taxonomy considers text external criteria, such as number of authors and writers, anonymity or non-anonymity of the sender, media

and others. They are the starting point and also the domain that is the focus of the analysis.

Of critical importance are the two following questions: first, in what way and to which extent do the variants defined by text external features relate to text linguistic classifications proper, based on text linguistic criteria as defined e.g. by Beaugrande and Dressler (1981); Biber (1989); Brinker (1988; 1993) and others? It is salient to examine whether the variants defined in the previous section show specific differences in terms (expression) of coherence and cohesion, nature and distribution of deictic elements, anaphoric characteristics and others. Do the bundles of features as a whole or certain "sub-bundles" of features show specific correlations with special text linguistic categories, classes and criteria? For forensic linguistic analysis, the converse analytical perspective would be of special interest: do the variants show a specific similarity or difference in terms of coherence relations, anaphoric and cataphoric elements, and other features?

Most importantly, however, a truly interdisciplinary perspective can be established by relating the variants defined by matrices of linguistic features to legal categories, text types, and perspectives that people working in the area of (language and) law would work with and would be interested in following up. This would, in fact, help build the bridge between the domains of criminalistic and criminological work and research on the one hand, and the contribution that applied linguistics as an auxiliary science can make on the other. It could be asked, for example, whether there is any specific distribution of the types of variants described in the previous section and the legal definitions of text types used in criminology, such as "extortion", "blackmail", "threatening", "defamation" and "libel".

The question as to what extent extortion, blackmail and threatening letters can be considered as one text type or three sub-types of one text type, and defamation and libel letters as two sub-types of the same type or two different text types would benefit from an examination of the distribution of the variants described on pp. 152ff. If a more or less exhaustive list of textual features can be given, this may also help the linguistic classification of texts as "real" or fake letters, which is of criminalistic importance.

For the endeavors of "law people", a more detailed and empirically sound description of the linguistic structures and qualities of texts defined as extortion, blackmail, threatening, defamation and libel letters in the law may, together with additional linguistic perspectives (such as a pragmatic analysis of speech acts), help to arrive at more proficient legal definitions.

There are two important concerns to be kept in mind here. Definitional work to be done is (1) language-specific, and (2) culture-specific. The definitions would have to be developed for a particular language and culture first; for example, for the American, the British and the German systems. It is an empirical question as to what extent this can be extended to other cultures using “the same” language; for example, the British and the Australian system. In an even broader perspective, one would have to ask to what extent this would be extendable to a different language and culture, such as the German system. One could start with the legal definitions of the offenses in focus, given in a very elementary form in Gilbert (1984).

Legal definition of “blackmail” (see Gilbert 1984: 13):

Extortion of money by threat to do bodily harm, expose a wrongdoing, or disgrace the character of another.

Legal definition of “extortion” (see Gilbert 1984: 44):

Offense committed by someone who employs threats to unlawfully obtain money or other property from another. At common law, extortion applies to public officials who use the color of their office to corruptly take money or property. Extortion differs from robbery in that the threat to the personal safety of the victim is less in extortion than in robbery. Synonymous with “blackmail”.

Legal definition of “threat” (see Gilbert 1984: 120):

Statement of intention to injure person or property through an unlawful act, along with the ability to injure, in order to coerce another.

These definitions show that the main job to be done here concerns the people working in the area of law rather than in linguistics. Put in a highly simplified way: in order to compare the terminology and linguistic features with regard to some output usable for the judiciary and others, it is important to find a *tertium comparationis* for comparing the legal matters involved. In order to check the differences between extortion, blackmail and threat as outlined in the legal definitions, one would have to contrast the judicial and legal facts in an expansive manner and relate these to the linguistic descriptions discussed previously. Only on this basis – that is, a comparison of the linguistic classification of features and

of the legal and judicial definitions given and the legal matters involved – can inter-language and inter-cultural contrasts (for example, between the German and the American systems) be accounted for.

These few remarks on an interdisciplinary perspective and the applicability of the linguistic taxonomy discussed above show how much work still has to be done to allow for a full-fledged description of incriminating texts and their legal and linguistic treatment across cultures. It also shows that this would be an important, perhaps the most important, way to proceed on this rocky interdisciplinary road.

7

The System and Diagnostic Potential of Orthographic Data in Forensic Linguistic Authorship Attribution

This chapter deals with the intrinsic structure and the diagnostic potential of orthographic data in the forensic linguistic analysis of anonymous authorship. The focus is on “feature configurations” of an encompassing nature between different levels and kinds of linguistic data. All generalizations discussed are based on authentic data of the analysis of real life forensic cases. It is argued that orthographic data can never be considered sufficient evidence *alone*, but may be of relevance indirectly and are a source of data that must not be neglected in forensic linguistic authorship attribution.

Introduction

This chapter provides a closer look at the role of writing, the validity and argumentative function of written texts, in particular orthographic data, for forensic linguistic authorship attribution (henceforth FLAA). Other general and theoretical questions are discussed in Kniffka (1996a; 2003a). Some more basic questions of methodology of FLAA are discussed elsewhere (cf. Kniffka 1981; 1989; 1990b; 1990c; 1992; 2000a; 2000b; 2001; Levi 1982; Levi and Walker 1990; McMenamin 1993; Shuy 1984; 1993a; 1998a; Solan and Tiersma 2002; 2005). All theoretical questions proper cannot be discussed here in detail. The entire area of handwriting analysis (“Gerichtliche Schriftvergleichung”) is not discussed here at all, since it belongs to a different forensic field with this name, not to forensic linguistics.

All data discussed here are from authentic incriminating texts (“Tatschriften”, henceforth TS) and comparison data (“Vergleichsschriften”,

henceforth VS) from real life forensic cases. It should be mentioned that it is not possible in all cases to represent sections and quotations of texts from TS and/or VS in full detail because of reasons of confidentiality. In most cases this has no relevance for the argumentation. Readers are asked to bear with the basic fact that in forensic linguistics there are some restrictions as far as publication of authentic textual data of TS and VS are concerned. It is understandable that the offenders in criminal cases involving anonymous written texts are very often not interested in, and actually forbid further use of, their data, even in scientific papers.

In this chapter, first, some definitory clarifications and explanations concerning the nature of “frozen” written data of linguistic behavior are given. Of particular importance is a clarification of the relationship of authorship and writership (typistship) of a text. This relationship is much more complex and multifold in real life than laymen (including linguists) would expect; for example, concerning the modalities and constellations with regard to time in the actual production process of a text.

I then offer some authentic data from incriminating texts, and comparison texts of orthographic behavior are discussed and operationalized in an exemplaric fashion to illustrate the various methodological implications and caveats of an analysis of these data in FLAA.

The chapter then focuses on an exemplaric clarification of the following questions:

- Which argumentative status can be generally assigned to data of orthographic behavior?
- Which methodological insights can be gained from the analysis of orthographic data in incriminating texts and comparison texts?
- Which generalizations concerning a continuum of diagnostic potentials are possible?

In addition, some general results of research on forensic linguistic cases and a brief outlook for future research are given.

Data of orthographic behavior and anonymous authorship analysis

Terms and caveats

Forensic text analysis is concerned with three domains of data: phonetic speaker-specific data, linguistic speaker-specific data, and orthographic writer-specific data. As shown in Table 7.1, they are of a different theoretical and empirical status.

Table 7.1 Domains of data of forensic text analysis

Domains of data of forensic text analysis	Data quality status
Phonetic data of speaker-specific behavior	Iconic
Linguistic data of speaker-specific behavior	Symbolic
Orthographic data of speaker-specific behavior	Indirect symbolic

This difference in the theoretical and empirical status implies substantial differences in argumentative potential, stringency and validity of the analysis.

The differences by definition do *not* imply that only phonetic data of speaker-specific behavior are reliable and valid, and that linguistic and orthographic speaker-/writer-specific data are not. Neither do the data imply, as one can find on occasion in forensic linguistic texts, that orthographic data are by definition irrelevant for the description and explanation of speaker-specific linguistic behavior. Such generalizations are unjustified and invalid. The practical results of FLAA refute such an assumption. It will be shown that generalizations in this regard need much more differentiation and consideration. It will also be shown that orthographic data are of considerable importance for FLAA, in particular when larger feature configurations of orthographic behavior are taken into account, and even more so when these can be related to feature configurations of linguistic behavior.

For the almost trivial basic facts that phonetic data are empirically “harder” than linguistic data (grammatical, semantic and so on), and that linguistic data in general are “harder” than orthographic data, one consequence should be emphasized for orthographic data: the analysis of orthographic data within FLAA needs to be undertaken with very great caution. This will be evident time and again in the course of this chapter. Furthermore, orthographic data in text products can only be assigned *indirect* evidence. This means that they are only of interest as concomitant and, possibly, supportive data in relation to data of linguistic behavior. This results primarily from the fact that the writer (typist) and the author of the text can be different persons (see p. 171f.). Linguists, generally speaking, show less interest in the question of who typed a text than in who is the author of the formulated text. Text analysts also have to be concerned with the former question, and the question as to whether both are one and the same person or two different persons. Even if one can say that the (mechanical) act of typing, the writing production of a text, is not exactly the focus in the linguistic analysis (within the framework of authorship attribution), it has to

account for the writer/typist of a text as a descriptive entity. The linguistic analysis has to be compatible also with an analysis of the writing production process supplied by different sciences (handwriting analysis, typewriter script analysis, typewriter analysis, which each require a particular forensic expert). There is overwhelming empirical evidence that the good rapport between the various forensic sciences involved and the extent to which they are able to practice good teamwork are of critical importance for the successful analysis of a case.

In more general terms, the relevance and importance of orthographic data in FLAA can be stated as follows:

- Orthographic data can never be taken as directly valid for the authorship of a text
- Orthographic data – more exactly, the various data of written language behavior (see below) – are always of indirect relevance for FLAA only
- Data of written language behavior are by their very nature to be included in FLAA, since the textual data of TS and VS are given as written material, as “frozen” text products on paper, when they are analyzed in forensic linguistics
- Data of written language behavior are very heterogeneous and complex.

Data of written language behavior may imply highly idiosyncratic habits of writing, internalized routines of writing within or of a group as a whole, and established and traditionalized writing habits of a whole generation. They also may refer to the various ways of writing established by an orthography reform within a certain time span. Finally, they may refer to so-called “Schreibvorlagen” (‘writing traditions’) of a particular culture, standards of writing training in a particular school system. This refers to the basic fact that in each particular culture a set of general prescripts, norms and traditional handed-down collective practices of writing exist, which differ (for example, in Turkey and Germany) in how to design texts and text types in writing. These facts cannot be described in one single continuum but as a combination of overlapping continua.

If it is true that orthographic textual data are never directly valid for the question of text authorship, or that, in simplified terms, data of orthographic behavior never contain a direct trace to text authorship, but, possibly, to the typist of a text, this implies several technical caveats for the investigation. One cannot overestimate these because, intuitively, for a given relatively large number of coincidences between the writing data of TS and VS, one is inclined to give a premature analysis of authorship, or even the person of the author as a whole. To equate a text product

with the total personality of an anonymous author is a forensic linguistic malpractice and blunder, however, leading to artefacts. Any direct generalization from a single TS or a (mostly very) limited corpus of TS to the author is likely to fail. An even larger blunder would be to base such a generalization on orthographic data *alone*. Sensibly and *lege artis*, such generalizations cannot be made. This means that, even if a large or an almost total coincidence of written language data between TS and VS can be shown, it does not necessarily mean anything with regard to the authorship of the texts. A parallel of written language behavior (in TS and VS) only means that, for the *writing* process of TS and VS, partial identity can be assumed. Writer/typist of a text does not necessarily mean author of a text.

Another reason to be extremely cautious with data of written language or writing behavior results from the fact that most external data of the writing and production process of a text are not subject to scrutiny by the linguistic expert. The linguistic experts have to accept the texts given to them by the people who request the expert opinion; that is, their authenticity has to be taken for granted. This sometimes bears the risk of creating artefacts. In the well-known case of the “Hitler-Tagebücher” (so-called ‘Hitler diaries’), the handwriting analysis administered by several internationally renowned experts failed. It was correct, but not valid: the Hitler diaries and the comparison data had, indeed, been written by the same writer: the forger.

Another problem results from the fact that the forensic linguistic expert normally cannot determine to what extent additional hints given in the texts themselves – for example, dictation marks in printed or typewritten letters, notes, signs – are authentic or are added later and by whom. There is an almost unlimited reservoir of possibly wrong information, unauthentic data and distortion factors. Every linguistic expert with any experience has run across so-called “authentic” incriminating letters, in which all kinds of handwritten notes could be found that were *not* originally in the incriminating letter, but were added by others; for example, professional readers such as police investigators and maybe even other linguistic experts who previously worked on the case. If papers with such entries are copied a couple of times, it is almost impossible for the forensic linguistic expert to find out what the authentic incriminating letter looked like. There even are errors, unauthentic data and other distortion factors originated by sources that should be absolutely free of such faults; such as official transcripts produced by court clerks. I have given expert testimony in cases in which the so-called “Leseabschriften” (‘reading copies’) officially produced by the court

showed many types of unauthentic deviations from the original texts. Such “Leseabschriften” are usually made by a court when the original material cannot be read very well, when there are some parts missing or for other reasons. All such factors have to be taken into account properly, and have to be checked and double-checked by the linguistic expert, because they can generate non-validity and non-reliability of data and create artefacts.

At the end of this long and discouraging list of intrinsic problems of data representation in written language and graphic form, one should nevertheless state clearly that the very fact that data of orthographic behavior and data of (other) linguistic behavior vary, in principle, independently of each other provides a potential source of information for the analysis of anonymous texts. Its relevance cannot be overrated. The heterogeneity and complexity of forms and dimensions of communicative behavior in written language, as has been pointed out many times, may under special circumstances be a very important source for authorship attribution of an anonymous text, and thereby for criminalistic investigation.

Text authorship and writership of an anonymous text

The job of FLAA can be described as follows: a detailed systematic analysis of data of the linguistic and orthographic behavior in the TS and in the VS is to be provided, concerning the various mistakes, errors, and deviations that can be seen, individual and idiosyncratic ways of expression, linguistic quirks, other noticeable facts *and* the textual instances in which *no* such deviances of one kind or another occur. In other words, all mistakes and errors and all *non-mistakes* and *non-errors* have to be taken into account. What someone does *not* do wrong in a text linguistically and/or orthographically can be of critical importance for the overall interpretation and the matching process of TS and VS. This allows for a few further alternative classifications: the same mistake may occur in the TS and in the VS, and the same non-mistake may also occur there. All of this, along with the systematic and “performance” dimensions, and the occurrence and distribution patterns of errors or non-errors in the texts, is of great value for the analysis.

A sound comparison of textual data is only possible by comparing text type-specific variation, a postulate that is not always fulfilled or fulfillable in FLAA. One has to provide a “linguistic differential diagnosis” procedure in which one does not falsely ascribe text type-specific differences as author-specific and writer-specific variants. Every adult native speaker of German (or any other language) writes a letter of condolence differently

than an application letter to a company. This applies analogously to the differences between anonymous and non-anonymous writings (and their contextual characteristics), as well as to the various subclasses of anonymous writings in a judicial and forensic context.

The general methodological precautions that apply in the analysis of anonymous authorship of texts apply in principle also for the questions of writer- or typistship of texts. One of the most important questions at the beginning of a forensic linguistic authorship analysis is the question: Are the author (originator of the text) and writer (typist) of a TS identical or different? The people who request expert opinions as a rule take it for granted that (1) author and writer (typist) of an anonymous TS are one and the same person, and (2) that only *one* author and *one* writer of a text exists. In reality, this is frequently not the case. There are numerous different possible constellations for author-and/or writership of an anonymous TS, which can occur as a complex continuum. This continuum ranges, on the one hand, from a simple one-to-one relation of author and writer identity for one and the same form of text as a whole to, on the other hand, a complex arrangement of various (also heterogeneous) authorships and partial authorships occurring in combination with various writers for one and the same text, for different parts of a text, and for different "editions" of a text (with or without change of the media and/or channels involved). Some special forms are not even included in this continuum, as in a dictated text, a text dictated at gunpoint, a text with multiple authorship and writership with free or with forced roles, a text with alternating authorship and typistship within it, a text with disguised or faked authorship and/or writership. Experience shows that of 100 anonymous TS in relation to the total of the textual component constellation, no two texts are exactly identical. The problems of text authorship and writership of the VS have been mentioned above briefly. In real life, the linguistic expert has to examine these questions rigorously and meticulously, independently of the wording that the request for the linguistic expert opinion contains and what the people who make that request think about the data. The most frequently occurring questions are: Is there only one author, or are there more authors involved? Is there only one or are there more than one typist of the (total) text or of specific sections or parts of a text? What configuration of partial author identity, writer identity, or combined author and writer identity can be assumed for the TS and the VS? Neglecting a detailed examination of these questions increases the potential for creating artefacts.

Status and repertoires of data of written language behavior in anonymous authorship analysis

The complexity, polyvalence and heterogeneity of data of written language behavior mentioned above are described in this section with some detailed examples. There is also discussion of some general characteristics followed by a heuristic taxonomy of selected data of written language behavior, in particular orthographic behavior, which has turned out to be relevant in linguistic expert opinions.

Some notes on data representation and legends

Original text data of TS and VS and other text examples are given in *italics*. Original data occurring in the text (representing, for example, grammatical and/or orthographic errors or deviant forms) are written in front of a slash (.../), the intended correct forms are given after the slash (/...).

Example: *ihr Firma/ihre Firma* or *Ihre Firma*. The form actually occurring in the text is *ihr Firma*, which is ungrammatical in German; the correct form should be *ihre Firma* ('their company') or *Ihre Firma* ('your company', the second person possessive pronoun, polite form).

Data of written language behavior relevant for forensic linguistic analysis in cases of anonymous authorship

In order to gain sensitivity about the complexity of "data of orthographic behavior" (henceforth: DOB) and to problematize the terms in FLAA, a few pre-theoretical comments concerning terms such as "dimensions of striking examples", "peculiarities", "deviations", "mistakes and errors", and "features" should be made. First, it should be noted that the term "feature" ("Merkmal") has been used in forensic linguistic texts in a somewhat inflationary manner. Due to the ambiguity and polyvalence of the term, it should be avoided as much as possible, except in the combination of "feature configuration", for which there seems to be no alternative.

What can be stated in a pre-scientific and pre-theoretical way concerning "peculiarities" or "deviations" of human behavior?

(1) A person may have specific peculiarities. She/he may be very tall or very small. Such visually perceivable data of outward appearance are undoubtedly of a different nature than the characteristics described below, and they seem somewhat "more invariant" and perhaps "more

solid". One should be very cautious, however, about declaring them to be generally invariant, since extremely tall or small human beings may appear, depending on the situation, to be of a different size to different observers. It cannot be presupposed that it is an invariant size without first making comparative analyses of observational data of different people and with different measurements.

(2) An adult man can have an extremely high voice, almost like a castrato. This acoustic quality has to be judged differently than the aforementioned and also the following.

(3) Human beings can talk extremely fast in situations for which no indication is given of such outward conditions as stress, threat or emotional engagement. They may speak like this normally. This property seems to represent a more dynamic behavior than the characteristics mentioned under (1) and (2), which intuitively and *a priori* would be described as more solid properties. The scientific analysis and the explanation of the differences, as such, are not of concern here. It is the difference in the dynamic or non-dynamic status of the qualification that is to be described.

(4) All three peculiarities or properties mentioned in categories (1) to (3) have in common that they can be found in relatively large populations of speakers, and these probably would not be described as idiosyncratic or highly speaker-specific properties by the average adult speaker of German. The latter, highly idiosyncratic speaker-specific feature can, however, be attested to persons with a so-called articulation defect ("Sprachfehler"), such as sigmatism.

A sigmatism can be decoded by interlocutors quite clearly, even if not in exactly the same way by all. Interlocutors who are in close or daily contact with the sigmatic speaker may pay less attention to this particular phenomenon than interlocutors who are listening to him or her for the first time.

No claim is made here that sigmatism is a factually person- or speaker-specific invariant property of the speaker in question (differently from the peculiarities listed under (1) to (3) above). All that is envisaged here is that adult native speakers of German would probably consider this a speaker-specific and somewhat invariant feature of articulation. What is important here is only that, for orthographic behavior, one would have to differentiate between features that are *factually observable* data (by scientific instruments and in valid and reliable observation) and features that are *thought* to be speaker- or writer-specific. This is an important difference. Some speaker-specific idiosyncrasies or deviations seem more obvious than others. This difference in impression cannot be shown to

actually exist in a sound scientific analysis, however. On the other hand, there are scientifically sound highly speaker-specific features of written language and/or orthographic behavior that are not perceived by laymen in the same way as speaker-specific data. The perception in general can be rather heterogeneous, and people can more or less consciously assess them. What is important here is that all these phenomena and more or less obvious peculiarities of verbal and orthographic behavior need a sociolinguistic and/or ethnographic dimension for an adequate description and explanation.

(5) Yet another “peculiarity status” can be assumed for stuttering, for which several dimensions and types can be distinguished intuitively. Some persons may stutter when they are nervous but not stutter in other situations. Other persons may stutter even when they are not nervous. One and the same feature cannot just be assumed to be invariant for one and the same speaker, let alone across larger speaker populations. Features and peculiarities of written language behavior are by no means *ipso facto* invariants of writing behavior. It is a basic fact – albeit an almost trivial observation – that any speaker of a language sometimes makes mistakes, errors and deviations in orthography that they do not make on other occasions. It would actually be necessary to assume different styles of orthographic behavior, as Labov (1972a) has defined for linguistic behavior, depending on the attention that speakers give to their manner of speaking (such as casual, formal, reading, word list and minimal pair style). It seems useful to define such different “attention styles” for orthographic behavior as well. To do this, different forms and dimensions of correction behavior would also have to be taken into account. In different text types and different situations, people produce mistakes and errors in different ways in writing. In a formal test or in an exam in class, a syndrome of errors and deviations is likely to occur that is different from those found in an informal letter or an e-mail addressed to a personal friend. In an official application letter, one can expect yet a different error and deviation syndrome in orthography than in a letter to the editor of a newspaper. What is also implied here is that the more attention the writer/typist dedicates to the writing/typing process, the fewer errors and deviations will occur. With a smaller amount of conscious attention, more orthographic errors and deviations are likely to occur. This is a rather small part of the data. In addition, different repertoires of mistakes, errors and deviations in and for different text types and situations of writing behavior, and different situational contexts have to be accounted for.

(6) There is yet another complication in terms of dimensions of peculiarities. It is possible, for example, that someone belongs

socio-economically to the upper class, and shows behavioral syndromes in large parts of everyday life that represent this social class (such as in clothing, eating and contact with other people), but that the same person may also show behavioral syndromes on other occasions that one would call atypical, or even exceptional, for his/her class and also the rest of his/her everyday behavior. It may be possible, for example, that she/he wears working class clothes, associates with people of lower class, eats in a fast-food-chain and so on. This behavioral peculiarity is, by dimensions and range of variation, structurally different from all those previously mentioned. It is more complex and allows various heterogeneous syndromes.

The heuristic criteria of differentiating characteristics, properties and features of human behavior *mutatis mutandis* can, in principle, also be assumed for written language behavior. This does not mean that a direct transfer from the dimensions (1) to (6) is intended, since this is probably not possible. The aim here is to bring to attention that one cannot simply base the analysis on static material invariant properties of the orthographic product as represented in a written text. Certainly, this is not possible in FLAA. It also is questionable whether it is possible to assume just *one* continuum of peculiarities and features of written language behavior. There are likely to be various complex continua that have to be defined and determined for their differences and parallels in a more complex, theoretical and empirical analysis. It is also possible that there is a hierarchical structure of orthographic deviations, as in linguistic behavior, for the occurrence of mistakes, errors, deviations and peculiarities, as well as the non-occurrence of each of these. In addition, interrelations between orthographic and linguistic behavior have to be accounted for.

Towards an operationalization and systematization of deviations in orthographic behavior in an anonymous authorship analysis

An adequate description and explanation of orthographic data that seems relevant for FLAA requires clarification of prerequisites concerning the performance status of "data of orthographic behavior" (DOB). Without any claim of exhaustiveness, the following points can be made:

DOB are given printed on paper, wrongly suggesting that they are unambiguous concerning their descriptive and explanatory status. One and the same written product may have to be explained heterogeneously. Such an explanation can signal entirely different things. This

has been recognized for decades in error analysis and other areas of applied linguistics, and is briefly summarized here.

First, one can distinguish items that are only orthographic deviations; second, one can identify deviations that are orthographic and linguistic (grammatical and stylistic) deviations at the same time. This means that there are four theoretical possibilities as given in Table 7.2.

Table 7.2 Deviations in orthography and language

(1)	(2)	(3)	(4)
+ linguistic	- linguistic	+ linguistic	- linguistic
+ orthographic	+ orthographic	- orthographic	- orthographic

A German example for (1) would be *das Grösse/die Größe*; for (2), *die Grösse/die Größe*; for (3), *das Größe/die Größe*. Classes (1) to (3) need no further explanation, but class (4) in itself is heterogeneous and worth discussion. It refers to ways of behavior from a relatively large “gray area”, for which one cannot clearly define a status of deviation, let alone an error, but which nevertheless evidences some kind of a continuum of peculiar or unusual behavior. For this, the following paragraphs give numerous incidences; for example, the writing of <ss> and <ß> in German; the writing of two separate words or one word in a manner different from the existing norm; the use of a comma instead of parentheses and vice versa; the writing with or without a hyphen in expressions with numbers (such as *3-Jahresvertrag*) and others. As stated above, graphemic deviations may be indicators by themselves and, at the same time, be indicators of linguistic deviations. The methodological question of critical importance is that in no case do the data themselves show clues that indicate which exact status applies. For each case, a “reconstruction” (Corder 1971; James 1998) has to be made, in addition to a “description”. Every example needs some reconstruction of the intended form with an “aitiology” description and analysis of the reasons why a particular mistake in this particular form occurs. Generally speaking, analyzing a forensic linguistic context is no different from analyzing student papers in class, at least as far as the general methodological postulates are concerned. In both contexts, one and the same form can have several alternative explanations, and some additional effort is needed to determine which alternatives are present. A well-known example from error analysis is that for the written product **die Schüler geht* two intended versions can be reconstructed: (1) *Die Schülerin geht*, (2) *Die Schüler gehen*. In the first case, there would be an error in gender

(*Schüler* instead of *Schülerin*). In the second case, there would be an error in number agreement (*geht* (sg.) instead of *gehen* (pl.)).

Orthographic behavior may be an adequate graphemic representation of grammatical and linguistic knowledge of the writer (for example, of a learner of German as a second language, who does not know better, so to speak). It is also possible that it is an error generated by the writing behavior, a so-called “*Flüchtigkeitsfehler*”, which does not represent the actual and factual level of knowledge of the writer. For the latter, which is frequently referred to as a “slip of the pen” (see Fromkin 1980), the most probable explanation is that it is a writer who is a native speaker and writer of German.

This example shows that there are numerous factors that have to be taken into account and have to be analyzed very carefully concerning their covariation, if the production of a particular graphemic product is to be analyzed sufficiently well. This is of critical importance for an adequate description and explanation of a TS and the VS.

In this chapter, more detailed statistics of the occurrence and text frequency of orthographic errors and classes of errors in TS and VS cannot be given, but a few comments will clarify the differences of text types. One general statement seems necessary: Deviations of the type *ihr Firma* instead of *ihre Firma* or *Ihre Firma* occur with relatively high frequency in German in TS and VS. It is hard to find any threatening letters, defamation letters and extortion letters in which such errors do *not* occur. Since there are so many possible alternatives, it is impossible to analyze the particular data given here without taking further contextualized data into consideration. A form *ihr Firma* may indicate insufficient knowledge of German grammar by a non-native speaker (for example, a speaker with a native language that does not have a grammatical category of gender). It may also indicate an emotional engagement of the written texts that leads to a slip of the pen by a native speaker. There are various combinations possible between the two, which is complicated by the fact that the second person polite form of address in German is written with capitals by some speakers and with no capitals by others. Another complicating factor, which also frequently occurs in TS and VS, is the fact that the writing behavior in a particular text item is inconsistent, which means that no simple explanations in the reconstruction of the deviations can be given. Finally, in quite a few cases the material form of many TS and VS is of such bad quality that no clear decision can be made whether a form *ihr* or *ihre* is given. This also holds for copies that are given to the linguistic expert by the authorities.

Table 7.3 Three heuristic classes of data of orthographic behavior in German

-
- (1) Correct, normal spelling, for example: *wir* ('we'), *innerhalb* ('within'), *schwarz* ('black'), *wäre* ('would be')
 - (2) Acceptable, but somewhat unusual variation of spelling, e.g.: *ggfs./ggf.* ('in case'), *zB./z.B.* ('e.g.'), *Mitarbeiter/innen/Mitarbeiter/Innen* or *MitarbeiterInnen* ('(male and female) employees')
 - (3) Entirely wrong spelling severely violating orthographic norms, e.g.: *wier/wir* ('we'), *inerhalb/innerhalb* ('within'), *Imobilie/Immobilie* ('real estate'), *währe/wäre* ('would be'), *Bahrzahlung/Barzahlung* ('cash payment'), *swarz/schwarz* ('black')
-

One comment should be made concerning the relation of orthographic data in TS and VS. Table 7.3 represents somewhat simply the general structural possibilities that exist here, to the extent that one can distinguish correspondences and non-correspondences of orthographic behavior in the TS and in the VS.

In real life, a complex continuum rather than these (simplified) three classes has to be accounted for, representing different kinds, modalities and degrees of deviations or correctness of orthographic behavior. Generally speaking, non-correspondences of orthographic behavior in TS and VS represent a more complex structure. If there is inconsistency of writing behavior in both the TS and VS, there is a yet more complicated covariation pattern of occurrence of forms. To name just one example concerning text type-specific rules of orthographic data: in a letter addressed to the president of a university or to the Chancellor of the Federal Republic of Germany, one would not use an abbreviation for the final formula of a letter. In other words, one would not abbreviate *Mit freundlichen Grüßen* as *MfG*, which is absolutely normal in short notes between colleagues/students.

The data discussed so far show two basic facts. First, in reality only continua exist, not clearly separable classes of deviations or correct forms.

Second, deviations such as those identified in (2) and (3) may occur *concomitantly in one and the same text*; for example, in a TS. Each deviation taken on its own may not mean much. But in their total covariation, the deviations, reconstructed as indices of orthographic behavior and analyzed together with various other factors, may indicate a *syndrome of orthographic behavior* that can correlate significantly with linguistic data for the total criminalistic evaluation. Such a global configuration of data can yield probability arguments, suggesting, for example, a non-native speaker of German as the writer of a text.

The core questions in the analysis of DOB are no different from those in the analysis of data of linguistic behavior: Which variants varying, in principle, independently of each other occur *in combination* in a particular text? What is the total configuration of data of orthographic behavior in this text? Linguists used to call this “Merkmalkonfiguration” (‘feature configuration’), when what was meant was “Merkmalausprägungskonfiguration” (‘feature value configuration’). It should be noted that the analysis is not only concerned with relatively large-scale and global differences, such as the writing habits of writers of different mother tongues (“scripts”), but also the writing habits of smaller categories, such as those used by people of different levels of education, different generations, different local regions of a speech community and different professions. There is a large number of characteristics which, taken as configurations on a larger level, may yield important indicators for the analysis of the writership of a text. They concern seemingly unimportant side instances of orthographic behavior, such as the usage of abbreviations, punctuation marks and similar characteristics. In this context, the specific cultural tradition described as the “Schulvorlage” (‘model of writing taught in schools’ as described, for example, by Wagner 1997 for Turkish) would have to be taken into account. It would have to be determined whether and how this is of relevance for the description of orthographic behavior. Also, the differences in orthographic standards between the German-speaking countries (Germany, Austria and Switzerland) should be taken into account.

There is yet another dimension of orthographic behavior that does not appear to have been accounted for to date. By this, a class of DOB is meant which, as such and even in its co-occurrence, does not show any particular peculiarity and is completely in line with the general rules and forms of orthography. If used with *extremely large frequency of occurrence*, however, it shows a highly remarkable peculiarity. This excessive use of orthographic features has to be differentiated further; for example, in relation to different text types and parts of a text. This refers to features of orthographic behavior, for example, in the use of punctuation marks, as represented in Table 7.4.

It is safe to say that features and peculiarities of orthographic behavior concerning punctuation marks have been largely neglected in FLAA to date. They are of particular importance concerning the occurrence in different text types. As described elsewhere (Kniffka 2001), it seems possible that these indicate differences between defamation and threatening letters on the one hand and extortion letters on the other. In defamation and threatening letters, an excessive usage of exclamation

Table 7.4 Features of orthographic behavior: excessive and (simultaneous) repeated use of graphemic signs (mainly punctuation)

-
- Hyphenation
 - Graphemic representation of numbers instead of words in a text discourse
 - Graphemic representation of words instead of numbers in tables, in the date line of a letter and so on
 - Excessive usage of quotation marks for single words (without quotation status)
 - Excessive usage of abbreviations (in text discourse)
 - Excessive simultaneous usage of several (instead of one) punctuation marks, such as question marks (???), exclamation marks (!!!), parentheses (- - - -) and so on.
-

marks (!!!), question marks (???), and so on, can almost be considered a text type-specific stereotype of these two classes of incriminating letters. It is next to impossible to find any defamation and threatening letters in which these features of punctuation do *not* occur. On the other hand, an excessive usage of punctuation marks seems extremely rare, if occurring at all, in extortion and blackmail letters. It is possible and necessary to describe characteristics of orthographic behavior in covariation with these in linguistic behavior. Perhaps one could set up another more general continuum concerning the degree of emotional involvement of an incriminating text.

Even the latter-mentioned excessive simultaneous use of punctuation marks is not without exceptions that can be shown to exist in certain text types. As stated before, there are hardly any extortion or blackmail letters in which exclamation sentences, requests or directions are followed by several exclamation marks at a time, such as *Keine Polizei!!!* ('No police!!!'). Far more typical and frequently occurring in extortion letters are very non-emotional, distanced forms like *Keine Polizei!* ('No police!'). Defamation letters abound in punctuation marks of this sort. This means that also in text types in a criminal context some conventionalized standards exist concerning how some means of the graphemic repertoire are being used. This probably represents a domain of non-conscious usage. Here is a field for future research: it is necessary to distinguish typical and atypical writing habits for each particular text type and subtype such as extortion letters, threatening letters and defamation letters. When such standards, as briefly outlined above, can be empirically and clearly defined, it may be useful to note exceptions of orthographic behavior concerning the "standard" usage in a criminal context. It could indicate a non-professional way of speaking and writing opposed to a professional one. This, in turn, may give clues to the writing production and the writer of a text. In quite a few criminal cases there are

Table 7.5 Styles of orthographic behavior

-
- (1) Informal, “careless” writing (an attitude that “it doesn’t matter”);
 - (2) Informal writing;
 - (3) Formal writing (realization of norms, how you write in a speech community);
 - (4) Formal writing with corrections (“correct orthography matters”);
 - (5) Formal writing with particular attention to correctness (e.g. students’ exam papers, application letters and so on);
 - (6) Very strongly accurate formal orthographic writing (documents, legal texts and so on).
-

many imitators (“Trittbrettfahrer”) who can be distinguished from the “real” professional criminals with the help of the orthographic and linguistic data of the kind mentioned.

A particular class of orthographic behavior of importance for the analysis of the writership of TS depends on a special thematic area of the information given, such as the written representation of names in a foreign language. An example of this is the Chechnyan town name *Groznyj*. A German writer may write *Grosny* and a British writer may write *Grozny*, both instead of the correct transliteration, which is used by people capable of transliterating Slavic idioms correctly: such differences in written representation can also be used by criminals as a way to disguise their authorship and writership.

As mentioned before, it may, and should, be possible to define “styles of orthographic behavior” depending on the attention the writers give to this activity, comparable to the linguistic styles of verbal behavior defined by Labov (1972a). Table 7.5 lists some tentative heuristically defined styles of orthographic behavior.

A heuristic continuum of “styles of orthographic behavior” would have to be drawn up for each particular medium and text type. It would have to be drawn up for handwritten texts, typed texts, e-mails, and other text types and sorts and registers, which would have to be distinguished according to “ethnographic components of writing”.

Exemplaric description of ambiguous orthographic data in FLAA

In this section, just a few examples of ambiguous or, more precisely, *polyvalent* DOB are discussed. Strictly speaking, the term “ambiguous” is somewhat redundant here. In fact, all orthographic data are ambiguous or polyvalent by their very nature. This holds in a multiple and rather heterogeneous sense as shown below. Frozen orthographic data of a text

can, similarly to linguistic data, be described as products of various different behaviors. This rather elementary form of polyvalence is not at all trivial. One example is the way that triple consonants or double consonants respectively are written according to the orthography norm in existence before 1996. Examples such as *Schiffahrt* (a compound from *Schiff* + *Fahrt*), *Wettturnen* (a compound from *Wett-* + *turnen*) with double consonants before a vowel; examples such as *fetttriefend* (a compound from *fett* + *triefend*) and *Balletttruppe* (a compound from *Ballett* + *Truppe*) with triple consonants before such consonant doubling or tripling existed as a rather clear-cut orthographic norm in German (or more than one norm, if the customary lifespan of a German writer is considered). This is the official norm as such. One has to assume, according to the data of some written products that are the object of forensic linguistic analysis, that this norm was known by a certain population X of German writers perhaps theoretically, but was not actually applied or followed in actual everyday writing behavior. One can suspect also that there is another population Y that did not apply this norm consciously and/or conscientiously, and possibly, a third population Z that did not know anything about the existence of such a norm at all. It is a well-known fact, and every experienced forensic linguistic expert has many examples of this, that in TS as well as in VS all theoretically possible orthographic variants occur, in addition to some deviations that one would not theoretically think possible.

In the type of example represented by triple versus double consonants of a kind in compounds, for which at least a norm existed, which was followed or violated by various populations of writers in rather different ways, a class of examples of orthographic behavior can be defined for which such a clear norm has not been in existence in German to date. This refers to cases such as the usage or non-usage of a hyphen with expressions or words with numbers, such as *5-Jahresvertrag* (see Kniffka 2003a). Teachers of German in elementary schools, teachers of German as a foreign language and people that have to deal with the teaching of orthographic norms will not appreciate this fact. Criminalists and forensic linguists acting as their auxiliary science helpers will like the fact that there are relatively wide areas of linguistic and orthographic usage that are not governed by unambiguous norms or by any norms at all. There is a rather wide spectrum of possible realizations, even in the standard variety of German grammar and German orthography.

If there are no unambiguous but only polyvalent data of orthographic behavior, it is safe to assume at the same time that orthographic (and linguistic) data show *gradual* differences in ambiguity and polyvalence.

Together with the aforementioned fact, orthographic data can be distinguished into data that are only and purely orthographic data and orthographic data that also pertain to other linguistic levels of description, such as the grammatical levels of morphology, syntax and semantics. This means that the descriptive category “data of orthographic behavior” cannot be simply characterized as ambiguous, but that heterogeneity, ambiguity and polyvalence are the rule for this class of data in FLAA.

All this is just the tip of the iceberg, however. The facts mentioned so far bear some consequences for the operationalization and differentiation of the instruments of the description. The most important seems to be a two-fold distinction. One can distinguish (1) a systematic, structural diagnostic potential; and (2) a factual diagnostic potential (empirically data and case oriented). It is an essential fact that, within both (1) and (2), several different levels and steps can be distinguished in a gradual scale. For (1) the different levels for the systematic structural diagnostic potential on the level of language system, and also the intrinsic regularities of language use, the various differentiations – text linguistic, text type-specific, genre-specific and so on – have to be accounted for.

In (2), case oriented data are to be described. This means (a) regularities that can be generalized across different cases of FLAA that apply to all kinds of extortion, threatening, and defamation letters; and (b) case-specific regularities that are only valid for a corpus of texts or for one particular real life case. It may also apply to case-specific data that are applicable to one single item of a text of TS and/or VS only.

The systematic structural diagnostic potential refers to the validity that can be derived from the orthographic system (or systems) of a language in the diagnosis of DOB. For (2) it renders the actual and factual argumentative power that the totality of DOB has *together* with text-external data existing in a particular case. The last-mentioned determination is decisive here: all data of the orthographic system and also writing usage in a speech community alone are not valid for the analysis of the writership of a text product. Valid indicators in FLAA, as single items and in complex bundles of features, have to be related to external (or non-text-internal) data of various kinds and dimensions.

All DOB that could be set up for the system and the usage of writing have to be related to a certain number of writers/persons >1 . Depending on the “idiosyncrasy coefficient” (henceforth: ISC) of the particular class of DOB, it will be of different size and text frequency (see below). These data are used by a population of speakers/writers counting in the millions, tens of thousands, thousands and hundreds. They can never apply to one single speaker and/or writer alone. This means that speaker- and/or

writer-specific or even speaker- and/or writer-identifying DOB do not exist. The same applies to data of linguistic behavior. If we were to communicate in “unique language specimens”, it is very likely we would not communicate at all. This does not mean that there are no unique text products in terms of language and writing (consider some experimental modern poetic texts). It does mean, however, that this type or these classes of data do not really play a significant role in FLAA. I am not aware of any real life TS or VS that have unique forms of orthography and/or language in such a number that they would be of any importance in a practical analysis. In most cases, so-called “idiosyncratic” data are (more or less large) group-specific orthographic data that can be described in their interrelations with other data in terms of a differential diagnosis, and may in this way be helpful for a successive delimitation of the number of possible alternative hypotheses. Decades ago, Sapir (1927) observed that most features that are claimed to be idiosyncratic or idiolectal are in fact (mostly) genre- and group-specific features.

The descriptive categories of the “structural diagnostic potential” and the “factual diagnostic potential” have not been described and analyzed in precise scientific terms to date. In particular, the description and explanation of certain classes of data of language use still seem to be underdeveloped. It is not enough, in a case of type (1), to describe the opposition of triple versus double consonants of a kind, or state that in a certain TS the orthography norm is being followed or is not being followed. It is not even sufficient to claim that, in a certain TS X, the writing of double consonants of a kind, or, in a TS Y, the exclusive writing of triple consonants of a kind – both of which are, in any case, very rare – are evidence of individual authorship. One has to present a rather detailed picture of the total distribution of items and features in all TS and VS and describe the structural (positive or negative) deviation status in a more complex and differentiated way. Even then, such an exact and adequate description would not be sufficient for the aims and tasks that an anonymous authorship analysis has to solve. It is well known that certain syndromes of orthographic behavior cannot be matched to particular types of writers, although this is done frequently by all of us in an unscientific way every day. If, for example, the writing of triple consonants of a kind and double consonants of a kind would exactly follow the norm, many laymen would not hesitate to consider this an accurate or pedantic way of German orthography and conceive of a writer in this way. If, in the very same TS, one also encountered a large number of orthographic errors – for example, in capitalization and

non-capitalization, in spelling one word as opposed to two separate words, in the use of punctuation marks and so on – such a hypothesis would become absurd very quickly. If, conversely, someone would violate the then existing norm of German orthography (in effect before 1996) by writing double and triple consonants in compounds, this would not establish anything useful regarding their factual knowledge of German orthography. Even if an orthographic representation massively deviating in the “amplitude” occurred – such as, as a theoretical example, the writing of quadruple-consonants (for example, **fetttttiefend*) – this would not allow valid matching of this behavior to a certain type of writer, although it undoubtedly would suggest a rather high degree of an idiosyncrasy coefficient (see below).

It has been described elsewhere in more detail (Kniffka 1981; 1990a; 1990b; 1990c) that a direct one-to-one matching of individual items and/or features of writing syndromes to one particular writer/typist is, on the whole, scientifically unsound and constitutes an analysis not done *lege artis*, although there are cases of handwriting analysis and forensic linguistic analysis that show such a procedure. One can sum up that, to date, too few data of the writing system and writing use(s) have been analyzed adequately to be available and applicable for FLAA in any practicable way.

This applies even more to the factual diagnostic potential (2), which is illustrated by redundant repetition of one and the same sign in punctuation. As an element of the systematic description, one must state that only some punctuation marks – such as question marks (???) and exclamation marks (!!!) – are used simultaneously in multifold form and that others (such as commas, full stops, colons, semicolons and so on) are not used as “several of a kind” in normal (non-computer) communication. In some cases, the repetition of the semiotic form (such as dots) would create a sign of different semiotic status (omission marks), and in others a repetition of the same sign in one place seems possible but is not customary. We know far too little about all of these regularities and systematic structures of semiotic systems, such as those of punctuation marks, today. Certainly, what we do know is not empirically based. If one can prove by hard empirical data that there is a clear text type-specific occurrence of a single punctuation mark as opposed to multiple simultaneous use of punctuation marks (such as !!!; ???) in complementary distribution for extortion and blackmail letters on the one hand, and threatening and defamation letters on the other (and, maybe, other text types that are not the object of forensic linguistic analysis), one could be able to describe configuration patterns of data from this semiotic area

(punctuation) with other semiotic dimensions of language (cf. Kniffka 2001). If, for all these, a total configuration can be determined and stated in quantifiable terms, it would be a useful working hypothesis for a syndrome of orthographic behavior of and in texts. It could be correlated with other data of text-internal and -external denomination. This would have to be related to external criteria, such as a suspect who did not finish grade school successfully, someone who had a reading/writing legasthenia syndrome, an aphasic patient, and someone who had successfully passed an exam in linguistics and/or is working successfully in their profession. Such information would add additional perspectives to the analysis.

Such a configuration syndrome could be related to “abnormal” or non-customary distributions within particular text types. If, for example, the occurrence of a multifold usage of exclamation and question marks is atypical in extortion letters, an exceptional occasion of this (that is, three or more exclamation and question marks) in this text type could yield clues for a non-professional writing syndrome and/or a writing disguise (some very interesting remarks on the argumentative potential of clues as an operational principle of forensic linguistic handwriting analysis are discussed in Davis (1996)). Whether the people who request linguistic expert opinions on authorship attribution would be able to successfully use such an analysis with consecutive differential diagnostic steps remains to be seen.

The general aim of FLAA is to derive the most complex configurations of data that, in principle, vary independently of each other. The principle of cumulative evidence (the more data you have, the more you can state about a text) is more or less the core notion of FLAA. The limits of such an analysis are more easily determined than the prospects. The most important restriction is undoubtedly that a one-to-one matching of text products and text writing syndromes to (types of) text authors and writers is at present unscientific and cannot be made *lege artis*. There are numerous examples of German forensic linguistic expert opinions that have unjustified classifications and direct matchings of this kind. At one time, a so-called “linguistic expert” defined “*weil*-types and *da*-types of text authors” (sic) (both meaning ‘because’) depending on whether an anonymous author used one or the other as a causal conjunction predominantly, or exclusively, in his letters. In a similar way, other forensic linguists have defined types of writers depending on whether they wrote the German conjunction *dass* (‘that’, representing the present standard of German orthography since 1996) or *daß* (representing the standard of German orthography before 1996). The problem is that, in

several instances (if the incriminating data are large enough) both ways of writing may occur. In most cases, the simple fact that there is no clear-cut distribution of one form or the other forbids the definition of a type, which also cannot be justified for general theoretical and empirical reasons. Persons may, for example, use different ways and syndromes when writing, depending on the degree of formality of the situation, the text type and so on, or even to create a variation to disguise their writership.

A note on the concept of “idiosyncrasy coefficient” (ISC) of orthographic behavior, which is introduced here as a heuristic working hypothesis: the ISC of different classes of DOB is by definition of different size. Graphemic representation X_1 can be rarer, more exceptional and more peculiar than the graphemic representation X_2 (as a representation or way of writing of one and the same variable X in a class of texts, or as graphemic representations of a variable Y).

The ISC of a comma with a curvature towards the left (waning or decreasing moon) instead of a curving to the right (waxing or increasing moon) – as is customary with German writers in a (German) text – is considerably higher than the written representation of a less idiosyncratic use of two or three question marks instead of one, the use of parentheses instead of a comma or vice versa, and the repeated use of exclamation marks instead of a full stop at the end of a sentence. If, for example, Arabs writing German use the first-mentioned curving of a comma (conditioned by the way of writing from right to left), this alone does not represent writer-identifying orthographic data. A comma curving to the left can also be created by different classes of writers; left-handed people writing with their right hand, people under alcoholic intoxication or people writing on a bad or moving surface (for example, in a train) and so on. This shows that only together with other linguistic text-internal and text-external data can some working hypotheses for delimitating the writership of a text be administered. It presupposes that, quantitatively and qualitatively, an unambiguous set of data has been given, and also that enough instances of usage of commata in a text occur and that they all follow this characteristic.

Another example of a relatively high ISC is the writing of omission marks (three dots, called “ellipsis” in some Anglo cultures and to be distinguished from the linguistic term “ellipsis”) at the end of a quotation. In an anonymous TS, over a dozen instances of these three dots in invariant form occurred. The form was <..>¹ (instead of standard <...>). It seems that the writer of this text uses a relatively high representation of the three dots in a quotation occurring at the end of a

sentence (and/or paragraph). The German standard is to leave out the fourth dot marking the end of the sentence after the three dots and the quotation marks (cf. DUDEN 2000, §100: 1151). Although the orthography- DUDEN still has a high authority, in particular with laymen German speakers of today, it is a theoretical guide and not based on the actual usage norm. To obtain some more reliable data on the factual “official” use of this form, a spontaneous inquiry was administered with the teaching staff of courses for professional secretaries at several chambers of commerce (Industrie- und Handelskammer (IHK)) in Germany. The result was that subjects stated unanimously that the variant occurring in the TS <..”.> was *not* known to them, that the correct form would be <...”>, and that in the “DIN-norm 2008” (which is the collection of norms of spelling for professional secretaries) such a form was not mentioned at all. It should be noted here that various linguistic colleagues that I also asked had never encountered or even used a variant <..”.>. One should also mention that all linguists I asked did not know of any rule about how to use German orthography properly in this instance.

This, undoubtedly, is an example of a rather high ISC that has no significant impact for the analysis of orthographic behavior in anonymous authorship. In other words: DOB with a relatively high ISC do not yield a guarantee for a *liquet* concerning the forensic linguistic question of authorship and/or writership. In the particular case in question, there was no comparative peculiarity <..”.> in the VS. If it had been found in a similarly clear distribution and frequency of occurrence as in the TS, it would perhaps have been an indicator for partial identity of writership of the TS and the VS.

It should also be mentioned here that a relatively high value for the ISC does not correlate in any significant way with the number of alternative hypotheses – for example, in a negative reciprocal way – or with the extent of ambiguity of a form. The deviations in both cases (a comma curved to the left and an idiosyncratic way of writing three dots at the end of a paragraph) may have been done by a relatively large number of writers. If, for example, all participants in a class of prospective secretaries had been taught by a teacher that <..”.> is the correct norm, they probably would have used this in the particular case. Static characteristics of orthographic data, as mentioned in the introduction above, are never *ipsa natura* valid for the analysis of writership of TS and VS. It is important to arrive at generalizations in this regard by reconstructing dynamic entities of DOB and to check the configurations with other dynamic data of linguistic behavior.

There is one final comment on the empirical status of orthographic data such as the two last-mentioned examples with a relatively high ISC. It is, beyond doubt, a fact of real life forensic linguistic analysis that data of this kind are relatively rare, if not very rare, in anonymous TS and VS. In many thousands of pages of TS and VS in cases in which I or my colleagues gave expert testimony in the last 35 years or so, I have found only these two examples to which, possibly, can be assigned such a relatively high ISC-value. The usually occurring DOB in TS and VS, including errors and deviations as well as correct forms, occur by the thousands, tens of thousands and millions. In German these are, for example, writing deviations of capitalization and non-capitalization, writing in one word or more than one word, writing of <ß> and <ss> according to various norms, among others. As forensic linguistic data, they occur at a much higher frequency than the ones mentioned above. The decisive question is to what degree and in what way DOB can be related to text-internal and text-external data, and to what extent these configurations can be described in quantitatively unambiguous terms for TS and VS.

Summary

The discussion above shows that a systematic analysis of DOB is, indeed, an important intrinsic part of forensic linguistic analysis of anonymous writership. From these data, indirect evidence can be gained for the question of text authorship, depending on the nature and quantity of the particular data under investigation. This is to be done *cum grano salis*, and there appears to be no other generalization possible across cases. In other words, to include orthographic data in an authorship analysis does not, by itself, mean that the analysis is closer to a satisfactory solution. Generally speaking, one can say that a holistic analysis of all data of semiotic behavior in a text can be relevant for the criminalistic investigation of questions concerning authorship identification.

It has also been shown that a considerable effort has been made, and promising results have been reached, in forensic linguistics, but that considerable research and practical analysis still needs to be done.

One does not know whether and to what extent DOB, as a structural diagnostic potential and as a factual diagnostic potential, may be relevant for anonymous authorship analysis because this is dependent on a thorough analysis of all the dynamic data available being conducted. It is necessary to state (due to several erroneous claims in the German

forensic linguistic literature) that it is not justified to exclude DOB generally for forensic linguistic authorship analysis.

It should be noted again that the analysis of DOB, in the analysis of data of linguistic dynamic behavior, allows only generalizations concerning the text corpus investigated. It is not *lege artis* possible, and scientifically not justifiable, to correlate single items or even syndromes of orthographic behavior with personality data of speakers or a particular speaker.

As indirect evidence, DOB may be of relevance in a context that cannot be seen at the start of the analysis. The forensic linguistic literature contains numerous examples of this. But a cumulative increase in knowledge can be achieved through analysis of orthographic data in many cases.

Linguistics, forensic linguistics in particular, will have to take into account the regularities, modalities, variations and dimensions of writing use from a much deeper and much more expansive frame of reference than it has to date. It would certainly be useful to have a much larger basis of text corpora and the practical forensic linguistic work administered to them – of real life TS and VS. The linguistic analysis of norms and the rather large area of theoretical and empirical problems related to them are of somewhat minor importance when compared to the forensic linguistic tasks characterized above.

8

Orthographic Data in Forensic Linguistic Authorship Analysis

The role and status of orthographic data (henceforth OD) in an anonymous authorship analysis is a much debated issue in a much debated field of forensic linguistics. Several methodological, theoretical and, even more so, practical-analytical questions have not been dealt with satisfactorily. Some have not been given adequate answers to date. Others have not even been stated properly as questions. Yet others have not received the general linguist's attention, even though the practitioners' concerns and interests have been articulated thoroughly (and vice versa, practitioners have not taken general linguistic data into account properly). Obviously, at times there is some kind of miscommunication going on between the two. This chapter elaborates on the methodological and theoretical status of OD by giving heuristic taxonomies of classes of OD from a systemic grammatical, sociolinguistic, and text linguistic perspective. The focus of the chapter is on the "diagnostic potential" (Kniffka 1996c) that may or may not be assessed for OD in the context of authorship analysis. *All* data (items, features, analyses) presented are taken from real life forensic cases, dating from 1974 to the present time, in which the author gave expert testimony for German courts and other authorities. The aim of the chapter is to illustrate the complex status of OD by introducing an extensional definition of the phenomena and some additional necessary distinctions, and by clarifying the position of OD in the context of argumentation in forensic linguistic authorship analysis.

Introduction

The first section of the chapter provides a brief overview of various types of orthographic data (OD) taken from authentic incriminating letters from German cases in which I have been asked to provide expert testimony

over the last 30 years. The aim is to show examples of the polyvalence of orthographic text products in a material, operational and methodological sense, along with the difficulties that they create. In many cases, it is not even clear *what is said*. With some data, one cannot even take the grammatical data given and the semiotic status of the utterance for granted. Depending on the orthographic analysis, an entirely different utterance and text in form and meaning would have to be investigated. In addition, orthographic features may relate or apply to several different levels of linguistic description, which may have considerable consequences for any linguistic analysis of these levels. Most of the cases referred to below have been mentioned or dealt with in other publications (Kniffka 1981; 1993b; 1996c; 2000a; 2000b; 2001; and the chapters in this volume). The reader may be referred to these for a more detailed discussion.

In the second section, a heuristic taxonomy or, preferably, a list of taxonomies for describing the potential “idiosyncrasy status” of OD is given. As in the previous section, the aim is not to give an exhaustive overview of classes of orthographic data and features from authentic forensic cases. Rather, the aim is to provide illustrations of a wide range of variation of orthographic features. The main hypothesis here is that OD are of critical importance and have to be described in terms of a continuum. This is a new perspective that has not been applied to OD to date. It seems that viewing OD as a continuum is of great advantage for producing an adequate description and explanation of OD in texts.

The basis for this continuum is described in the third section, where the methodological impact for the description of a potential “idiosyncrasy coefficient” is discussed. This seems to be one of the most essential prerequisites for an exact quantitative description of OD, which would help to validate their empirical strength for forensic linguistic anonymous authorship analysis (henceforth FLAA). This section also deals with some text type-specific features of extortion and blackmail versus defamation letters; on the textual, orthographic and layout levels.

The overall hypothesis of this chapter is that it is not only worthwhile, but also absolutely essential to take OD into account systematically and thoroughly when undertaking AA. It is not adequate to neglect OD. On the contrary, the cases all show empirical evidence for the hypothesis that OD have to be taken into consideration with great care and *cum grano salis* in AA.

The term “orthographic data” is used here as a cover term for all kinds of textual data given in graphematic representation; that is, written texts, which, by definition, follow one orthographic system or another. Also, the more general term “writing system” is used here.

The simplest explanation of why we have to deal with data of the orthographic system and the structural as well as factual ingredients pertaining to it is that some incriminating texts and comparison texts are given as data in written form. Therefore, forensic linguists are bound to deal with the data from an orthographic perspective as well, even if this is in no way sufficient to answer questions of textual authorship and in spite of the fact that OD have to be handled with the greatest care. More specifically, analysis of the OD is a prerequisite before trying to answer the first of the three famous questions to be dealt with in AA: (1) What is said?; (2) What is meant?; (3) Who is the author of the text?

A note on data representation and keys: original data of incriminating texts, comparison texts and other object language examples are given in *italics*. Original data occurring in a text are written in front of a slash (.../), the intended “correct” forms (in S.P. Corder’s 1971 terms: the “reconstructed” forms) are given after the slash (/...).

Example: *Geschäftführer/Geschäftsführer* means that in the original incriminating letter the (incorrect) form *Geschäftführer* occurs, which corresponds to a standard form *Geschäftsführer* in Germany.

Examples of orthographic data in anonymous letters

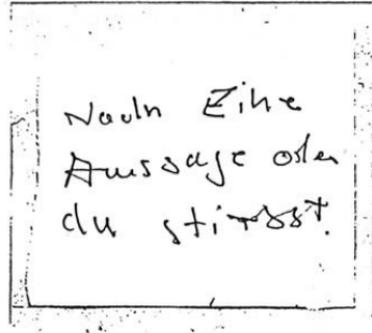
Case 1: An anonymous life-threatening letter

In the case of an anonymous threatening letter (cf. Figure 8.1 (a), (b) and (c)), the forensic linguistic expert’s job was to examine whether the letter was written by the German wife of an imprisoned Arab terrorist (who claimed that his wife had written and addressed the letter to herself), or whether it was written by a fellow Arab accomplice of his with whom he was planning to get rid of his wife. The police took this threatening letter very seriously due to various factors and reasons that will not be explained further here (for details, see Kniffka 1992).

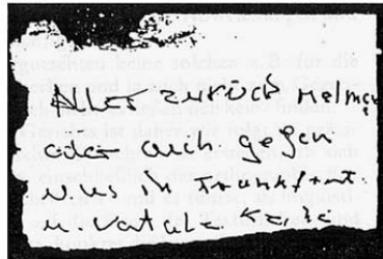
It is necessary to illustrate two classes of OD that should be analyzed before the question “What is said?” can even be approached. (1) In one line it says *Frankfurt u. vatale*, leaving it open as to whether this refers to the name of a town, a village in Germany or in an Arab country, or whether the German adjective *fatal* (in wrong orthographic representation) is meant. The latter poses a problem especially to non-native speakers and writers of German, due to the fact that in some words <v> is pronounced /f/, as in *Vater* (and in some special pronunciations, such as Bavarian *Vesper* /'fespe/), and in some words <v> is pronounced /v/, as in *Vase*.

(a) Front of envelope

'One more
evidence or/and
you [will] die.'

**(b) Back of envelope**

'Take everything
back or/and no
evidence against
us in Frankfurt
u. vatale (?)'

**(c) Inside of envelope**

'[evidence], or
you are dead.'

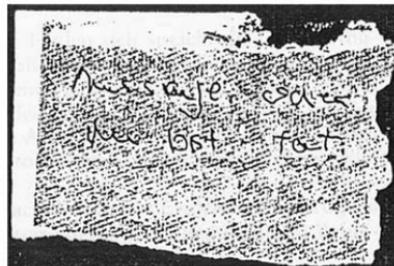


Figure 8.1 Anonymous threatening letter

An even more “potentially writer-indicating” feature is the comma in Figure 8.1(b), first line. It is curved to the left under the second letter <e> of the word *zurücknehmen* (similar to an opening parenthesis (... (...)). Curving a comma to the left rather than to the right (similar to a closing parenthesis (... ...)), as is customary with native German writers), frequently occurs with Arab writers writing German due to their normal direction of writing from right to left. This is not necessarily

a one-to-one correspondence, however, and certainly not a “writer-identifying” feature *per se*. It could also result from writing with the left hand rather than the (customary) right by writers who wish to disguise their writership (which also occurs not infrequently; cf. Hecker 2000; Michel 1982), by writing on a moving surface (as on a train), or by writers with some degree of intoxication. The important point is that this feature can be of specific evidentiary value in connection with other data.

As has been described elsewhere in more detail (see Kniffka 2003a; 2003b; and p. 187ff. in chapter 7 of this volume), there is rarely or never a one-to-one correspondence between particular OD of frozen text products and a writer’s personality as a whole, or even her/his individual and habitual characteristics. Rather, generalizations apply only to the totality of the writing behavior that can be reconstructed for a particular text (product) in a systematic analysis. This means that the writing of the comma curved to the left rather than to the right, as is customary in Latin script with a writing direction from left to right, certainly is a noteworthy graphematic feature that has to be interpreted in full detail by the field of handwriting analysis rather than linguistics. It does not necessarily mean that the author is an Arab, however, though many Arabs show this behavioral feature when writing in Latin script. The fact that most text products represented in written form are polyvalent as to their explanation applies here also.

As mentioned above, features and data of/in written texts can only be assessed as *indirect* evidence of text authorship. The writer/typist of a text does not necessarily indicate the author of a text. OD are, at any rate, only concomitant and, possibly, supportive data for questions of authorship. They may lead to significant results together with data of linguistic behavior.

An interesting question – which also appears to lack in-depth systematic research to date – concerns the systematic relationship of orthographic and linguistic behavior in terms of the range of individual variation. Although handwriting analysts have looked in great detail at the phenomenon of the variation(s) of writing that one and the same person can produce (cf. Hecker 2000; Michel 1982), there is little research about this range of variation compared to that in language behavior and language use by one and the same person. From a commonsense point of view, one would be tempted to assume that variation in handwriting is more restricted and more limited than variation in language use and ways of speaking. This may, however, be an assumption that only holds, somehow, for the totality of subjects and languages. It lacks empirical

validation with individuals to date. The question of whether one and the same person has a wider potential for variation in his/her language use than in his/her handwriting has not been researched and needs thorough empirical study.

Case 2: German “Fugen-s” in court “reading copies”

This case illustrates, from an entirely different perspective, the critical importance of OD for the collection and analysis of forensic linguistic data. From my own experience in expert testimony, it is a very important area of everyday routine handling of OD by the judiciary. The case in question illustrates two different instances of how the judiciary was inaccurate and negligent in handling textual and orthographic data relating to incriminating and comparison letters:

(1) Linguistic experts in many cases have no way to examine the authenticity of the data of the original incriminating letters. They have to work with the texts given to them by the people who request the expert opinion. They have to accept the data in the form that reaches their desks, taking their authenticity for granted. This holds, with an even stronger methodological impact, for comparison texts, which the linguistic expert also has to accept as authentic data stemming exclusively from the suspect. In many cases in which I have given expert testimony, artefacts, or even several layers of artefacts, were generated by the fact that the textual and orthographic data of the incriminating and/or the comparison texts were *not* authentic but were, in fact, created by others and involved various types of methodological blunders. The most spectacular instance in Germany in recent years involved the well-known “Hitler diaries”, in which several internationally reputed handwriting analysts gave expert testimony. This was, in itself, a perfectly appropriate procedure. The experts testified that the incriminating and the comparison texts had indeed been written by the same person – the forger (who had supplied the wrong comparison data).

(2) A second problem is of a similar type. It also has to do with the authenticity of written incriminating and/or comparison texts. Every linguistic expert probably has come across so-called “authentic” letters containing many kinds of hand-written (and sometimes even typed) additions by the police, the court, the prosecution, other investigating authorities and even by linguists previously employed with the case. After being reproduced some three or four times, it is practically impossible to discover the authentic version of such texts or even to notice that the data are not authentic.

Case 2 illustrates this in a rather unspectacular and typical way. This case involved a series of defamation letters, including anonymous legal charges brought against a large southern German company. Since some of the original anonymous letters and the comparison letters were barely legible, the court produced what it called “Leseabschriften” (‘reading copies’), which, as it turned out later (after quite a disturbing correspondence), did not represent the originals accurately. In the incriminating letters, some forms of the so-called “Fugen-s” in German (‘joining-s’, which is inserted between two constituents of a nominal compound, such as *Übergang* + *Zeit* = *Übergangszeit*) were misrepresented. In instances in which, according to standard Modern German, one would expect a “Fugen-s”, no such element occurred in the reading copy produced by the court clerks, whereas it did occur in the original incriminating letters correctly; for example, (reading copy:) *Geschäftführer*/(original:) *Geschäftsführer* (‘manager’). In other instances, the reverse constellation was found. The original anonymous letters (wrongly) did not show a “Fugen-s”, whereas the reading copy correctly did; for example (reading copy:) *Altersversorgung*/(original:) *Alterversorgung* (‘old-age pension’).

This case illustrates how everyday “slips of the pen” (see Fromkin 1980) can have considerable consequences for linguistic analysis. In other words, seemingly harmless casual misrepresentations of an original text, stemming from sources which should be absolutely free of such flaws, such as official transcripts produced by a court, may turn out not to be harmless at all. They may actually lead to serious methodological and factual errors in AA. In German, the “Fugen-s” is not just a spelling variant, but also has some special grammatical and sociolinguistic structural differences, as represented in Table 8.1 below. This is, at the same time, an illustration of the general postulate that every orthographic and linguistic analysis of incriminating and comparison texts has to provide a systematic grammatical analysis of the data as well.

Four basic classes of the “Fugen-s” can generally be distinguished, as shown in Table 8.1.

A deviation in the written representation (reading copy:) *Geschäftführer*/(original:) *Geschäftsführer* or (reading copy:) *Altersversorgung*/(original:) *Alterversorgung* can mean that, in addition to or instead of a “slip of the pen”, there was a grammatical mistake produced by the writer; namely that she/he does not know German grammar well enough to differentiate these forms (for further details, see Fleischer and Barz 1992; Holst 1978).

This is not the whole story, however. Matters become even more complicated by the fact that there is, in addition, a geographical variation of

Table 8.1 Classes and types of the “Fugen-s” in German (cf. Fleischer and Barz 1992; Holst 1978)

Class (1): Complementary (asymmetrical) distribution of “Fugen-s” (only one form variant occurs for each word):

<i>Rindfleisch</i>	-	<i>Rindsleder</i>
<i>Schiffbruch</i>	-	<i>Schiffsrumpf</i>
<i>hilflös</i>	-	<i>hilfsbedürftig</i>

Class (2): Real double forms (for many speakers in Germany):

<i>Fabrikgelände</i>	-	<i>Fabriksgelände</i>
<i>Schadenersatz</i>	-	<i>Schadensersatz</i>
<i>Gesangverein</i>	-	<i>Gesangsverein</i>

Class (3): Forms with “Fugen-s” only (in first constituents of a compound which are derived words):

Freiheitskampf
Menschheitsstraum
Lösungsmittel

Class (4): Forms with or without “Fugen-s” with a lexical-semantic difference (representing two different words synchronically):

<i>Landmann</i> ‘peasant’	-	<i>Landsmann</i> ‘fellow country man’
<i>Verbandkasten</i> ‘first aid box’	-	<i>Verbandsbeitrag</i> ‘association fee’
<i>Verbandpäckchen</i> ‘bandage set’	-	<i>Verbandskasse</i> ‘association’s takings’
<i>Schiffahrt</i> ‘shipping’	-	<i>Schiffsfahrt</i> ‘sea journey’

forms with or without “Fugen-s” in identical words in the three German-speaking countries, as represented in Table 8.2.

Table 8.2 represents the standard forms in use in the particular country. An interesting fact of this distribution is also that the regional alternatives are usually *not* well known or not known at all by speakers of the other region/country. Experiments showed that oversimplified alternatives

Table 8.2 "Fugen-s" in identical words in the three German-speaking countries

Switzerland	Germany	Austria
	<i>Zugführer</i>	<i>Zugsführer</i>
	<i>Gepäckaufgabe</i>	<i>Gepäcksaufgabe</i>
	<i>Gesangverein</i>	<i>Gesangsverein</i>
	<i>Fabrikmarke</i>	<i>Fabrikmarke</i>
	<i>Fabrikhalle</i>	<i>Fabrikshalle</i>
<i>Auslandmission</i>	<i>Auslandsmission</i>	
<i>Beileidtelegramm</i>	<i>Beileidstelegramm</i>	
<i>Sportsmeldung</i>	<i>Sportmeldung</i>	
<i>Zugsverbindung</i>	<i>Zugverbindung</i>	

used in countries Y and Z are rated ungrammatical or non-existent, and are heavily rejected by the speakers of country X, and vice versa. For example, speakers from Germany insist that *Zugführer* is the *only* grammatical and acceptable variant, whereas speakers from Austria are equally vehement that it is *Zugsführer*. The Swiss insist that *Beileidtelegramm* is the form of the word for a telegram of condolence, whereas it must be *Beileidstelegramm* in Germany.

Although the "incorrect" use of a "Fugen-s" in a text could be a linguistic and/or orthographic indicator of the home country of the author and/or writer, no such regional explanation is possible in the case discussed above. The deviations found are clearly just "slips of the pen" originating from the slovenly production of a reading copy by a court clerk.

It is clear that the data of the grammatical and the sociolinguistic variations of the "Fugen-s" in German cannot simply be taken for granted as described in grammars of German, no matter how reliable they and the sources (such as court transcripts) are. The forms found have to be tested against actual usage with speakers of the three standard varieties of German. Such an empirical investigation is necessary in order to determine whether the data found in grammars are reliable and valid after all. In the case in question, inquiries were made with speakers of German in three university towns in Germany, Austria and Switzerland. The results showed that there are some additional variations for some of the words listed but that, in general, the picture described above is correct. In my own north German "regiolect" I (like many other speakers from that area) make no difference between *Verbandkasten* ('first aid box') and *Verbandskasten* ('first aid box'), which can also occur with a "Fugen-s".

The variation of forms of this type, all with a rather low degree of regularity and normative strength, has to be determined thoroughly if these data are to be used in forensic cases. The example given is also an illustration of the general postulate that linguistic evidence of this nature should be restricted to the “Ermittlungsverfahren” (‘investigation section’) rather than to the “Hauptverhandlung” (‘trial section’) in German court proceedings in cases of AA. In other words, nobody should be convicted or acquitted on the basis of such linguistic evidence alone.

Together with data from neighboring sciences (handwriting analysis, forensic phonetics) and other investigative forensic sciences and methods used by criminalists, this type of linguistic evidence can be very helpful.

Case 3: Analysis of an extortion letter without comparison data

Case 3 and Case 4 are also discussed here in exemplaric fashion primarily for their methodological impact, each case illustrating a different problem.

Case 3 is, in many ways, the most exceptional anonymous extortion/blackmail letter among hundreds that I have come across to date. It was advertised in a newspaper in the form of a demand to buy an item of real estate near Frankfurt a.M., Germany, for some 11.65 million DM, of which ten million was to be the real estate agent’s commission (the anonymous author/writer). The anonymous extortion letter, represented in Figure 8.2, is a stencil-written letter of almost one A4 page (35 lines). Unusual also in its length (see pp. 225–33), a more detailed analysis is given in Kniffka (2000b).

The higher police authorities, who had requested the expert opinion, were very concerned at the time, since there had been anonymous extortion and threatening letters by the Southeast European Mafia, then active in the area. In fact, the letter makes reference to the Bosnian and the Chechnyan wars, mentioning that the authors are veterans of these wars. The letters were mailed from three different towns in Hungary. The police wanted the linguistic expert to determine whether the author and/or writer of (the three different, but very similar variants of) the extortion letter was a native German, a native Bosnian, a native Russian or a native Hungarian speaker and/or writer.

The most exceptional characteristic of this case is a methodological one. There were *no* comparison data from any suspects, which made the “selection” of one within the theoretical possibilities of natives of

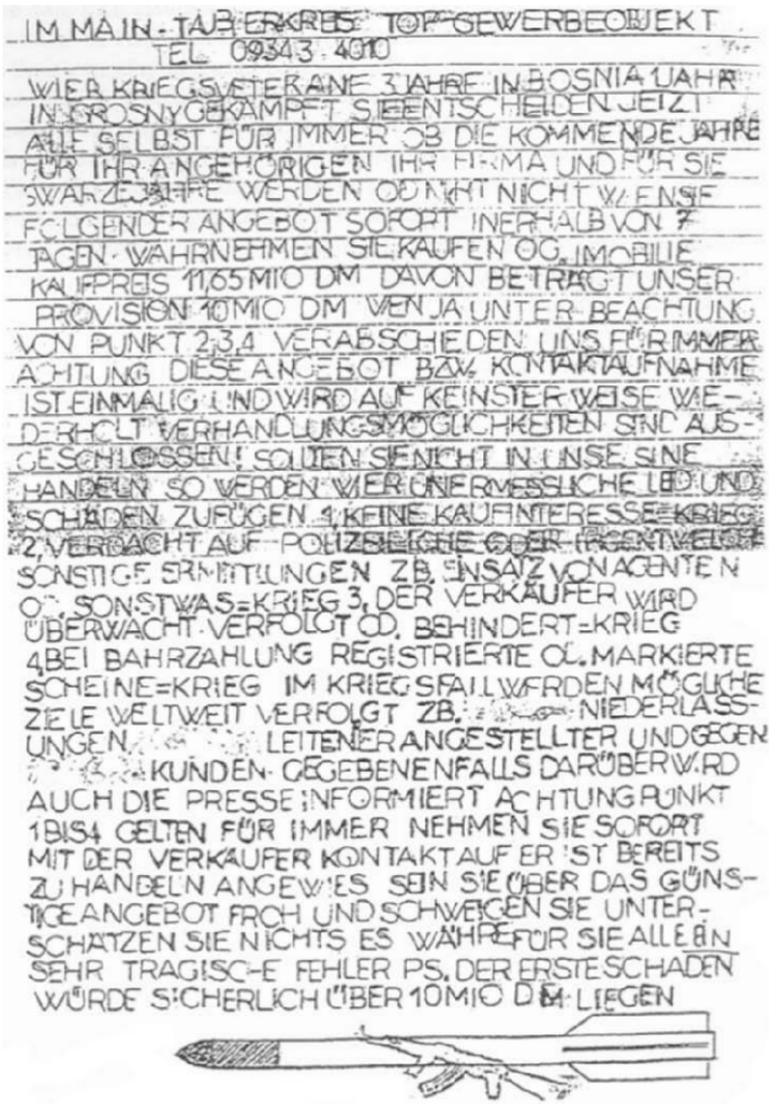


Figure 8.2 Example of an exceptional anonymous extortion letter

different languages even more difficult. In addition to grammatical and/or stylistic characteristics (including deviating features, errors and mistakes), orthographic and layout features were the only data available. It is not exceptional, of course, that there were no comparison data and no known suspects available. This is actually a quite frequent constellation of data. What is exceptional here is the nature of the data of the

anonymous letter, the overall investigation context and the specific questions addressed to the linguistic expert (see below).

This case was a strong indication of the importance of OD as clues for the “nativeness” of the writer. The analysis of the OD of the incriminating letter is definitely not a clue, let alone a guarantee, for determining who the *author* of the text is. It concerns only the anonymous *writer* of the text, more exactly an analysis of which native traditions of writing and spelling can be assumed. This, in turn, suggests hypotheses concerning the “writership”.

As basic alternatives, a native German, Bosnian, Russian and Hungarian were possible as the anonymous writer. In addition, it was necessary to examine whether some kind of disguise had been used; for example, whether the text had been written by a native German posing as a Hungarian writer or vice versa, or whether someone was trying to create the impression that she/he was a Serbian, Bosnian or Russian writer.

The technical and physical aspects of the (hand-)written text products were analyzed by handwriting experts. The forensic linguist’s job was to analyze the orthographic (and grammatical and textual) data. Since these analyses represent semiotic dimensions varying, in principle, independently of one another, they allow a systematic description of co-occurrences of (bundles of) features of each analysis. If patterns of feature configuration on a larger scale can be discovered, the analysis is much stronger in terms of reliability and validity.

The handwriting experts’ analysis of the data proved inconclusive, except for the fact that the analysis strongly suggested that the letter as a whole was written by one and the same person. This was confirmed by the linguistic analysis of the text, which showed that there was a consistent and symmetrical distribution of orthographic errors and deviations across the whole text. The initial, the middle and the final parts did not display any special differences and mistakes. On the contrary, they showed exactly the same structure and frequency proportion of errors and mistakes of the same kind.

The detail of the analysis and the interpretation is available elsewhere (see Kniffka 2000b: 187–98). Summarized, the entire pattern of deviations, spelling mistakes and errors, suggested that a native speaker/writer of German was very unlikely to be the anonymous writer. Some mistakes, which any third-grade student at a German elementary school would not produce with such frequency, were (consistently) made; for example, *wen/wenn* (‘if’, ‘when’), *Imobilie/Immobilie* (‘real estate’). Other double consonant writings (with letters other than *-mm-* and *-nn-*) were spelled correctly in the anonymous text.

This suggested either a non-native writer of German making elementary spelling errors or a native German posing as a non-native, intentionally disguising (downgrading) his command of German spelling. On the other hand, the text shows several examples of “more difficult” spellings in German (for example, words with an umlaut) that are generally difficult for non-natives of German to perform. The numerous examples of correct spellings of umlaut forms might suggest a native German or a non-native writer in whose writing system umlaut is a common and widespread phenomenon. From this evidence, a native writer of Croatian, Serbian, Russian, Czech and other Slavic languages could be excluded. There is no umlaut writing in these languages (Slovak being the only exception with the letter <ä>). Natives of these languages have considerable difficulties with umlaut spellings in German and other languages. In Hungarian, on the other hand, umlaut letters are widespread. Hungarian has the short vowels <ö>, <ü>, represented by two dots (“Pünktchen”) above the letter, and the long vowels <ő>, <ű>, written by two strokes (“Strichelchen”) above the letter. Hungarian does not have a letter <ä> with dots or strokes. An equivalent to the German long vowel written <ä> does not exist in Hungarian (except for a few borrowings) and the equivalent to the German short vowel written <ä> is spelled <e>. Correct spelling of all umlaut forms makes it unlikely that Russian, Serbian or Croatian was the native spelling system of the anonymous writer. Hungarian or German would be much more likely.

In addition, Serbian and Croatian do not have double consonant writing at all, so many more errors than the ones found in the text might be expected. In Hungarian, double-consonant writing has a different status (that is, from German). It marks the length of the consonant. Graphemic representation of double-consonants is generally C + y, so <ny>, <ty>, <gy> corresponds to German <nn>, <tt>, <gg> in writing.

There are several other examples of a significant tendency in this regard. The writer of the anonymous text obviously has problems with the spelling of diphthongs (Hungarian does not have diphthongs).

These data taken together and, most of all, the fact that there are absolutely no errors in the text in the representation of <ö> and <ü>, but several in that of <ä> – which is not indicative of the writing behavior of a native writer of German but is very easily explainable as being that of a native writer (and speaker) of Hungarian – suggest that the writer of the anonymous text was most likely a native Hungarian, with a rather good command of German.

This was confirmed by the analysis of the data of grammatical, textual and stylistic behavior, which also suggested a native speaker of

Hungarian with a rather long and extensive exposure to and knowledge of German. On this basis, the linguistic expert's opinion concluded that, as the most likely option, a native Hungarian with a good command of German wrote the anonymous extortion letter.

A local native Hungarian who, for other reasons, had been in contact with the police was taken into custody, and was confronted with the linguistic expert's opinion. When charged with the authorship and writership of the letter, it was pointed out to him that this was the result of an analysis by a forensic linguistic expert. He then wanted to read the conclusion of the expert opinion (yet another very unusual feature of the case). After he read it carefully, he confessed that he indeed had written the anonymous extortion letter.

This case illustrates the importance of OD for its own sake and in connection with grammatical data. The analysis of the covariation patterns of the orthographic and the grammatical data allowed extensive hypotheses concerning the writership and, in turn, the authorship of the anonymous text. Natives of several other languages could be excluded as writers and authors, even without comparison data from particular suspects.

The general reference and *tertium comparationis* in this case were not comparison data of texts written by known persons, but more general implications of the orthographic system and the language system working in tandem. In such cases of broad or large-scale alternatives, it is possible to analyze anonymous authorship even when no specific comparison data are available.

Case 4: Misspelling in incriminating letters

The fourth case illustrates the operational and methodological impact that data of orthographic behavior may have in FLAA. In the case in question, an enormous corpus of anonymous letters (in total, more than one hundred typed A4 pages) and comparison texts of about the same quantity existed, supposedly produced by a suspect who was already in custody at the time. Orthographic features turned out to be the key features in the "discovery procedure" that the successful AA was based upon. The case is described in more detail elsewhere (cf. Kniffka 1990d: 438ff.). The linguistic analysis turned out to exonerate a suspect who was already in custody and was released from jail after the linguistic expert opinion was presented to the court. As mentioned above, the case is methodologically of particular relevance, since it illustrates that features of OD may be the key features for alternative hypotheses in AA.

To summarize Case 4, a large number of threatening and defamation letters to a company were assumed to have been written by an employee

who had been convicted of writing anonymous defamation letters to another company (a previous employer) some 15 years previously. In addition, he had a psychiatric record. All of this made him sufficiently suspicious to be charged as the anonymous author. The police put him in “Untersuchungshaft” (‘detention pending trial’). The 57 year-old suspect had been living in the city of Cologne his entire life, did not learn any foreign languages either at school or later, and had vocational school training without baccalaureate. He had never visited any other foreign countries except France and the Netherlands, each for only a few days. In addition to his personal record, he probably was arrested because the anonymous threatening and defamation letters happened to refer to insider details of the section in which he was employed.

In spite of an intense effort by a team of linguists for three months, it was not possible to come up with any kind of working hypothesis that would give clues to the authorship.

Shortly before arriving at a *non-liquet* as conclusion of the AA, yet another examination of all data, including OD, was made. In this examination, the attention of the linguistic experts was caught by the (frequently occurring) writing of *Die 3-Affen* (‘the 3-apes’) and also *Die drei-Affen* (‘the three-apes’), cf. Table 8.3a. (The expression refers to an emblem well known to people in Germany for ‘nothing heard’, ‘nothing seen’ and ‘nothing said’, with three apes covering their ears, eyes and mouth with their hands respectively). The very fact that this expression was hyphenated was the key feature that triggered a substantial working hypothesis that a non-native speaker was the potential writer/typist of the anonymous letters. A native citizen of Cologne who had no knowledge of foreign languages and had never been outside his home country would rarely be tempted to use a hyphen in such a context, not in the version with numbers, let alone that with the word *three*. Not so in the US, for example, where it is not at all unusual to write expressions such as *the 3-apes*, *a three-way procedure* or other expressions containing numbers with a hyphen. This cultural knowledge triggered a hypothesis that the anonymous writer may be a non-native speaker of German with a near-native command of the German language.

With this working hypothesis in mind, other spelling data consistent with this hypothesis, together with a considerable amount of grammatical data suggesting non-nativeness of the author of the incriminating letters, were found. Spellings such as those given in Table 8.3b (*Äüßerung*, *Außerung*, *Aüßerung*) are usually not made by native speakers/writers of German. Quite frequently, however, they are misspellings made by

Table 8.3a Spelling of a well-known emblem in German

<i>die 3-Affen</i> 'the 3-apes'	(11 times in incriminating letters)
<i>die drei-Affen</i> 'the three-apes'	(3 times in incriminating letters)

Table 8.3b Concomitant exceptional spellings

<i>Äüßerung, Außerung, Aüßerung</i> instead of (correct) <i>Äußerung</i> 'utterance'	
<i>Pfüi Teufel</i> 'boo'	instead of (correct) <i>Pfui Teufel</i>

non-native speakers. From a phonological and graphemic point of view, these misspellings show that it is a bit difficult for non-natives. The diphthong [ɔʏ] has three graphemic representations in German, <äu> as in *Äußerung* ('utterance'), <eu> as in *heute* ('today') and <oi> as in *Ahoi* ('ship ahoy'). Also, the misspellings represented in Table 8.3b cannot be explained by a confusion of letters on the keyboard: <u> and <ü> are not placed adjacent to each other. The frequency of occurrence, the types of deviations, and also the occurrence of the word *Pfui* misspelled as *Pfüi* seem to be relatively rare with native German writers and suggest that the writer/typist of the anonymous text very likely was not a native German.

With these results for the writership of the anonymous letters, a thorough linguistic analysis of the grammatical, textual and pragmatic behavior was undertaken. It turned out that there were many semanto-syntactical expressions that could be explained as anglicisms, such as *Diese Maßnahmen werden mit weiteren Techniken begegnet*, obviously influenced by *These measures are met by ...* Furthermore, there were also non-idiomatic quotations of German proverbs. A native German will, very likely, quote a well-known German proverb in the form *Eine Krähe hackt einer anderen kein Auge aus* ('A crow does not peck another's eye.') (which may have a standard variant *Eine Krähe hackt keiner anderen ein Auge aus* ('No crow pecks another's eye.') in some regions). In the incriminating letters, however, the non-idiomatic version *Eine Krähe hackt einer anderen nicht in die Augen* occurs. Also, some errors representing non-idiomatic German were found. The expression *Auf Geschäftsunkosten nach Düsseldorf fahren* is quite logical, but unidiomatic. Idiomatic German would be *Auf Geschäftskosten nach Düsseldorf fahren* ('to go to Düsseldorf on company expenses').

The entire analysis of the data was repeated with this hypothesis in mind, and the amount of Anglicisms found was so high that it was concluded, with a high probability, that the writer and author was a non-native speaker and writer of German. The criminal investigation department of the police went through the records of all employees in that section of the company. Only two non-native speakers of German were found, a Frenchman and a German-American, both of whom spoke German very well. After re-examining all data in this case with French in mind, it was felt that there was a stronger hypothesis suggesting that the German-American could be the anonymous writer and author. He was, in turn, monitored by the police and was caught writing yet another anonymous letter.

Case 4 also shows that forensic linguistic AA is not only capable of insinuating and creating suspects, but also of clearing people of criminal charges and exonerating them, which is also an important intrinsic characteristic of the field.

The four cases briefly described above show that OD, in a broad sense, can be of importance – sometimes of critical importance – in forensic linguistic AA. In fact, if the sociolinguistic principle of “cumulative evidence” is taken seriously, one cannot but administer (in each case, involving written text material) a systematic analysis of textual, linguistic *and* orthographic data. In all four cases, a “solution” – or even a reasonable working hypothesis – would not have been possible without analyzing OD. The only conclusion one can draw from this is that OD, although never of any direct evidence for questions of textual authorship, may be of very important indirect methodological relevance for AA.

In the remainder of the chapter the nature and status of OD is described in a more systematic way. The description is, necessarily, also somewhat brief and exemplaric but, nevertheless, it offers a theoretical dimension for explaining the data described inductively on the basis of the four cases above.

Continua of orthographic behavior and the methodological impact for a “idiosyncrasy coefficient”

This section has two main concerns:

(1) OD rarely have a plain dichotomic structure. Rather, they represent *continua* of orthographic behavior ranging from “right” to “wrong”, that

is, from more or less right to more or less wrong, from customary to less customary, from standard to less standard;

(2) There are text type-specific rules of orthographic behavior that represent standard or non-standard forms of orthographic behavior and which also apply to all sorts and types of incriminating texts. On this basis, it is possible to derive methodologically salient criteria for the analysis of incriminating letters.

It is also argued in this section that the continua described have a hierarchical structure in terms of a lower or higher “idiosyncrasy coefficient”, which one should eventually be able to describe in statistical quantitative terms. This undoubtedly offers a great advantage in describing and evaluating forms of orthographic behavior as types and tokens.

Basic theoretical and methodological assumptions

In the first section of this chapter, it was shown that OD and features of orthographic behavior are of relevance and, sometimes, critical importance for AA. The evidence supplied by authentic data from four different cases should be sufficient to support this. If this is so, what do orthographic data and features look like? What is their structure? What is their evidentiary significance? What are the major methodological implications for AA resulting from them?

(1) The most salient observation in dealing with OD and features of orthographic behavior is that there are considerable differences in the nature of the data themselves and in the ways they can be used as evidence in AA. The predominant overall characteristic is that the differences can best be described in a heuristic taxonomy of a continuum, or more than one linear continuum of orthographic features. In addition, one has to describe both orthographic features and structure on the one hand, and usage and frequency of occurrence on the other. And, most importantly, orthographic features, as all linguistic features, have to be described as text type-specific entities. Both facts – first, that OD are to be described as continua and, second, as text type-specific regularities, have not been accounted for sufficiently to date in forensic linguistics or, for that matter, in linguistics altogether. A preliminary account of the empirical and methodological implications for different types of incriminating letters is given elsewhere (cf. Kniffka 2003a). Some theoretical and methodological aspects have also been described elsewhere (cf. Kniffka 2003b; and chapter 7 in this volume). In this

chapter, only the gist of the central concepts and methodology are discussed.

The basic characteristics of OD and the main reasons for describing them as continua are that OD show considerable differences in their occurrence in quantitative terms and differ at the same time (though not in a one-to-one relation) in the diagnostic potential they have or may have for AA. This means, in simple practical terms, that some orthographic features occur by the thousands and millions in everyday written communication, and are used daily by millions of interlocutors in their written texts. In the realm of punctuation, for example, this would hold for such things as the (non-)use of a comma, the use of a full-stop instead of an exclamation mark or vice versa at the end of a sentence, the use of dashes, the non-use of space between letters. Other OD are rare as types or as tokens or as both; for example, the continuous and categorical omission of a full-stop or any other punctuation mark at the end of a sentence in an entire text.

A rare orthographic variant (as type and/or token) may be useful for identifying a particular writing usage and, potentially, a particular writer/typist, but this is not always the case. A relatively frequently occurring orthographic variant on the other hand may, under certain circumstances and in certain linguistic and situational contexts, be equally useful for identifying a particular writing behavior and, potentially, the writership of an anonymous text. This means that forensic linguists have to internalize the rather unpleasant and unpopular fact that there is no direct one-to-one correspondence between any orthographic or linguistic variant and its methodological and evidentiary relevance in forensic linguistic writership and authorship analysis. Methodological impact and evidentiary significance of an orthographic variant have to be determined in a *procedere sui generis*; that is, in a rather encompassing and thorough linguistic analysis of the systematic and factual “diagnostic potential” (see Kniffka 1993b).

(2) In addition to the fact that OD are not simple dichotomies but are also rather complex entries on a continuum with gradual differences, there is another even more challenging methodological complication. Orthographic variants and orthographic features of the structure and the occurrence of these variants do not have one and the same value once and for all. There is no general opposition of widespread and frequent occurrence versus uniqueness and exceptionality. This depends on text type-specific constraints and other text linguistic regularities, and only on this level are generalizations possible. For example, abundant and redundant (simultaneous) use of punctuation marks (such as

exclamation and question marks) might possibly work well in an ad for a rock concert, but not at all or not quite as well in a newspaper obituary. If such a form should occur against the general standard and expectation of a particular text type, the exceptionality or “uniqueness status” would represent a marked occurrence.

(3) From this follows yet another criterion for which one cannot base a forensic linguistic analysis on the assumption of a one-to-one correspondence between a particular orthographic (static) variant or orthographic feature and a certain methodological or operational consequence. In other words, a particular orthographic variant cannot be said to have a specific diagnostic potential that can be seen as an indicator attached to this particular orthographic variant. Nor can a particular orthographic variant be said to have a particular diagnostic potential once and for all, independent of text type-specific constraints. Instead, it is necessary to determine a *dynamic* diagnostic potential in a rather complex analysis in which the structural entities of the type and the distribution in certain text types and situations are determined. This diagnostic potential has to be reconstructed in an analysis comparable to that used in error analysis a few decades ago (cf. S.P. Corder 1971), including the etiology of errors in addition to the description.

(4) In an even more complex analytical procedure, the diagnostic potential that can be provided by the analysis of orthographic features can be combined in a holistic perspective with a description of a dynamic “Ganzheit”, indicated and stated by the “idiosyncrasy coefficient” (ISC, cf. the third section of this chapter).

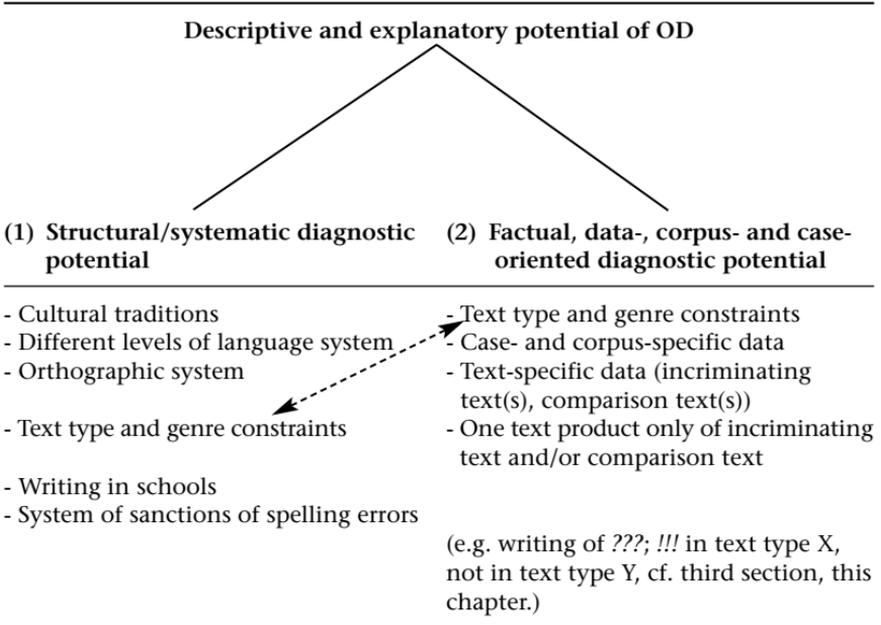
The data of orthographic behavior can be combined with data of linguistic behavior, together leading to a holistic interpretation of the ISC of verbal and orthographic behavior. Since both of them, in principle, vary independently of each other, an analysis of the entire feature configuration may allow some likelihood assumptions about a potential individual speaker’s behavior.

The diagnostic potential of orthographic data

The diagnostic potential of OD, which is derived from the descriptive and explanatory potential of OD in a particular text, can be described in a simplified manner, as represented in Table 8.4 (a more detailed description is given in Kniffka 1996b; 2003a).

In essence, the distinction of (1) structural or systematic diagnostic potentials; and (2) factual, data-, corpus-, and case-oriented diagnostic

Table 8.4 Descriptive and explanatory potential of orthographic data



potentials can account for the fact that a forensic linguistic analysis of orthographic (and, similarly, linguistic) data listed in (1) refers to data that can be achieved by an analysis of the various systems and subsystems that a language, a text in a language, and a special text type in a language, in a certain sociocultural situation, provide. They are superordinate to (2), the particular sets of data that are given in a special corpus, a particular text of a corpus, or a particular item of one text product in a larger text or collection of texts.

The distinction between (1) and (2) is based on the practical needs of methodology. The factual diagnostic potential envisages the case-oriented data; that is, regularities that can be generalized from the particular text or text corpus, including generalizations that can be made for the text type of, for example, extortion and blackmail letters as opposed to other incriminating letters, and as opposed to non-incriminating letters that otherwise would fall into the same text type. In addition, the factual diagnostic potential would cover all case-specific constraints, regularities and characteristics that can be defined for a particular corpus of texts, such as incriminating texts only. Such a regularity or feature of orthographic behavior may apply to one single item of a text, an incriminating text or a comparison text alone.

To both diagnostic potentials, the hypothesis applies that a much greater range of linguistic and orthographic variables and regularities, constraints and specifications has to be taken into account in a more consistent, thorough and expansive approach than has been done to date in forensic linguistics. The realms mentioned in Table 8.4, although not exhaustive, can be seen as areas that have attracted only limited attention in the past. They represent features that occur in several real life cases and are of particular importance for the “solution” of the forensic linguistic cases. For example, the data of “cultural background knowledge” of linguistic and writing behavior, the use of particular “copy book styles”, as identified by handwriting analysts, in schools, in a particular culture and others should be described systematically.

The dotted line in Table 8.4 between text types in the structural and in the factual diagnostic potentials illustrates that rules, constraints and specifications of text types and genres have to be described on each level and in every direction. It is not only necessary to understand how a particular extortion letter is to be judged in (absolute) terms of “commonness” and “uniqueness”, but also that the general rules applying to personal letters of non-criminal denomination and criminal denomination each have to be described systematically. This includes the structural similarities and differences between extortion letters and other incriminating letters; for example, defamation and libel letters, and the “exceptionality” or “professional” status and standard of extortion letters. This is not trivial. It can be of great importance in differentiating a “typical” from an “atypical”, a “serious” from a “non-serious” (mimicry), a “professional” from a “semi-professional” writing behavior in extortion letters.

A forensic linguistic analysis of both the structural diagnostic potential and the factual diagnostic potential seems to be a prerequisite for an adequate qualitative and quantitative analysis of the ISC of a particular text and, eventually, of the idiolectal status of the text in question. This can be of great criminalistic importance in narrowing down the number of suspects as authors and writers.

Continua of orthographic features

Concerning the continua of orthographic features, it should be pointed out again that not only the material characteristics of a particular variant of linguistic and orthographic behavior matter, or are the most interesting and methodologically most rewarding forensic linguistic “facts” of a text. Of higher evidentiary significance for forensic linguistic analysis are the conditions, constraints and regularities of the *distribution* of features. In contrast with the aforementioned, the latter have been accounted for far

too little in forensic linguistic research. The idea that a text of a criminal can be “identified” by forms of a very “peculiar” or, eventually, “identifying” nature is, at best, wishful thinking. This is a false interpretation of basic data of language and language behavior, as well as of the specific data constellation in an anonymous letter. Only in special genres of poetic text types is one likely to find a relatively large amount of “highly specific” language data, such as neologisms and unique words.

Everyday language probably would not function the way it does if the degree of its neologisms, “exotic” forms and meanings were very high. Incriminating letters, written by criminals with a criminal purpose (such as extortion or blackmail) do not, as a rule, show specific forms or peculiarities that could be used as operational indicators that identify a particular person’s way of writing. In my experience in analyzing a great many extortion and blackmail letters, such language items, ways of expression and orthographic features are, in fact, extremely rare. It is a common misconception of lay people to assume that criminal authors and writers can be identified by any material peculiarities of this sort in their texts.

Rather, the performance aspect (ways of occurrence of a particular form, a system or a subsystem of forms) may be more indicative of a particular authorship and/or writership. For OD this may be relevant in different ways for different dimensions and kinds of orthographic variation. Table 8.5 lists different (types of) spelling variants of a word in use in Germany today. It refers, as a kind of an orthographic shortcut, to a male and a female person designated by one and the same word/expression.

As a result of feminist and gender study activities, a couple of “new” forms (1)–(5) have gained ground in German spelling, in addition to (6) and (7), which were the only forms for a long time.

Such a variation in texts is not highly marked or indicative of a writer – generally speaking. The variants (1)–(5) in Table 8.5 may be said to have a rather low ISC. However, the occurrence of these variants may be of high significance in a special linguistic and extra-linguistic context and with a particular author. An 80-year-old man with little exposure to

Table 8.5 Different spelling variants of a stereotyped expression

-
- (1) MitarbeiterInnen ‘co-workers’ (male and female)
 - (2) Mitarbeiter/innen
 - (3) Mitarbeiter/Innen
 - (4) Mitarbeiter/-innen
 - (5) Mitarbeiter(innen)
 - (6) Mitarbeiterinnen und Mitarbeiter
 - (7) Mitarbeiter und Mitarbeiterinnen
-

written language would not be expected to use any of the first five variants at all, so if he chose (1), (2), (3), (4) or (5), rather than (6) or (7), this could be seen as an indicator identifying his authorship. Conversely, a feminist activist who would *not* use one of the variants (1)–(5) but only the form *Mitarbeiter* for males and females (which would be a feature of linguistic and orthographic behavior) and would not use the variant most en vogue with feminists, might be expected to use a highly marked feature of orthographic behavior.

The data in Table 8.6 are yet another indication of the fact that values cannot be assigned once and for all to orthographic features, no matter of what type. The sociolinguistic ingredients of the structure and use of these forms in particular given contexts that can best be described in an ethnographic frame of reference have to be taken into consideration (cf. Kniffka 2003b; and chapter 7 in this volume). In the forensic linguistic analysis of orthographic behavior in authentic cases, this is much more important (and rewarding) than the search for highly idiosyncratic grammatical and spelling variants. An “excessive” use of punctuation marks (or any other special mark) is one of the areas that has been much neglected but which is of potentially great importance in analyzing writership of an anonymous letter. Excessive hyphenation, excessive use of quotation marks in one language or in different languages, excessive use of abbreviations and other punctuation items may be of great methodological and operational significance. If one can verify or exclude a similar punctuation pattern for the comparison data, the likelihood of identity of the writership in the incriminating and the comparison texts is rather high or low respectively. It goes without saying that, in this case, statistical measurements and methods of quantitative analysis have to be addressed. In short, whenever a quantitative analysis of the data of the incriminating and comparison texts can be done, it should be.

It should be noted in this context that very rarely indeed (at least, in certain text types) do incriminating letters allow for a statistical analysis, since they are usually very short. In my opinion, the general device in FLAA has to be: “Whatever can be quantified and measured should be”. It should be kept in mind, however, that not everything can be quantified.

What is “excessive” and what is not is determined by text type-specific constraints and subject to individual language variation. It is definitely possible to calculate intra- and interpersonal invariant syndromes and attitudes of orthographic behavior in this respect. The usage of quotation marks in written texts of everyday communication seems to be relatively higher in German than, for example, American academic writers. This can be confirmed by empirical quantitative studies. There

Table 8.6 Spacing with parenthesis dashes ± empty space (ES) before and after parenthesis dashes

- Occurring many times in incriminating anonymous texts exclusively in one constellation		
(1) with ES before and after dash:	x – y	x – y
(2) without ES before and with ES after dash:	x– y	x– y
(3) with ES before and without ES after dash:	x –y	x –y
(4) without ES before and after dash:	x–y	x–y
- Combinations of (1)–(4) in asymmetrical distribution in a text		
- “One and the same” punctuation mark, but various heterogeneous forms/versions and distribution patterns in a text		

Key: “x” stands for the last word/letter *before*, “y” for the first word/letter *after* a parenthesis dash.

Dimensions of variation:

- Language-specific range of variation
 - Culture-specific range of variation
 - Writing/speaking tradition(s)
 - “Schulvorlage” (‘School version/norm/standard’; copy book (style))
 - Competing (new) norms
-

are many other language-specific, culture-specific, group-specific and also idiolectal characteristics of the use of punctuation marks. As an example of what a quantitative analysis in principle could analyze, Table 8.6 gives an outline of the use of parenthesis dashes with or without empty spaces before and after in a text. The advantage of these data is that they can be analyzed and measured in a quantitative analysis quite easily. The diagnostic problem is that, in many real life cases, the amount of data of incriminating and comparison texts available is too small to undertake statistical analysis.

Empirical analysis of (linguistic and) orthographic “idiosyncrasy coefficients” (ISCs)

It is not the material characteristics of linguistic and orthographic data that are, in one way or another, of greatest relevance for AA. In laymen’s terms, the general maxim “keep looking for rare (enough) and ‘unique’ forms in an incriminating text and you will find clues as to who the author/writer is” has no real scientific bearing as far as I can see. First, there are very few “unique” forms in our everyday communication

system provided by language. Secondly, criminals do not really use them that much, knowingly or unknowingly. Perhaps if there were more unique features than there actually are, criminals would recognize this and try hard not to use them. Be this as it may, this hope for unique data would not work methodologically. One would have to apply this maxim both to the incriminating texts and to the comparison data. Only if such rare or unique forms occurred in large numbers in the incriminating and in the comparison texts of one single author and writer, they could, possibly, have some methodological impact. Such a constellation of data is very rare, however. It is safe to say that, in real life forensic linguistic AA, such data do not play any important role.

Below, reference is made to orthographic variants with tentatively lower and tentatively higher ISC. The most important empirical facts that have to be accounted for are that: (1) it is not the substance but rather the dynamic pattern and distribution of linguistic and orthographic forms that matter in an empirical analysis; (2) there is significant variation and considerable difference between various (linguistic and) orthographic features. Even when not all these differences can be stated in exact quantitative terms, the very fact that such significant differences exist is hard empirical data.

In addition, most forensic linguists would probably agree that a demanding statistical analysis would have to be made in order to describe and calculate the facts of linguistic and orthographic behavior in terms of exact data. There is no question that this will be a requisite of forensic linguistic analysis in the future. In recent years, there have been interesting and promising attempts along these lines (in particular by the studies of Carole Chaski (2001) and Tim Grant (2001) – both of which indicate that it is generally sensible to undertake quantitative computerized analyses of large amounts of data when analyzing authorship attribution). At the same time, it should be noted that we still have a long road ahead of us to apply and adjust these analyses to the real life data of forensic cases. On the other hand, it also should be mentioned that not everything that is of critical importance in AA is related to questions of statistical quantification and, certainly, it cannot always be “reduced” to measurement. Nevertheless, Chaski (2001) and Grant (2001) appear to be leading us in the right direction. Obviously, corpus linguistic data will have to be consulted in a much more intensive way than they have been to date.

To apply quantification to the example of spacing parenthesis dashes described in Table 8.6, one faces an enormous statistical job. The exact

amount of variants of spacing or non-spacing needs to be described in relation to:

- (a) one particular incriminating threatening letter, for several incriminating threatening letters of one case, from a larger set of cases and, perhaps, even for all incriminating threatening letters available to the linguistic expert at a particular time;
- (b) the entire genre of threatening letters of various authors available in quantitative form to the linguistic experts' community in- and outside academia;
- (c) similar occurrences of spacing, i.e. [+ spacing] and [- spacing] before and after punctuation marks other than parenthesis dashes; such as before a question mark, an exclamation mark, a comma, a full stop.

These examples show that at least two descriptions have to be given here: one for parenthesis dashes and their occurrences, and one for spacing or non-spacing in relation to parenthesis dashes and other punctuation marks. The orthographic features mentioned here may seem to be of a somewhat "trivial" nature, at first sight. It should be pointed out, however, that the analysis of various punctuation marks, including their occurrence in a text, has been of prime importance in the analysis of forensic cases for which I have been asked to give expert testimony. Mostly, this analysis made it possible to come closer to or to actually "solve" the question of an anonymous writership, which in turn was of critical importance for the analysis of certain cases of anonymous authorship (cf. the list given below).

The theoretical difficulties in defining idiosyncrasy in/of linguistic and orthographic behavior cannot be discussed here in greater detail. Ever since Edward Sapir's reflections on "Speech as a personality trait" (Sapir 1927), linguists have been intrigued by idiolectal forms of linguistic (and orthographic) behavior, and forensic linguistic analyses have made reference to this (cf. Kniffka 1990b). But little systematic research concerning the idiosyncrasy of linguistic and orthographic behavior has been made to date.

This is an absolute desideratum for future basic and applied linguistic research, mainly for forensic linguistics. Many elementary questions and basic notions will have to be (re-)examined, if not modified: What are the basic ingredients of "idiolect" on the various levels of linguistic description? What of orthographic "idioscript"? Are an idiosyncratic way of speaking and an idiosyncratic manner of writing behavior identical

or are they a subset of “idiolectal features” of language behavior? Or is it, rather, a cover term for idiolect or for data seemingly specific of idiolect, but in fact belonging to group-specific language and orthographic behavior? Is idiolect and idiosyncratic behavior definable by the notion of “deviation” or “variation” alone?

One may wish to borrow the distinction between “oppositional definition” and “identificatory definition” from the old days of phonology, which has played an important role and paved the way for an enormous progress in this science ever since the Prague School. Undoubtedly, the need for an oppositional definition of “idiolectal” and/or “idiosyncratic” linguistic and orthographic behavior exists. Most linguistic and, even more so, forensic linguistic studies seem to have been content with giving oppositional definitions of the idiolect, rather than truly identificatory definitions, which will be needed for an adequate description and explanation.

For orthographic behavior, much more truly interdisciplinary research is necessary to make any progress in reaching an identificatory definition of idiolect, including data from social psychology, cognitive sciences, handwriting analysis, cognitive linguistics, developmental psycholinguistics, sociolinguistics, computational linguistics and other sciences.

Some of the key questions are: Which empirical predications for which quantitatively definable entities can be given? Should the idiolectal linguistic and orthographic behavior (each to be analyzed in its own terms) and the ISC be empirically stated and calculated for a single text (item), a text as a whole, the totality of texts of a corpus (for example, of incriminating letters of a suspect), a text genre altogether, or units larger than genre and text type? All these questions show that there is an immense amount of work waiting to be done in forensic linguistics in the next decades.

Leaving the more theoretical questions aside for the moment, what are the next steps for making some modest progress in the future? It is of prime importance to provide at least a working definition of the ISC of orthographic behavior. Here is one suggestion:

The ISC of features of orthographic behavior is the dynamic value or set of values assessed to the frequency of occurrence of certain orthographic types and/or tokens a writer consistently uses/shows/displays in a particular text type in structurally similar linguistic and situational contexts.

The ISC is a dynamic entity; that is, it is not definable as a material orthographic item, enriched by some (accidental) data of its occurrence,

but its status is projected to the orthographic behavior of a writer in a particular text type.

The reason that “text type” is chosen as an empirical reference results partly from more general linguistic considerations. As discussed elsewhere (cf. Kniffka 1981; 1990b; 1992; 1998; 2000a; 2000b; 2001), the consistency of the linguistic and orthographic behavior of individual speakers can only be defined in relation to or for particular *text types*. It is an empirical question whether and to what extent it is possible to define speaker-specific ways of linguistic and/or orthographic behavior across text types or even across all text types a speaker has at her/his disposal.

All empirical data available suggest that it is justified to say that Mr. X and Mrs. Y each show a (distinguishable) speaker-specific idiolectal manner of linguistic and orthographic behavior in writing (for example, a letter of condolence, a letter of complaint to the post office or a letter of congratulations to a nephew who just passed an exam). The argument that text types are the main category of reference is made for orthographic behavior, *cum grano salis* and with all precautions necessary, by analogy to linguistic behavior. From a commonsense point of view, one could perhaps assume that writing behavior is somehow more specific. As stated above, this also has to be studied thoroughly from the point of view of handwriting experts (cf. Hecker 2000; Michel 1982).

The most important, decisive and rather new arguments come from the empirical data analyzed for the text type-specific occurrence or non-occurrence of certain orthographic features (cf. the third section of this chapter). As mentioned above, some features of orthographic behavior can be assigned a relatively low ISC, some a rather high ISC. Table 8.7 lists some examples for a rather *low* ISC.

A written text of everyday communication may belong to any text type, such as a business letter, a personal letter, a news story in a newspaper, an academic report, a course description in a university catalogue and so on.

Table 8.7 Low idiosyncrasy coefficients of OD in a written text of everyday communication

-
- Excessive use of quotation marks (in a normal written text of everyday communication)
 - Excessive use of hyphens (*SPD-ler*)
 - Excessive use of abbreviations
 - Preference of writing numbers rather than words for numerals
 - Parenthesis dash without empty space before and after text
-

In contrast, the date line of a business letter (for example, in the usage of words rather than numbers) could represent a rather higher degree of exceptionality or uniqueness. This example also shows that it is again the specification of a particular text type and its general regularities that matter here, defined in relation to certain writer populations, within a culture, within fractions of a culture, within regions, within age groups. Also, this is an appropriate illustration that a sociolinguistic perspective has to be addressed in addition to a text linguistic one.

If all these data and definitions are to be taken *cum grano salis*, if only a statement in relative terms in regard to text types can be made on safe empirical grounds, which data can be named as having a high ISC in the repertoire of orthographic features?

One example has been discussed on pp. 194–7 (Case 1), the customary as opposed to the exceptional way a comma is curved in German (Latin) handwriting: The customary way is that it is curved to the right, similar to a closing parenthesis (...) ...), the exceptional, marked, way is a curving to the left, similar to an opening parenthesis (... (...), as in the example given in Figure 8.1(b).

The orthographic features, which in several decades of expert testimony turned out to have the relatively highest ISC of written data I have come across, are the omission marks (called “ellipsis” in the US and UK) in the form of three dots in a particular linguistic and textual context. The occurrence of these three dots at the end of a quotation, a sentence, a text, or a text paragraph, is represented in Table 8.8.

Table 8.8 Omission marks (three dots) at the end of a quotation at the end of a sentence at the end of a text-paragraph¹

(1) German norm as stated in DUDEN 2000, §100, S. 1151 (DIN-Norm 2008):

Instead of x ...". you write x ..."

(2) Version occurring in an incriminating text (5 times):

x .."

(3) Norm reported by British natives living in Germany:

x"
↑ [empty space]

(4) Norm reported by British natives living in Britain:

x ... ". (taught in universities)
↑[empty space]

(5) x ..." (taught in schools)

(6) Norm reported by US forensic linguists for an “ellipsis”:

x"

The writing standard of the three dots in modern German as prescribed in DUDEN (2000, § 100, p. 1151), which is the highest orthographic authority in today's Germany, in particular with laymen, is a theoretical rather than an actual usage norm. This norm as stated in Table 8.8 (1) requires that instead of <...".> at the end of a quotation at the sentence ending, one should write <...".>; that is, leaving out the final full stop.

The version <...".> given in Table 8.8 (2) occurred in five instances in incriminating texts in a defamation case several years ago, and in no instances in the comparison data. This did not contribute much to answering the question about who the author of the anonymous defamation letters was – though by the very nature of the orthographic variant, it seemed to be potentially of high significance for the analysis of writership.

Informal inquiries with native British and American colleagues seemed to reveal various highly language- and culture-specific variants for this item. British colleagues living in Germany stated that the only variant in use in Britain is <... ".> (3). The salient contrastive aspect of the British and the German way of writing, as the sources claimed, is that, in Britain, the quotation mark would always be at the end of the sentence. Interviews with native British colleagues living in Britain showed that there is a double standard, one taught in universities <... ".> (4) and one taught in schools <...".> (5). So, the unsystematic analysis of the British system (see also Truss 2003) revealed three different variants in the perspective of (academically trained) linguists.

American (academically trained forensic linguistic) colleagues affirmed that, in the US, there is only one accepted standard used in universities (also represented in the MLA style sheet): usage of three dots known as an "ellipsis" and a fourth dot (even without an empty space) and a following quotation mark: <"> (6).

Such highly culture-specific features of orthographic behavior may be data of critical importance in a case. They could lead to a clear distinction among, for example, German as opposed to British and/or US writers, particularly as long as significant configurations of features of orthographic and linguistic data can be shown to exist. For the case in question, it was only the German situation that was analyzed in more detail.

The variant occurring in the incriminating texts <...".> seems to be an item of a relatively high ISC. I had never seen or used it myself. Neither had my wife or any of the German forensic linguistic colleagues I talked to. One has to realize that this usage (type) occurs relatively rarely in texts anyway. Many people, including me, would show some kind of an "avoidance behavior" in such a situation. If one does not precisely

remember the norm as given in the DUDEN (cf. Table 8.8 above), one would rearrange the parts of the utterance so that a possible spelling mistake would be avoided.

All data stated do not really amount to much if *hard empirical* data are lacking in a forensic linguistic AA, as has been stated by many statisticians, people undertaking stochastic analyses of texts (cf. Grotjahn 1979), and also forensic linguists (cf. Chaski 2001; Kniffka 1990b; 1996c). Data obtained through introspection and certain theoretical norms as given in the DUDEN do not help much. In any case, additional empirical analyses and tests are necessary. To obtain some more reliable information about the factual use of this form, an informal survey was made with the teaching staff of courses for professional secretaries at several chambers of commerce (IHK = Industrie- und Handelskammer) in two larger German cities. Briefly summarized, the result was that none of the teachers and the students knew the variant that occurred in the incriminating letters (<..>), and that they all knew the “correct” norm (the DIN-norm 2008, which is the collection of spelling norms for professional secretaries). All agreed that they had never seen this variant used, would never use it themselves, and found it very odd (see more explanations in Kniffka 2003b; and p. 188ff. in chapter 7 of this volume). These data of spelling and punctuation in German business communication, mainly business letters, acquired by participant observation and analyses of professional secretaries’ writing behavior, clearly show that the variant <..> is not in use, is unknown and judged to be very odd.

In yet another miniature experiment with German second-year students of linguistics, an attempt was made to find out how non-professional writers/typists would assess such a variant, and which they would use themselves. Table 8.9 shows the results of this miniature experiment with 86 subjects, participating in an introductory syntax class of linguistics at Bonn University in 2005, in which the subjects were orally instructed to write *Du dumme ...* (‘you stupid ...’) as a quotation at the end of a sentence at the end of a paragraph. This type of a linguistic context is common in German. Frequently, words such as *Du dumme Kuh* (‘You stupid cow’), *Du dumme Ziege* (‘You stupid goat’) occur colloquially. These nouns, being derogatory, are frequently omitted in every day written texts.

The most frequently occurring variant is that with an exclamation mark *Du dumme ... !*, which purposefully was left open in the instruction. Out of 86 subjects, 36 have this variant. They, at least in part, seem to apply some kind of an “avoidance behavior”, since the exclamation mark after the three omission dots solves the punctuation problem elegantly.

Table 8.9 Results of spelling experiment with German students (aged 21–28): *Du dumme ...*

Variant	Number of occurrences
(1) <i>Du dumme ... !</i>	36
(2) <i>Du dumme ...</i>	28
(3) <i>Du dumme</i>	5
(4) <i>Du dumme</i>	4
(5) <i>Du dumme ...</i> (Full stop exactly under quotation mark)	3
(6) <i>Du dumme ... !!</i>	1
(7) <i>Du dumme ..!</i>	1
(8) <i>Du dumme ... !.</i>	1
(9) <i>Du dumme</i>	7

Note: n = 86

Less than one third (28) of the subjects gave the “right” version as prescribed in the DUDEN (2000, §100, p.1151): <“*Du dumme ...*”>, and only five used the variant that is explicitly described as “wrong”: <“*Du dumme ...*”.>. Four subjects used a variant with a full stop directly adjacent to the three omission dots (that is, four dots altogether) and followed by the quotation mark: <“*Du dumme*”>, and three subjects used a variation with three omission dots, quotation mark and the full stop placed directly under the quotation mark: <“*Du dumme*”>. Seven subjects seemingly did not understand the instructions (9), and some subjects used even more “exotic” variations (three omission dots, one full stop, and an exclamation mark, followed by a quotation mark: <“*Du dumme ... !!*”> (6); two omission dots and an exclamation mark followed by a quotation mark: <“*Du dumme..!*”> (7); three omission dots, one full stop, an exclamation mark, and a quotation mark and another full stop at the end: <“*Du dumme ... !.*”> (8)).

None of the 86 subjects used the variant that occurred in the incriminating text <“*..*”.>.

As described elsewhere (cf. Kniffka 2003b), a relatively high value of ISC does not necessarily have an impact for a *liquet/non-liquet* of the analysis of orthographic and also linguistic behavior. OD with a high value of ISC do not guarantee a *liquet* of the analysis of an anonymous writership. For the particular feature in question, it is conceivable that the unique variant found several times in the incriminating text <“*..*”.> could, theoretically, have spread in a class of professional secretaries in

which the teacher, for one reason or another, chose to teach this “exotic” variant (which is very rare in real life, of course). If the teacher had used this instead of the correct norm, the prospective secretaries of that class would, perhaps, have internalized it as a correct form and there would be a more frequent occurrence.

It is sufficient to state here that there is considerable variation of the ISC of orthographic variants, which is to be described on the basis of an empirical investigation. Writing errors such as capitalization as opposed to non-capitalization, writing <ß> and <ss> in German (according to different norms and spelling reforms) are much more important and frequent, and less idiosyncratic in real life texts of any denomination, be they incriminating or not, than, for example, the curving of a comma to the left, or the use of an “exotic” variant for the three dots at the end of a quotation at the end of a text in German.

Text type-specific distribution and constraints of linguistic and orthographic features

When describing the heuristic continua of the features of orthographic behavior above, it was mentioned that text types are amongst the salient, if not the most salient, textual categories. Most forensic linguistic analyses *de facto* analyze data on a text type-specific level, be it explicitly or (in most cases) inexplicitly.

Contrasting extortion and blackmail letters versus defamation and libel letters: a heuristic taxonomy

It seems worthwhile to state some established orthographic and grammatical features of incriminating texts in terms of the contrast of extortion/blackmail letters and defamation/libel letters. Tables 10a and 10b list some of the most important features. They are an adapted version of the lists first given in Kniffka (2001: 89–91). All of the features are gathered from authentic incriminating letters on which I was asked to give expert testimony in the last 30 years or so.

It should be noted here explicitly that these are heuristic, simplified, abbreviated and non-exhaustive lists selected from a larger number of features described in expert opinions on AA. They appear to be almost (proto)typical. A large-scale quantitative analysis is neither possible nor intended here. Therefore, no exact proportions or numerical indices are given for the various types and/or tokens. Except for the more general contrast that extortion and blackmail letters tend to be shorter (less than 200 words) and defamation and libel letters tend to be longer (more

Table 8.10a Text type-specific contrast(s) of *grammatical*, *textual* and *semanto-pragmatic* features in anonymous extortion and blackmail letters versus defamation and libel letters

Extortion/blackmail letters	Defamation/libel letters
<ul style="list-style-type: none"> - Length <200 words - Less hypotaxis constructions, embedded sentences, etc. - Predominance of short sentences - Limited variation of syntactic constructions - Syntactically little “innovative” - Customary constructions - Predominance of prefabricated expressions - Limited amount of adjectival attributes - Few new coinages - Few unusual expressions - Frequently “telegram style” - Little emotional and emotive vocabulary - Few syntactical criteria of emotional way of speaking - No extensive characterizations of people - Few value judgments - Little delight and vanity in (own) formulations - Little preference for word play, puns, etc. - Little drive for language correction - With or without address and date line - Rarely non-explicit signator/signature - No metalinguistic, metatextual, and metapragmatic comments and explanations - No value judgments by picking on addressee or addressee’s characteristics 	<ul style="list-style-type: none"> - Length >200 (400?) words - Numerous hypotaxis constructions, embedded sentences, etc. - No predominance of short sentences - No limit of variation of syntactic constructions - Syntactically frequently “innovative” - Also uncustomary constructions - No such predominance - Excessive usage of adjectival attributes - Relatively many new coinages - Many unusual expressions - Rarely “telegram style” - Great deal of emotional vocabulary - Many syntactical criteria of emotional way of speaking - Very extensive characterizations of people - Many value judgments - Strong delight and vanity in (own) formulations - Strong preference for word play, puns, etc. - Strong and sophisticated language correction behavior - With address and date line - Frequently non-explicit signator/signature (e.g. <i>Some distressed employees</i>) - Frequently metalinguistic, metatextual, and metapragmatic comments (e.g. <i>We are sorry to be able to write to you anonymously only</i>) - Frequent value judgments by picking on addressee or addressee’s characteristics (e.g. <i>You dirty rat! You look like ...</i>)

Table 8.10b Text type-specific contrast(s) of *graphemic*, *orthographic* and *layout* features in anonymous extortion and blackmail letters versus defamation and libel letters

Extortion/blackmail letters	Defamation/libel letters
- Little use of typographic variation	- Frequent and intensive use of typographic variation (bold-face, different sizes of letters, italics, etc.)
- "Unmarked" use of punctuation marks	- Highly "marked", "excessive" use of punctuation marks (several of a kind, ?????, !!!!!)
- Not necessarily standard "letter" format/layout	- Standard "letter" format/layout
- No "graphemic variation and/or ornamentation"	"Graphemic variation and/or ornamentation" (e.g. use of bold types, capital letters, italics, etc.)
- Bare text with or without (rudimentary) punctuation	- Text with standard/customary punctuation
- No innovations in spelling of words, "new" abbreviations, etc.	- Frequent (use of) innovations in spelling of words, new (ly invented) abbreviations, etc.

than 200 to 400 words), it is not possible to give any definite numerical restrictions for either. I have worked on a defamation case with some 100 A4 printed pages of incriminating letters and also cases with an average of only one half of an A4 printed page (which is rather small for this text type). I have worked on extortion cases with a whole A4 page written in stencil script (cf. Case 3, pp. 201–5) and on extortion letters consisting of only two or three lines (covering the place, time and specification as to where a ransom sum was to be deposited). One can even say that, at this pioneering stage, an exact quantitative statistical analysis is not yet possible. It is definitely a challenge for forensic linguistics and a task of prime importance for the next generation(s) to come.

These lists, primarily, have the aim of giving a first-hand illustration of some important textual characteristics and of the text type-specific contrastive aspects of extortion as opposed to defamation letters. I have tried to choose rather unconventional and (purposefully) "pre-theoretical" terms to avoid the impression that these are well-defined and established textual categories in forensic or general linguistics. They are heuristic concepts modeled according to practical endeavors and concerns in forensic linguistic expert opinions, which allow for and need further modification. It is empirically safe to say that the oppositions stated represent yet another continuum each, not simple binary values.

Also, it is empirically safe to say that it would *not* be adequate to state them in the reverse distribution, when text type-specific general constraints, customary constellations and so on are described. I would not hesitate to describe the oppositions listed in Tables 8.10a and 8.10b as salient and “prototypical” for the two text types.

The most important empirical fact, however, is that all regularities, tendencies and distribution patterns described above represent something like “sociolinguistic rules”. In plain words, the distribution patterns stated do not have to be the way they are, and they do not work the way grammatical rules work. They could be different, but they are not. To give just one example, it is not impossible that an extortion/blackmail letter could contain personal evaluations. But according to my experience working with extortion and blackmail letters and other data I have seen, this is extremely rare. Defamation/libel letters, on the other hand, tend to have many personal evaluations as ways of specifying the addressee. German expressions such as *Du verdammtes Schwein!* (‘You damned pig!’), *Du dreckiger Heuchler!* (‘You dirty hypocrite!’), *Du schulterzuckender Lackaffe!* (‘You shoulder-shrugging fop!’) and others of a very large repertoire of “creative” labels for addressees in defamation letters are very frequent. Although it does not seem possible to give any detailed numerical descriptions (which would presuppose a large-scale corpus linguistic analysis of several thousands of original extortion and defamation letters each), general contrasts such as the ones stated in Tables 8.10a and 8.10b seem valid and salient approximations. A corpus linguistic quantitative analysis would be one of the prerequisites for an exact empirical determination of the general areas and dimensions of the investigation and also of the number and size of the “measuring instances” to be used.

It seems that as a general device (1) the introduction of oppositional and identifying definitions, (2) the definition of standard and “prototypical” versions of the features, and (3) the distinction of marked and unmarked features and distributions of features would be useful.

The last mentioned distinction (marked/unmarked) seems of particular methodological importance for the description of the contrast of extortion/blackmail letters and defamation/libel letters. As stated above, we are dealing with a system of sociolinguistic rules, which means that the features can be expressed in different degrees and values of a continuum, and also that the form, in principle, can vary. This may be of critical importance for AA and the criminalistic endeavors as a whole. To use the same example again; if an extortion or blackmail letter *does* make use of personal evaluations (which is extremely rare), it would

mean that it is an atypical rather than a typical variant of this text type. This could possibly be an indicator (in particular together with other atypical characteristics for this text type of linguistic and orthographic behavior) that, in simplified terms, this is not likely to be a “professional” extortion/blackmail letter but rather an “amateur” extortion/blackmail letter. This, in turn, could help criminalists in determining whether it is a serious or real extortion letter, or a fake written by someone who wants to join the bandwagon. This is quite a common phenomenon in many extortion and blackmail cases that have gained publicity. It could also occur when someone tries to make a (bad) joke. So, OD of this kind can be of great help in the investigating process for the authorities.

Marked forms, distribution patterns, and larger co-occurrence structures of linguistic and orthographic features, as opposed to unmarked forms, features and patterns stated in relation to a particular text type, may be the most important data of all for AA and the criminalistic goal of identifying criminals. Consequently, special deviations as marked features concerning a particular text type of incriminating texts potentially have a very important diagnostic potential in AA – much more so than deviations (such as mistakes, errors) *per se* could have measured against the grammatical norm as described in prescriptive grammars, the DUDEN and others.

What is stated here for linguistic features holds, in principle and *mutatis mutandis*, for graphemic and orthographic features as well. If five or six items of a punctuation mark – say, <?????> or <!!!!> – occur in an extortion letter, this may be a very important indicator of the “professional” or “amateur” status of the writer of that particular text.

Two important precautions already mentioned should be stated again:

(1) There is no simple one-to-one proportion for the contrast of features of one (incriminating) text type as opposed to another. Generally speaking, the contrast of no or very few value judgments as opposed to many value judgments and, similarly, of five question marks or exclamation marks in a row rather than the occurrence of one only is, in fact, a scalar continuum rather than a dichotomy;

(2) The second precaution is (cf. the second section of this chapter) that an unmarked form or distribution is not by definition less important, less relevant or less valuable as a diagnostic potential than a marked form or distribution. The latter *per se* is not necessarily more helpful in AA. In many cases in real life, however, it is the marked form, in particular if stated as a feature configuration on a larger scale, which can be the most helpful in the practical analysis of anonymous authorship and writership.

Some more general methodological concerns

The lists of text type-specific contrasts of extortion/blackmail letters as opposed to defamation/libel letters have been given to indicate the main endeavor that forensic linguistics pursues in AA. This is to set up holistic feature configurations for entire texts, sets of texts that occur in a particular case and, perhaps, text types across case boundaries describing a text type and a genre of incriminating text as a whole. Bundles of features and feature configurations on a larger scale can be combined to create "Ganzheit" pictures, which may reflect the linguistic and orthographic use of text types by particular speakers and/or writers. The more expansive and detailed such covariation bundles of features that vary, in principle, independently of one another, the higher the likelihood that certain dynamic entities of linguistic and orthographic use can be stated. They, in turn, may be matched to the overall holistic picture that is described for the incriminating and for the comparison data. Not infrequently, significant structural similarities of dynamic behavior can be stated that can be used as likelihood indicators for the authorship and writership of the texts as a whole, and be offered as valid auxiliary data to investigators.

I am a little hesitant to use the expressions "profile" or "linguistic profile" for such textual constellations, since we are dealing with text products, qualities, structures and other qualifications of texts, incriminating and comparison data, written texts – *not* with personality profiles of speakers and writers. The term "profile", from my point of view, should best be reserved for the description and analysis of personalities in psychology. If the term can be defined in a solid empirical manner for language data and texts (linguistic profiles of dynamic values of language and orthographic behavior), there can be no objection to it.

It has been expressed time and again in forensic linguistic literature (cf. Kniffka 1990a; 1996c; 2000a; Levi and Walker 1990; Rieber and Stewart 1990; Shuy 1993a; Solan 1993; Tiersma 1999; and others) that forensic linguistic analyses are services of an auxiliary science that are of prime importance for the criminal investigation. In German court cases, they are of much more relevance in the so-called "Ermittlungsverfahren" ('investigation stage') in the German legal system, than in the "Hauptverhandlung" ('main trial', where the verdict is fixed).

The diagnostic potential of a forensic linguistic expert opinion is generally acceptable, advisable and useful for the investigation process. I would not subscribe to the idea, however, that a verdict or a conviction

should be based solely on forensic linguistic, forensic phonetic or forensic handwriting evidence.

In the regular real life situation of a case with anonymous incriminating texts and comparison texts, the analysis and application of such general feature configurations have proved to be very helpful. They have to be backed by quantitative (statistical) data. If, for example, a very poor use of vocabulary in an anonymous incriminating defamation letter co-varies with an exceptionally flawless command of German punctuation rules of the latest date, this may be a noteworthy fact for the forensic linguistic authorship and writership analysis, and for the search for further configurations of features (maybe for fake or disguise of authorship). An excellent knowledge of German grammar would, normally, rarely go together with a deficient knowledge and actual usage of German punctuation. If such a covariation can be stated for both an incriminating anonymous text and the comparison texts, some interesting perspectives for the analysis of the authorship and writership of the texts could be gained. This means that there are no overall or *a priori* generalizations possible as to which features go together with which other features.

A clarification of the following questions by the forensic linguist seems of particular importance:

(1) Which textual, linguistic and orthographic characteristics can serve as a safe basis of an empirical distinction of “professional” (or serious) extortion/blackmail (or any other incriminating) letter as opposed to “amateur” extortion/blackmail (or any other) letters? On which textual linguistic and orthographic characteristics can it *not* be based?

(2) Are there “typical” and “atypical” linguistic and orthographic forms and features for particular text types; for example, for extortion/blackmail letters versus all other types of incriminating letters; for defamation/libel letters versus all other incriminating letters? (cf. Tables 8.10a and 8.10b).

(3) Which feature configurations or feature value configurations holistically describe extortion/blackmail letters as a text type altogether, and which describe defamation/libel letters?

(4) Which typical “borrowings” of ways of grammatical, textual and orthographic behavior of non-criminal text types are similar to those in incriminating texts? Which are taken over into or are represented in incriminating text types (and, if possible, vice versa)? Which general “paths of borrowing” exist?

These are some of the questions and prerequisites that would help to base AA on more solid grounds. Thereby, several oppositional and identificatory definitions of the various text types of incriminating and non-incriminating texts can be set up. Eventually, on this basis, “prototypes” of extortion/blackmail letters, defamation/libel letters, and other text types can be acquired for German that can be compared with and measured against those for other languages.

The heuristic list of features, which show a contrast of extortion/blackmail letters as opposed to defamation/libel letters given in Tables 8.10a and 8.10b should not lead to the assumption that there are only uncommon and different features for the two text types contrasted. The same descriptive, explanatory and, hence, diagnostic potential that can be gained from the differences and contrasts of two or more text types can, in principle, be drawn from the amount and structure of the similarities, correspondences and “non-contrasts”. In particular, the structural invariants for different text types are of interest. Almost trivially, it is not only important to study and describe the differences between text types, but also the coincidences and similarities. So, lists will also have to be set up for the following questions: Which features and feature values are found in *both* extortion/blackmail letters and defamation/libel letters? Which can also be found in letters of *non*-criminal denomination? Which overlap in which specific proportion and structure, and can be stated? Are there recurrent patterns of the distribution of features and feature configurations?

Obviously, though not trivially, extortion and defamation letters will show grammatical errors and mistakes, such as in morphological case forms or in number agreement. Also, they will share lexico-semantic deviations, such as regional variants of word meanings, characteristics of spoken language rather than written language. As far as orthographic behavior is concerned, one could state the occurrence of customary everyday orthographic errors and mistakes, the use of a non-professional layout in “official” incriminating letters and others.

Such forms and characteristics of linguistic and orthographic behavior found in incriminating letters have to be contrasted, on a large scale, with the general characteristics of orthographic and linguistic behavior in other than incriminating letters. The predominant analytical perspective in these forensic linguistic endeavors is to describe texts as holistic entities, and to give detailed qualitative and quantitative descriptions of texts as a whole in a systematic and empirically sound way. The data acquired for a particular corpus of a language X should be examined in

relation to other corpora in the same language, and in turn be contrasted with the results gathered for other languages Y and Z.

This may lead to a satisfactory qualitative and quantitative definition of measuring instances and textual parameters of linguistic and orthographic features, which may entail some kind of a field guide for the analysis of incriminating texts. Not just as a checklist, however. If it is used as a checklist in a real life case of AA, it may be of help for investigators, at least for the people that do not have a linguistic training. An incomplete list of some of the parameters and text type-specific measuring instances is illustrated in Table 8.11.

Table 8.11 Textual parameters and measuring instances

(1) Total length of the text product (various measuring instances);
(2) Mean length and typical length of a text product of a particular text type;
(3) Total number of words per text;
(4) Total number of words per sentence (sentence type, phrase, etc.);
(5) Total number of lexical morphemes and words;
(6) Total number of grammatical morphemes and function words;
(7) Proportion of function words versus lexical words per text;
(8) Proportion of function words versus lexical words per text section/paragraph;
(9) Proportion of function words versus lexical words per sentence;
(10) Proportion of regular versus deviating word formations;
(11) Proportion of hypotactic versus paratactic constructions;
(12) Proportion of nominalizations versus non-nominalizations;
(13) Proportion of finite versus infinite verb forms;
(14) Total number and classes of orthographic deviations (errors and mistakes);
(15) Total number and classes of deviations in punctuation (errors and mistakes);
(16) Total number of types of excessive use of punctuation marks;
(17) Total number and classes of deviating "layout types"
(18) ...

In addition, all the graphemic and orthographic features listed in Table 8.10b could and should be systematically analyzed.

Summary

The exemplaric discussion of orthographic features above reveals that a systematic analysis of data of orthographic behavior is a salient intrinsic part of AA. It is shown also that, in many cases, indirect evidence can be gathered from OD for the question of anonymous authorship. No generalizations of any kind can be made, however, except that OD pertain at any rate to writership only and, when used with extreme

caution and *cum grano salis*, may be of indirect relevance for the analysis of authorship. Data of orthographic behavior may prove to be a very important diagnostic potential that can be of help in or for an AA as a whole. OD themselves do not by any means guarantee a solution to questions of anonymous authorship. The analysis of a partial identity of the writership of anonymous incriminating texts and comparison texts may be of methodological interest for linguists, and be of an auxiliary value for investigating criminalists. Generally speaking, the “cumulative principle” of relevance in other areas of linguistics (for example, socio-linguistic description) is usable as some kind of a core category of forensic linguistic analysis.

Whether and to what extent the analysis of the structural and the factual diagnostic potential of OD is relevant for anonymous authorship and writership analysis, one does not know *a priori*. It can only be learned from a systematic empirical analysis of all data available. There is no justification, by methodological criteria or language data of any kind whatsoever, for the total exclusion of OD from AA. On the contrary, orthographic (as well as linguistic) data, stated for a certain set of incriminating and comparison texts, frequently allow solid empirical predications on and about the texts. There is no scientific empirically safe way of jumping from texts to conclusions concerning the personalities of the author and/or writer of these texts. This would be absurd at any rate, considering how atypical a text of three or four pages is for the total linguistic and orthographic behavior of any person as a whole. In a forensic linguistic expert testimony, that would be a practice not performed *lege artis*.

FLAA will have to take data of orthographic behavior into account in a much more solid, intensive, and thorough fashion than it has to date. The exemplaric description of certain data of orthographic behavior given above also illustrates the enormous amount of work that forensic linguistics will have to undertake, and can accomplish, in the years and decades to come.

Part IV

Outlook and Future Tasks

This page intentionally left blank

9

Language and Law: Some Needs

This chapter deals with misunderstandings (mainly) between linguists and lawyers inside and outside the courtroom. It describes and explains how and why they happen, and points out how they could be addressed in terms of a preliminary needs analysis in a truly interdisciplinary perspective. It is hoped that this will contribute to an understanding of how to deal with certain salient questions of forensic interaction in a more satisfactory fashion, and eventually help solve some problems linguists and lawyers have (with one another) in their daily routine.

Introduction

The main concern of this chapter is to make a modest exemplaric contribution from linguistic basic research and applied (forensic) linguistic research that will lead to a better understanding between linguists and lawyers or, more specifically, between “linguistic people” and “legal people” who are required to cooperate in performing their jobs in a forensic setting (see below). The overall hypothesis is that there is much more than is generally assumed that can lead to improvements in interdisciplinary forensic understanding in terms of prerequisites, dimensions of perspectives to be incorporated and depth of reflection. There are hardly any systematic endeavors available in terms of a needs analysis – at least, not in terms that can be realistically adapted by law and language practitioners. This chapter is neither concerned with the teaching of legal terminology to law students and/or linguistic students, nor with teaching linguistics to law students. Instead, it deals with some prerequisites, some of their underlying principles from the (general and applied) linguist’s perspective and gives exemplaric illustrations of them. I would argue that the prerequisites discussed, and many others not discussed, are of critical

importance for interdisciplinary action of one kind or another between linguists and lawyers; only a very few can be discussed here.¹

Leaving aside all theoretical concerns proper, a heuristic taxonomy as given in Table 9.1 below can be set up for needs for linguist/lawyer-interaction and their scientific analysis. It is based on a more general division of three classes of needs in the area of language and law:

- (A) Needs of/in scientific knowledge of one's own and of others' academic fields;
- (B) Needs of/in theoretical and empirical methodology from an intra- and interdisciplinarity perspective;
- (C) Needs of/in practical operational know-how of linguistic expert testimony.

In each class a threefold distinction, as given in Table 9.1, is made.

This taxonomy is to serve as a first heuristic grid to interpret the data described below for the formulation of needs. I refrain from describing each need by a particular example because, while this would create a neat picture in terms of systematicity, it would be too artificial, distorted and "too good to be true" in terms of real life forensic interaction. The relevance of the particular point shines through, however.

In the first section of this chapter, instances of miscommunication between linguists and lawyers, including misconceptions of each other's work, will be discussed. This is the main objective of this chapter as a whole. The majority of examples given are located in this realm.

The second section provides specific problems in forensic interdisciplinary interaction in more detail, focusing on examples of "traps" and "tricks of the trade" and their reception by the people involved.

The final section gives a brief demonstration of a few (types of) intrinsic language problems that turned out to be of particular concern in the understanding of non-linguists, representing larger deficits of linguistic applied and basic research. The examples discussed are mainly from German orthography and lexical morphology. All linguistic and forensic data discussed are authentic data stemming from actual forensic cases in which the author was asked to give expert testimony.

Many of the problems in communication arise from the very fact that (at least) two different sciences and their corresponding "Praxisfelder" ('fields of practice') interact in a forensic context. Another equally large number of problems, widely neglected in linguistic and other research to date, arise from the fact that neither of the two sciences represent a homogeneous block but, rather, a mixed conglomerate of heterogeneous

Table 9.1 Heuristic taxonomy of needs of cooperations in the field of language and law

A	Needs of epistemological scientific knowledge (desiderata of academic and scientific endeavors)
A1	Lack of knowledge of the other field(s) (for example, linguists' ignorance of the requirements of legal sciences and vice versa); <i>interdisciplinary needs</i> ;
A2	Lack of experience and knowledge within one's own field (for example, general linguists' ignorance of applied linguistics concerns; applied linguists' lack of width and depth in general linguistics); <i>autodisciplinary needs</i> ;
A3	Lack of general academic/scientific background and basic science knowledge (level of reflection in empirical science, test theory, theory of statistics, "current theory"); (<i>general scientific needs</i>).
B	Methodology-oriented scientific needs (desiderata of scientific method)
B1	Lack of general empirical methodology (for example, of analyzing data of language use); (<i>general methodological needs</i>);
B2	Lack of technical expertise (for example, in know-how of sampling, measuring and analyzing linguistic data); <i>technical needs</i> ;
B3	Lack of interdisciplinary "pedagogical" design (for example, of addressee-specific adaptation and application of linguistic findings, terminology, style); <i>interdisciplinary adaptation/application needs</i> .
C	Practical operational needs (desiderata of real life performance of the applied scientist)
C1	Lack of professional "realities", expectations, projections of practitioners (for example, linguists' ignorance of courtroom routine, judges' preoccupation with language and language use, behavioral traits, mutual (in)sensitivity regarding salient issues of own science, practical errors in dealing with semiotic syndromes in behavior of the police, the defendants, adaptation to courtroom interaction); <i>professional forensic needs</i> ;
C2	Lack of practical experience in dealing with the "tricks of the trade" (for example, linguists' ignorance of and blindness towards lawyers' routines to have linguistic experts declared biased, cross-examination behavior); <i>"tricks of the trade" needs</i> ;
C3	Lack of experience in and of adaptation to clients' needs (for example, linguistic expert testimony's account of judges' and/or lawyers' interests, needs, wishes, concerns (e.g. strive for exact percentages data), linguistic training, capacity to read linguistic texts); <i>users'/clients' practical needs</i> .

constituents and various subdisciplines operating on different levels. Several groups and subgroups of the law and of linguistics and representations thereof are involved.

The expression “linguists and lawyers”, which is used throughout this chapter for reasons of simplicity, is in fact too global and inadequate a label for such a heterogeneous array of participants. One could use (purposefully non-established) terms such as “law people” and “linguistic people”. “Law people” would include judges, lawyers, and others professionally involved with the law. “Linguistic people” includes all people who write and present expert opinions in courts on language matters. In Germany, only a fraction of the latter are professional linguists, or people who are actually working as linguists in universities. Many linguistic experts or so-called “experts” are people from other sciences, from law and “criminalistics”, philologists of one denomination or another, or people with no or little academic training; for example, policemen and lay assessors. They have to deal with language data in a forensic context as a necessity of their job. Lay assessors, who are asked by the judge for an opinion (for example, whether a word X used in a special context is an “insult” or not), are asked to give an opinion on a language matter as though they were actually experts in language.

Whether professional linguists like it or not, there will always be “linguistic expert testimony” from people other than linguistic experts. This also holds for other sciences involved in forensic interaction. There will also be less competent experts among linguists as well as among other forensic scientists and practitioners. Linguistic experts must respond to both less competent linguists and non-linguists whose claim of expertise in linguistics is doubtful.

Finally, it is of critical importance to appreciate that what linguists regard as conclusive may *not* be regarded as conclusive in a court. Convincing a fellow linguist and convincing a judge can be two very different matters. This is a pivotal point around which many theoretical, metatheoretical, practical and “metapractical” considerations, including reflections on a needs analysis, are centered.

Linguists and lawyers: misconceptions of each and of each other

Linguists’ and lawyers’ interaction

The relationship can be described by a fourfold distinction of “attitudinal perspectives”:

- (1) How do lawyers view their own and other lawyers’ practice?

- (2) How do lawyers view and assess linguists as experts in court, including their attitude towards (their own and other lawyers') language?
- (3) How do linguists view lawyers?
- (4) How do linguists view themselves and their activity as court experts?

Based on my own experience, only perspectives (2), (3) and (4) can be dealt with here. Misunderstandings and misconceptions perhaps describe best the overall relationship between lawyers and linguists. The two terms do not reveal much about the types and structural interdependencies of the language behavior involved, however. Detailed linguistic descriptive analyses must be made in a thorough systematic study of the "Ethnography of Communication" (see Gumperz and Hymes 1972) to understand how communication does and does not work between linguists and lawyers (cf. Table 9.1: A.1, A.2, A.3).

The point is not only that a lawyer has problems understanding a linguist and vice versa. Court interaction is much more complicated and has several layers of structure. There are at least three classes of "legal experts" (the judge, the defense and the prosecution) and, at least in Germany, three classes of "language experts" (for the court, the defense and the prosecution) in a criminal case, interacting with one another and across their disciplines. "Misunderstanding" is a multifold and rather complex phenomenon by itself, depending on who miscommunicates with whom, about what, when, why and how. At any rate, it is a wide continuum of verbal interaction of various forms on different levels.

It would be necessary to study systematically the relationship between lawyers and linguists from a culture-specific and from a culture-contrastive perspective; for example, between England and Germany, Germany and France. The attitudes of linguists towards lawyers are relevant in various kinds of court proceedings; for example, in disputed statements, press corrections, questions of (similarity of) product names, insults, libel, slander, issues in criminal cases involving author identification of anonymous (extortion, threatening, defamatory) letters, analysis of authenticity of documents, manipulation of texts, plagiarism and others. Each country and culture has a specific overall picture and system of routines in court cases and the ways they are handled. Therefore a culture- and locale-specific set of needs is of great importance. In Germany, linguistic expert testimony is generally requested more frequently in libel and slander cases than it is, relatively speaking, in such cases in England. Linguistic expert testimony is requested much more frequently by the court (the judge) in Germany than by the prosecution or the defense as is the case in England (cf. Kniffka 1996a: 21–50;

see also Introduction to this volume). For this, the needs outlined in Table 9.1: A.1, B.3, C.1 and C.3 are of prime importance.

Linguists' misconceptions of the lawyers' job

One of the most basic misunderstandings seems to be the linguists' lack of knowledge about the needs listed in Table 9.1: A.1, C.1, C.3, and the objectives of lawyers (even those laid down in legal texts). An example is the construction of the so-called "unvoreingenommener/unverbildeter/unbeeinflusster Durchschnittsleser" ('ordinary/average/uninfluenced reader'), which plays an important role in deciding, for example, whether an utterance or a word is to be taken as a "verbal injury" or an insult.

Linguists sometimes simply say that there is no such thing as an "uninfluenced average reader". This is true, but it does not help lawyers. Linguists need to try to understand what it is that leads lawyers to constructions of this kind. It is not helpful at all if linguists simply state that this category does not exist. Instead, they should give clues about what concepts could replace it and what practical consequences this would have (cf. needs listed in Table 9.1 B.3, C.1, C.3).

An even more widespread and basic misunderstanding of linguists in (German) court cases is that they misjudge their role as experts, which is to aid the judge. Sometimes, rather than confining themselves to an expert opinion on linguistic facts, experts take on the role of the judge and give a judicial evaluation of the linguistic data (cf. needs listed in Table 9.1: A.1, B.3, C.1, C.3). In some cases, the linguistic expert even supplies a subsumption under a section of the law. I have seen expert opinions in German insult, libel and slander cases in which a linguist stated "word X is [sic] an insult as stated in Section 185 of the German Penal Code (§185 StGB)". An assertion of this kind is not consistent with linguistic terminology and practice (cf. need listed Table 9.1: A.2). Furthermore, there is no once and for all "insulting status" of any word in any context in a language. A word may have an entirely different denotative and connotative status and range of usage depending on the general (educational, professional, personal) characteristics of the speaker and the ingredients of the linguistic and the situational context. Who says what, to whom, when, why, how, under which conditions and with what intentions may be of more decisive semantic significance than the word (the lexical entry) as such.

A general assertion as quoted above is also contrary to the terms of the law itself. The term "insult" is not defined in the German Penal Code. The judiciary, in this respect, seems wiser than linguists think – and also

wiser than linguistic experts are at times. By not providing a definition, German law tries to account for the fact that there hardly is any utterance or action that can be regarded as insulting *per se*.

A constructive offer in terms of a needs analysis (cf. needs listed in Table 9.1: A.1, A.2, B.1, B.2, B.3, C.1, C.3) would be that linguists' help to lawyers should not be confined to stating *that* the context is of prime importance for the meaning, but would have to consist of a detailed analysis *how and why* elements of the linguistic and situational context are decisive, how they can be analyzed in sound empirical fashion. A linguistic description of the particular case/item in question should be given in a fashion that can be understood by non-linguists.

Another postulate of equal importance is that a linguistic expert opinion presented in court, in oral or written form, should be "clinically clean" from legal terminology, jargon and judicial evaluation (cf. needs listed in Table 9.1: A.2, C.1, C.3). Linguists in such cases are in danger of assuming the judge's role.

Research is also needed to discover why some linguists tend to use judicial evaluation in their expert opinions at all (cf. needs listed in Table 9.1: A.1, A.2, B.3). One reason seems to be that, in our culture, in everyday non-professional interactions, people tend to do this. For example, it is customary in Germany, when discussing daily personal and public events over a beer in a pub, to use conclusive accusations such as "this is fraud", "this is an insult" or "what Mr. X has done is a case of slander". In a linguistic expert opinion in court, everyday conversational style must not be used.

According to German law, it is entirely at the judge's discretion whether or not to request an expert opinion from linguists and whether to consider them experts in the field. Many linguists giving expert testimony are not aware of this (cf. needs listed in Table 9.1: A.1, C.1, C.3). In fact, this is stated explicitly in the German "Gerichtsverfahrens-gesetz", the German law for court proceedings (see Jessnitzer 1978; Sandler 1986) for expert testimony ("Sachverständigenbeweis") in all fields. The law states that the judge is the expert, and that it is she/he alone who decides whether to seek another expert's help. If the judge decides not to like an expert opinion, or does not want to use it, she/he may do with it whatever she/he sees fit. The judge may not use it at all, may not even mention it in the proceedings, may use it to the opposite effect of that which the linguist's expert opinion suggested. The notion underlying the German law term "Würdigung des Sachverständigenbeweises durch das Gericht" (approximately 'evaluation and assessment of an expert opinion by the court') may be hard for linguists

and other experts to understand in all its theoretical and practical implications.

It should be understood and accepted that (1) linguists' evaluations of language data; and (2) the judicial evaluation of what the linguists say about the facts of their reports, are two entirely different matters. But it sometimes is difficult, at least for people without a law degree, to understand that even if linguists provide conclusive evidence, and the judge accepts this evidence as being conclusive, she/he may still decide not to use the evidence as presented, or even not consider it for the verdict.

Based on my own experience, it is less of a problem to understand that judges may use the linguist's evidence from a different perspective, and may put different weight on certain findings in their evaluation (which normally takes a wider view of evidence in the case into consideration than only the language facts). A larger problem rests with cases in which judges develop an analysis of language data of their own, representing a "competing theory", an alternative description and claim of explanatory adequacy than the one that the linguists presented. Obviously, the needs listed in Table 9.1: A.1, A.2, B.3, and, in particular, C.1 and C.3 work both ways. One viable way to ensure that the needs of both are met seems to be for the judiciary to offer courses for applied linguists about the nature and status of the expert proof ("Sachverständigenbeweis"), and, conversely, for the linguists to offer courses for the judiciary explaining the ingredients of an empirical analysis of data of language and language behavior. In fact, all needs listed above have to be accounted for with further additions concerning culture-specific data of everyday life, including cross-cultural compatibilities and incompatibilities.

If such a step is taken, particular reference should be made to the attempts that each science has made in describing and explaining facts of the other science (cf. needs listed in Table 9.1: A.1 and C.1, which are of critical importance here). There are, for example, almost ritualized and stereotypical judicial references to "context" as the decisive category for understanding an utterance. It is not surprising at all that lawyers have used the term "context" in a magical, non-operationalized way, if linguistics – the science that ought to be able to give a satisfactory explanation of (1) what "context" means, and (2) how it can be verified and operationalized – has failed to do so. Dealing with pragmatic and sociolinguistic matters scientifically, securing an empirical analysis of language behavior that really deserves the name, has been a standard objective of linguistic inquiry for only a few decades. To be sure, linguists also have to be taught what lawyers have done with and written

about “context” and why they do what they do. Lawyers have to be taught that there is a large amount of linguistic research on context and contextual meaning, why this is relevant for their work and how it can be operationalized to serve the purposes of law. Even more basic distinctions such as that of linguistic context and situational context have to be observed.

In short, deficits are as large on the linguists’ side concerning legal issues as they are on the lawyers’ side concerning linguistic facts. Elementary lack of information about the other science, in a non-trivial sense, is the predominant characteristic (cf. needs listed in Table 9.1: A.1, C.1). Linguists, above all, lack information about the science of law and about what judges do and want. Judges lack information about the opportunities and endeavors of linguistics as a forensic auxiliary science and about what linguistic expert testimony can and cannot do.

Lawyers’ misconceptions of linguists’ work

Several examples of linguists’ misunderstandings were given on p. 240ff., where “corresponding” misunderstandings of lawyers shine through. In this section, some additional comments and examples will be pointed out briefly (a more detailed discussion and a larger number of examples are given in Kniffka 1981; 1994).

As far as lawyers’ practical work with the notion of “*der unbeeinflusste Durchschnittsleser*” (“the uninfluenced average reader”) is concerned, the main problem is not that the judiciary works with fictions of this kind, but rather how it operationalizes the concept (cf. needs listed in Table 9.1: A.1, B.1, B.2). That is, how the data of language use are being analyzed and used as criteria for judicial judgments, and in which way the data of language use are considered to be available to the judge. The “uninfluenced average judge”, one might say, argues that: “The ‘uninfluenced average reader’ is me. How do I understand the utterance in question?”

In such cases, judges are in some way confusing their own competence as native speakers of their mother tongue with the result of a systematic linguistic analysis using intersubjectively valid criteria and methods of scientific analysis. In principle, there is no difference between judges’ (linguistic laymen’s) statements on language data and their statements, for example, on forensic medicine or chemistry. In deciding whether a word “is an insult” or “is not an insult”, judges would do the same as if they relied on self-experiments and introspection in cases of forensic medical or chemical expert testimony. They adopt the same misunderstanding that is customary with students in introductory linguistics

classes. It always takes them some time to understand that linguistics is a descriptive, not a prescriptive science and that linguistic analysis means eliciting data from native speakers and making their internalized (mainly unconscious) knowledge of language explicit.

Some judges think that linguistic experts should be consulted because of their “personal authority” in language matters; that is, because they are more competent native speakers than the judges consider themselves, not because linguists are trained in scientific analysis in basically the same sense other forensic experts are trained in their fields. This has to do with the nature of language and with the status of linguistics as a science, and also with the role that language plays in the judiciary.

Almost all judges I have dealt with were not aware that self-assessment (“auto-stereotype”) of (one’s own) language behavior, assessment of one’s language behavior by others (“hetero-stereotype”), and observable language behavior (as an object of systematic linguistic analysis) are three different matters. Linguists have also ignored this basic fact for decades, before sociolinguistics pointed out its importance for the description of language use.

It is difficult to convince a judge that we all, including the judge and the linguist as everyday language users, are not consciously aware of what we say and how. And even if they were aware of this, they would have only imperfect ways to describe it without training in linguistics. It also is difficult to convince a judge that most people do not know how language functions, even though they communicate easily every day. It is a prevalent task of linguists, forensic and general, to pass along the word that a major portion of the linguist’s job is the explicit description of the regularities and rules of the system and the use of language internalized by native speakers. Furthermore, linguists should inform all interested parties that this process presupposes a scientifically sound description and explanation of how language works as (1) a system in itself, and (2) a system of human communication. It is not enough to “see” or “feel” that a word uttered in a certain context is an insult or not, whether a word or an utterance stems from speaker A and not B.

Lawyers frequently claim that forensic linguistics (sometimes including handwriting analysis and phonetics), unlike forensic psychiatry, psychology and medicine, does not yet have “a fully established and standardized method” (this has been a stereotypical statement in German courts and articles in law journals by non-linguists). Lawyers use this to advise against consulting a linguist or are afraid that a linguistic analysis might harm their client’s case in one way or another. This is a misunderstanding that must be challenged. It has to be taken seriously, considering the

effects it may have in actual court cases. This perception by lawyers is both unjustified and unscientific (cf. needs listed in Table 9.1: A.3, B.1, B.2, B.3).

The fact that forensic linguistics is a comparatively young science does not necessarily imply that it is unable to supply results based on solid empirical analysis. Moreover, it is unjustified to state any systematic difference in this respect between forensic linguistics and other forensic sciences such as forensic psychiatry, psychology and even DNA-analysis (after all, differences in data and data measuring are taken into account). Each forensic science has its own problems of validation of results and refinement of methods. There is no doubt that data of forensic linguistics are less hard than, for example, data of forensic chemistry. The quest for an interpretation of the findings is basically the same, however, in all sciences. It is not the field as such, but the question of how the field is practiced by whom and in what way that determines the reliability and validity of the results.

Another widespread misunderstanding of linguistics and the role of the forensic linguistic expert concerns the general accessibility and intelligibility of scientific data. Judges seem, in this respect, no different from the rest of the population. There are German judges who believe that author identification is an area in which the judge, *ipso facto*, needs the linguistic expert's help, more so than in the area of disputed utterances, insults, libel and slander. There is nothing to support this assumption from a linguistic point of view, however. It seems to reflect the general attitude of the public towards scientific data available in a non-formal versus a formal, a nomenclatured versus a non-nomenclatured format. Many judges seem to believe that they are as able to judge objects of psychology as a forensic psychologist, because they have enough "experience of life". The same holds with language.

These findings point in the same direction as those of the previous section. Lawyers and judges should be offered courses by linguists about linguistic basic research, methodology of empirical data analysis, aims, goals and limits of applied linguists who provide expert testimony in court.

Excursus: Lawyers' view of the role of language

A brief example from an authentic case is given below to illustrate: (1) the importance of questions of language structure and language use in court proceedings; (2) the extent to which lawyers get into, are occupied with, and indulge in questions of linguistics proper; and (3) how necessary linguistic help is in dealing with such questions. The linguistic data in

question concern such things as verbal behavior and assessed verbal behavior, “language attitudes”-data and opinions held about language. The data show that a grammatical perspective alone is inadequate and has to be extended to include a sociolinguistic (ethnographic) dimension of the description and explanation (cf. needs listed in Table 9.1: A.1, A.2, A.3).

The case centered on an anonymous defamation letter. The two defendants worked for the same company. The plan of the plaintiff (the company) was to exonerate one of the two potential authors in order to charge the other. The attorneys of the plaintiff argued, among other claims, that the German word *Konzeption* (approximately ‘concept’, ‘set-up’, ‘design’, ‘structure’) would and, in fact, *could* not be used by a cleaning lady, as with several other words and expressions in German, due to her “lack of linguistic expertise and competence because of her poor schooling and due to the fact that she had been born and raised in Strasbourg” [sic]. In excluding her as an author, the company tried to identify *ex negativo* the other defendant (with an academic training in economics) as author of the anonymous letters. This would have enabled them to dismiss him from his job as a member of the company’s council. He admitted to have typed (but not (co-)authored) the handwritten notes of his fiancée, the cleaning lady, on the typewriter in his office, after a carbon paper used for writing the letter had been found in his waste paper basket by the police.

His attorneys argued in his favor that the word *Konzept*, as any other word of any language, “could even be taught to a parrot” (original quote), let alone a cleaning lady, and that she could very well use it at her discretion any way she liked, even if she would not fully understand what it meant.

It is worth looking at the original wording of the statement made by the lawyers when assessing the cleaning lady’s language capacity, as given in Table 9.2 below.

Here, the lawyers for the plaintiff pass judgment on what the cleaning lady can and cannot do linguistically; that is, to write, formulate or produce certain words and phrases. In addition, they explicitly point out the language material that she supposedly could not have authored. They also give explicit reasons why she could not do so (lack of formal education, language skills, her upbringing across the border in Strasbourg and so on). From a linguistic perspective, this is not valid. The lawyers of the defendant(s) are by no means less creative in their own statements. They take pains to argue that “even a parrot” could be taught to produce certain words and phrases including those in question.

Table 9.2 Lawyers' assessment of language proficiency of a cleaning lady (original wording of statement sent to the court)²

Frau ____ wäre aufgrund ihrer sprachlichen Gewandtheit nicht in der Lage, ein derartiges Schreiben zu verfassen. Dazu fehlt Frau ____ die notwendige Vorbildung und sprachliche Gewandtheit. Da Frau ____ Elsässerin ist, ergeben sich diese Schwierigkeiten allein schon aus dieser Tatsache. Frau ____ wäre zu folgenden Formulierungen nicht in der Lage:

1. *“In unserer Konzeption steht wörtlich:”*
2. *“Die Betreuungssituation widerspricht ebenfalls der Konzeption des ...”*
3. *“Weiter möchte ich auf die m.E. unzureichenden sanitären Anlagen aufmerksam machen.”*
4. *“Da in den Zimmern Rauchverbot besteht, ist die Situation in der Küche für die Nichtraucher unerträglich.”*
5. *“Ebenso betreibt m.E. Herr ____ eine menschenunwürdige Personalpolitik, auf die ich hier nicht mehr eingehen möchte.”*

English translation (word by word):

‘Ms. ____ would, on the basis of her language skills, not be able to write a letter of this kind. To do that, Ms. ____ lacks the necessary education and language skills. Since Ms. ____ is from the Alsace, these difficulties arise simply from this very fact. Ms. ____ would not be capable of the following formulations/wordings:

1. *“In our conception stands literally:”*
 2. *“The reality of the care situation also contradicts the conception of ...”*
 3. *“Furthermore I would like to bring to attention the (in my opinion) unsatisfactory sanitary installations”*
 4. *“Since it is forbidden to smoke in the rooms, the situation in the kitchen is unbearable for non-smokers”*
 5. *“Also Mr. ____ , in my opinion, is practicing a degrading personnel management upon which I don’t wish to comment further here.””*
-

The linguistic matters debated here contain enough material for several linguistic dissertations on “opinions held about language” – if the questions are rephrased in a sound scientific way.

The above is quoted here to illustrate the extent to which lawyers engage in, and have to engage in, the analysis of language data – and actually perform the linguist’s jobs here, which probably is somewhat surprising to the latter (cf. needs listed in Table 9.1: A.1, A.2, B.3, C.1).

This is, by far, not all of the “linguistics” done by non-linguists. In the case reviewed, the court (a non-academically trained employee on behalf of the judge) interrogated the cleaning lady as a witness, trying (1) to elaborate on the actual production process of the letter; that is, what exactly the joint activity of the two people was; and (2) to determine

the lady's command of language more precisely; that is, to find out experimentally whether she knew/was able to produce the word *Konzeption*, understood it and was able to use, for example, a German compound noun *Heimaufsicht* – which had not come to her mind in a previous interrogation and thus was suspected not to “stem from her” and not be part of her repertoire.

It is also worth looking at the minutes taken by the court clerk of this part of the interrogation, a selection of which is given in Table 9.3. The reason for quoting the judiciary's “experimental analysis” of language matters is not to delight in seeing non-linguists' performance of linguists' jobs. Rather, it is to document the enormous quantity of questions about language, language use and command of language that, in one way or another, are handled by the judiciary and lawyers – and the urgent need to get linguists involved in this type of work (cf. needs listed in Table 9.1: A.1, A.2, A.3, B.1, B.2, B.3, C.1), which is of critical importance for the analysis and judicial evaluation.

The last quoted remarks of the cleaning lady contain sufficient information about the validity of the data gained in this experiment. They are invalid and do not yield anything for the question concerning whether the German lexical item *Konzeption* and the compound noun *Heimaufsicht* are at the cleaning lady's disposal and in her repertoire, let alone whether they were at her disposal at the moment when she was interviewed and/or when she wrote the letter. The lexicon of speakers and their “compounding” capacity cannot be determined in this fashion. Even if it has any linguistic significance at all, it would certainly be of no avail for the question of authorship of the anonymous letter.

In short, it takes a law degree to understand and handle the legal aspects of a case. It takes a linguistics degree to understand and handle the linguistic idiolectal repertoire of a person and to analyze and assess the significance it may have for authorship analysis of anonymous texts (cf. needs in Table 9.1: A.1, A.2).

Called in as a linguistic expert by the judge, I argued that the assumptions made by the parties involved were not based on any solid linguistic grounds and that it is not possible to “exclude” a speaker of a language for using (and understanding) a certain word of one kind or another. In this case, the contention of the plaintiff's attorneys was in fact counter-productive because the cleaning lady's use of the German word *Konzeption* could very well have been fostered rather than hampered or even suggested by the French word *concept* to which she had been exposed during her upbringing in Strasbourg.

Table 9.3 Official court transcript of taped notes of witness interrogation on command of language

Danach fragt³ das Gericht die Zeugin, was sie unter dem Begriff *Konzeption* verstehe. Die Zeugin erklärt:

“In diesem Falle verstehe ich unter *Konzeption* die Erklärungen, wie sie in einem Prospekt der Gesellschaft für ... im einzelnen enthalten sind.

Unter *Konzeption* verstehe ich weiterhin, wie man ... hilft durch entsprechende Maßnahmen.”

Das Gericht fragt danach die Zeugin, wie man die Aufsicht über Wohnheime bezeichnet. Zunächst erklärt die Zeugin, sie könne hierzu nichts sagen, sie wisse es nicht. Nach einigem Überlegen sagt sie dann, es handle sich um die *Heimaufsicht*.

[Details of the letter production process follow]

Nach dem Vorspielen der Zeugenaussage erklärt die Zeugin ergänzend:

“Ich bin vorhin nach dem Begriff der *Heimaufsicht* gefragt worden. Ich möchte hinzufügen, daß ich vorher auch jetzt noch sehr aufgeregt bin und mir aus diesem Grund der Begriff nicht gleich eingefallen ist.

Den Begriff *Konzeption* habe ich aus dem bereits erwähnten Prospekt entnommen. Dieses Prospekt war mir bekannt.”

Informal English translation (abbreviated):

‘Then the court asks the witness what *Konzeption* means to her. She answers:

“In this case *Konzeption* means to me the explanations as given in a brochure of the company. Furthermore, I understand by it how one helps ... by appropriate measures.”

The court then asks the witness what the term for the administration of ... is called. First she says she did not know. Then after some deliberation she says it was *die Heimaufsicht*.’

[Details of the letter production process follow]

‘After the tape has been played to the witness she adds:

“I have been asked about the term *Heimaufsicht* before. I would like to add that I was and still am now very nervous and therefore could not think of the term. The term *Konzeption* I know from the brochure of the company.”

Note: All names and identifying words have been omitted.

In addition, the fact that a person could be trained to repeat and use words “like a parrot” supplies little or no conclusive evidence about co-authorship. Co-authorship of a text would certainly imply more than repeating words one does not fully understand. “Parrot-like” use of words and phrases, with no understanding or a variable degree of understanding of what is said – a question to be studied in its own right and not very easy to answer, as recent research suggests – has nothing to do with the question of authorship at stake in this case.

A linguist using lawyers’ vernacular and judges using linguists’ jargon, quoting, for instance, a statement from a linguist’s expert opinion in a verdict, would not touch the question of authorship of the text under investigation. The lawyers of both sides and the court each had several misapprehensions and misconceptions of basic facts of language system, language use and idiolectal repertoires of speakers, which needed to be clarified by a linguist (cf. needs listed in Table 9.1: A.1, A.2, A.3, B.1, B.2).⁴ The question that was the focus of this case – is there *multiple authorship* or *single authorship* – could not be answered by the linguist consulted. The identification of a particular author and the determination of a specific “degree of authorship” in cases of multiple authorship are, as a rule, the most difficult jobs in authorship attribution. Not infrequently, they are impossible tasks for the forensic linguist. In other words: “Who wrote what” is a question that, in most cases, only the people involved can answer reliably (see Kniffka 1990d; 1994).

This represents yet another misunderstanding of linguistic facts by the judiciary, which I have experienced many times when giving expert testimony in/for German courts (cf. needs listed in Table 9.1: A.1, B.1, C.1). Judges feel at ease acting as language experts themselves; for example, in libel and slander cases (which, at times, take a substantial effort of linguistic-pragmatic analysis and much professional experience). They do this to a lesser extent in cases of authorship attribution, however. Moreover, judges tend to think that, for example, cases of disguised or pretended authorship are more difficult to analyze, whereas cases of multiple authorship are not. Based on my own experience, as a rule, the reverse is actually true (cf. Kniffka 1994).

The example illustrates that teaching introductory courses in linguistics to lawyers and judges, and teaching introductory courses in law to linguists is, undoubtedly, an urgent and useful endeavor. Detailed real life reports and case studies (relating to topics of all needs listed above) should be given in classes administered by the most experienced applied linguists and forensic linguistic experts and the most experienced lawyers and judges.

Problems in forensic interaction: “tricks of the trade”

The discussion so far concerns matters of (lack of) information, misunderstandings and so on, between linguists and lawyers, commenting occasionally on disputed matters between them. There is yet another dimension in the need for cooperation or strategic interaction between linguistic court experts and lawyers (including, to a degree, judges). This concerns the tricks of the trade employed by forensic linguistic experts, or “experts”, and by lawyers trying to win a case for their clients rather than applying the standards of science in general and of the science of “the opponent’s expert” in particular; in this case, linguistics. It would be an unjustified (and counterfactual) oversimplification to say, even from a layman’s perspective, that the matters referred to on pp. 237–52 pertain to cases in which “everything is in order”, with all sides involved; whereas in cases reported here “not everything is in order”, or something is wrong, against the rules and professional ethics.

There is no simple dichotomy between proper and improper performance of interactants, clean and unclean methods in court proceedings. Neither is there a dichotomy within each of the two classes, but rather a continuum of many shades and degrees from “professional” to “non-professional” conduct, from “good” to “bad” performance, from “excellent” to “very poor” practice and work in terms of one’s job as a linguistic expert. Linguistic court experts, generally speaking, are in a much weaker position compared with all others involved, due to the fact that all others have a law degree and a good knowledge of legal matters and proceedings. For example, linguists on their own would probably not know how to evaluate a lawyer’s performance with regard to the continua mentioned and in terms of the “tricks of the (law) trade”. They have to rely entirely on the help of lawyers. Forensic linguistic experts, as a rule, learn about the tricks of the trade of the legal professions “the hard way” – frequently after they have been the object thereof or when it is too late.⁵ Without a law degree, I cannot give a more general theoretical discussion here. I can merely report on an authentic case in which I was the linguistic expert – and the professional “victim” of a lawyer’s tricks of the trade. It should be mentioned that this was, in fact, the only case of the kind I have experienced to date. Reporting it here is not meant to assess any ill-doing or illegal practices used by the lawyers and the judge involved. As said above, I am not competent to judge this. This case is reported to illustrate the forensic linguist’s (that is, my own) ignorance of the tricks of the trade that lawyers use and to help others become better prepared for such in

their current and future cases (cf. needs listed in Table 9.1: C.1 and particularly C.2).

This case further illustrates that linguistic experts, even if they are doing a good job in their testimony, can fail by not knowing what lawyers have at their disposal. That is, they may be tricked out of giving expert testimony at all by the opposing lawyers' concern that such testimony might harm their client's case.

The case involved anonymous libel and defamation letters about a company. One of its employees, who had written non-anonymous letters of complaint to his superiors in the past, was suspected to be the author. The company requested and received a "linguistic" expert opinion (made by a marketing research institute, which did not employ linguists and had never given forensic linguistic expert testimony before). The non-professional testimony came to the conclusion that the employee's letters and the anonymous letters originated from the same author. As a result, the company dismissed the employee, without notice. He filed suit against this at a labor court, claiming that it was unjustified since he had not written the anonymous letters.

The labor court requested its own linguistic expert testimony, asking for a preliminary statement whether or not it was possible to make reliable statements about the anonymous authorship, and the percentage of likelihood of the assumption that the employee had written the incriminating letters.

In the court file sent to me there were several extensive statements written by the lawyer on the part of the employee. He was concerned that my testimony would come to the same result as that of the previous testimony, and so he took pains to make strong statements of an unscientific and, at the same time, highly derogatory nature about forensic linguistics in general, and the BKA's and my own work in forensic linguistic authorship attribution in particular. In his letters to the court, he quoted "authorities" he had contacted over the telephone who had supposedly supported his own low view of forensic linguistics. One of these was a colleague in (general) linguistics who had never seen, let alone written, a forensic linguistic expert opinion. He declared that forensic linguistics was not a successful field. A historical linguist specializing in ancient Indo-European languages was also quoted as giving an "expert" opinion on the "do's and don'ts" of the field, the chances of success of author identification and the nature of data required.

The lawyer summarized that forensic linguistics was a science "that has to fight hard for its academic survival" (original quote, in 1993) and

that I should be forced to give proof whether any of my testimonies had ever been successful before I was admitted to the court. He also stated that it was simply impossible, once and for all, to say anything about anonymous authorship, the question of similarities and discrepancies between the anonymous and the comparison data.

The legal representation of the other side, the company, held equally strong opinions about the role and scientific potential of forensic linguistics, not surprisingly in complete opposition to the opinion of the employee's lawyer. Since the judge's letter requesting expert testimony also contained various counterfactual assumptions about forensic linguistics and its scientific explanatory and argumentative potential for authorship attribution, I felt compelled to write to the judge explaining in some detail what could and could not be done from my point of view. I also argued that the results of an expert opinion could not be known beforehand, that no simple overall assessments were possible or legitimate, and that the majority of statements and assumptions put forth by the lawyers of both parties were counterfactual and unscientific.

This was exactly what the lawyer had wanted me to do. In describing that his statements were unjustified and unscientific – in a scientifically sound and factual fashion – I had proven myself as not being an impartial, unbiased court expert in his eyes. He then filed a complaint at the Superior Labor Court to declare me not impartial but biased (“Besorgnis der Befangenheit des Sachverständigen”; “Misstrauen gegen seine Unparteilichkeit”). The court accepted the complaint, and my testimony was not admitted as evidence; though, had the lawyer known that my testimony came to the opposite conclusion of the previous non-professional testimony, he would probably have given his complaint more consideration.

A detailed analysis of the linguistic statements in the letters of the lawyers, my own letter, and the court decision, and the way the judge interpreted the expressions would be worthy of study in greater detail, but this is beyond the scope of this chapter. He did not seem to be comfortable dealing with questions of language and language use.

The crucial points, in terms of “tricks of the trade” and the need for the linguistic expert to account for them, are the ignorance of (1) how to react or *not* to react to provocations by a lawyer; and (2) a distinction to be taken into consideration. It is not an issue here whether the court declares the linguist to be biased, or whether the linguist is in fact biased. The question only is whether one of the parties involved can have a subjective impression toward that end. Answering my letter of

inquiry, the presiding judge elaborated on this even more clearly (original quote):

“Ich möchte aber auf diesem Wege nochmals darauf hinweisen, daß es nicht darum ging, Sie wegen Befangenheit abzulehnen; vielmehr war zu prüfen, ob eine entsprechende Besorgnis der Partei aus deren subjektiver Sicht nachempfunden werden konnte.”

English translation:

‘I want to express to you again here that the court did not reject you on grounds of a bias. It had to decide whether a subjective impression of the party to this effect was justified or not.’

A court expert, according to the court order, is not entitled to state that the lawyer’s argumentation of linguistic matters is unscientific, that the lawyer has a basic information deficit of the field (original quote from the court order):

“Zu all diesen Stellungnahmen ist ein Sachverständiger nicht berufen, auch wenn er noch so sehr recht mit ihnen haben sollte.”

English translation:

‘The court expert is not entitled to such comments, even if he is absolutely right.’

The point here is addressed to prospective linguistic court experts in Germany. One should not write anything to anybody about lawyers’ writings, statements or arguments, no matter how strongly, persistently and violently one gets personally attacked, provoked or discredited. If you do react, you prove yourself biased – from the perspective of the lawyer (cf. needs listed Table 9.1: C.2). It is considered an absolutely legitimate part of the strategic plan of the lawyer to claim that the discipline of the expert is “basically questioned and discredited in its stringency”. It is not considered legitimate for the expert to engage in a rebuttal – no matter how justified this may be from a scientific perspective.

Intrinsic deficits of linguistic basic research and interdisciplinary projection⁶

As stated above, there are domains of linguistic research that have not been developed to an extent that their relevance for forensic linguistic expert testimony and for general concerns of the judiciary has been

adequately described. In some cases – for example, German orthography, lexicon and grammar – the descriptive data are “clear” as scientific (linguistic) data, but not “clear” as data themselves. They are, by their very nature, “fuzzy” data of a rather unclear transitional status. This quality of language data is usually not realized by lawyers and thus must be seen as part of their misunderstandings or misconceptions of the facts of language and language use.

This must also be listed under linguists’ misconceptions, however, since linguists have largely failed to supply lawyers with the proper view and assessment of the data and the argumentative potential in court. In fact, linguists have frequently been unsuccessful in convincing lawyers that they know what they are talking about and that their analysis of data is sound and executed *lege artis* (cf. needs listed in Table 9.1: A.1, A.2, B.1, B.2, C.1). Lawyers like clear-cut quantifiable language data. They request the statement of specific percentages – for example, of likelihood of/in authorship of anonymous texts – and blame the linguists and their poor professional competence if they do not come up with such percentages, but rather state explicitly that this would contradict the facts of language and, in addition, would be unprofessional and unscientific. Here, linguists are safe in what they are doing, and what they say about it, though substantial desiderata and needs of empirical research can be seen in many areas (see below). In several other domains of data analysis, a genuine lack of linguistic basic research proper and applied research seems prevalent (cf. needs listed in Table 9.1: A.2, B.1, B.2). Here, some linguists do not even seem to have realized that huge research deficits and desiderata exist, what they look like and how they can be dealt with most efficiently. In this section, a symptomatic example of each will be discussed briefly.

There is a widespread misunderstanding among lawyers and laymen in linguistics, particularly people who frequently have to deal with norms and standards of language, such as secretaries, clerical staff and school-teachers. They think that there is always a clear statement of a norm or “the norm” of language use, and that the degree of deviation from it is obvious and easy to analyze. A domain enjoying public controversy *par excellence* is German orthography (that is, long before any spelling reform was in sight). Domains such as vocabulary, grammar (morphology and syntax) and stylistics could be named equally well. Real life teaches us, however, that there are many instances in which it is not possible to state a clear-cut norm, or in which the norm itself is fuzzy rather than clear.

If, in a debate on spelling in Germany, somebody is able to support a version by a quote from the “Rechtschreib-DUDEN”, the case is usually

decided. The “Rechtschreib-DUDEN” is something like the “orthography bible” of laymen. Whatever is in the “bible” is right. Whatever is not in the “bible” is wrong, does not exist, and is not used. Many non-linguistic people go by this authority of the DUDEN; many, such as schoolteachers, secretaries and judges, feel they have to follow its authority. Confused by which other standard they should follow, people forget that the DUDEN is man-made, that it itself does not give clear norms in all cases of spelling and that the norms of usage as stated in the DUDEN may be worlds apart from actual usage; for example, spelling practice of fragments of the German population.

One example is illustrative: hyphenation in expressions with numbers in German compounds such as *5-Jahresvertrag* (which was, in fact, a critical orthographic item in an expert testimony on anonymous authorship attribution some 15 years ago). The “old” Rechtschreib-DUDEN (before 1996), p. 66 under R212 (rule no.212), states:

“Ableitungen, die eine Zahl enthalten, werden zusammengeschrieben, unabhängig davon, ob die Zahl in Buchstaben oder in Ziffern geschrieben wird. Das gilt auch für Zusammensetzungen (vgl. auch R 43).

achtfach, 8fach, Achtpfänder, 8pfänder ... ver307fachen, 80er Jahre, Dreikant[stahl], 3kant[stahl], Elfmeter[marke] ... aber bei Aneinanderreihung: 400-m-Lauf, 2-kg-Dose, 3/8-Takt”

English translation:

‘Derivations containing a number are written in one word, no matter whether the number is written in letters or in numerals. This applies to compounds also: [Examples]’

A definition of “Aneinanderreihung” (‘concatenation’) is not provided here, and it is not apparent whether this instruction is clear in itself or not. Independent of this issue, the rule does not help the man in the street much when uncertain how to write a word such as *Fünfjahresvertrag/fünf-Jahresvertrag/Fünf-Jahres-Vertrag/Fünf-jahres-vertrag/5-Jahresvertrag/5-Jahres-Vertrag/5-jahres-Vertrag/5-jahres-vertrag ...*

Interestingly, the “new” Rechtschreib-DUDEN (1996: 36) has a different reading:

“Ableitungen und Zusammensetzungen, die eine Zahl enthalten, werden zusammengeschrieben: ... *achtfach, achtmal, Achtpfänder, die Achtziger.*

Wird die Zahl in Ziffern geschrieben, setzt man bei Zusammensetzungen einen Bindestrich: ... *8-mal, 8-Pfänder, 8-silbig*.

Aber bei Ableitungen ... : *8fach, 17tel, 80er, 32stel*.

Bei Zusammensetzungen mit Ableitungen: *in den 90er-Jahren* (auch: *90er Jahren*) [sic], *auf ein 1000stel-Gramm genau*.

Bei Aneinanderreihungen werden Bindestriche gesetzt ... : *4000-m-Lauf, 2-kg-Dose ...* "

English translation:

'Derivations and compounds containing a number are written in one word:

[Examples]

If the number is written in numerals, a hyphen is used in compounds:

[Examples]

But differently in derivations: [Examples]

In compounds consisting of derivations: [Examples]

In concatenations hyphens are used: [Examples]'

In the same edition (DUDEN 1996: 880, §§42–45), there is an attempt to define "Aneinanderreihung" ('concatenation'): "§43: Man setzt Bindestriche in substantivisch gebrauchten Zusammensetzungen (Aneinanderreihungen) ..." [sic] ('Hyphens are used in compounds (concatenations) of nouns.'). Also, rules for the use of hyphens are spelled out: "Man setzt einen Bindestrich zwischen allen Bestandteilen mehrteiliger Zusammensetzungen, in denen eine Wortgruppe oder eine Zusammensetzung mit Bindestrich [sic] auftritt" ('A hyphen is used between all constituents of complex compounds in which a word-group or a compound with a hyphen occurs.'). The examples given, *800-Jahr-Feier* ('800-year-anniversary'), *35-Stunden-Woche* ('35-hour-week'), *8-Zylinder-Motor* ('8-cylinder-engine'), could suggest that §44 *5-Jahres-Vertrag* is correct – but what if one thinks that words should be used instead of numbers?

This example illustrates that even a (supposedly) professional linguistic account of German orthography is, in fact, anything but clear. It does not provide a norm – let alone a clear-cut norm – and a recommendation of how to write.

Against this background, it is not surprising that the actual orthographic behavior of German native writers is even more varied, unclear and fuzzy.

In some experimental testing with several classes of German students of linguistics, a maximum variation of spelling behavior was found even within this population. Table 9.4 contains the entries of just three subjects to illustrate this. To distort the subjects' attention, two French words were put into the sentence, pretending to find out whether the students were familiar with these French words (German students of linguistics and foreign languages are familiar with this type of test). The results were as follows. There seems to be a substantial random variation in the use of hyphens in the words tested as standard spelling behavior of German university students of linguistics and, very probably, for larger populations of German speakers as well. Several students showed inconsistent usage of the hyphen in the two counts of identical words in the test. This is what real life German spelling behavior looks like. Linguists and lawyers will need to take this into account.

Linguists must realize and pass the word along to other disciplines that, in many cases language data are not by any means clear. Not infrequently, they look the exact opposite. Analyzing language use in real life means dealing with a fuzzy, ever changing, ambiguous object – very bitter indeed to people who much prefer to count well defined tokens (and base verdicts upon it).

Table 9.4 Fuzzy norms: Data on hyphenization in German in the word 5-Jahresvertrag

(Original answers of test takers to illustrate idiolectal variation)

Female student, 20 years, from NRW:

*Ein fünf-Jahres-Vertrag en block ist mir lieber als zwei oder mehr 3-Jahres-Verträge
peu à peu.*

Female student, 22 years, from Böbingen/Pfalz:

*Ein 5-jahres Vertrag en Block ist mir lieber als 2 oder mehr 3-jahres Verträge
peu à peu.*

Male student, 24 years, from Southern Germany:

*Ein Fünfjahresvertrag en bloc ist mir lieber als zwei oder mehr drei Jahresverträge
peu a peu.*

English translation:

'I'd rather have a five-year-contract en bloc than two or more three-year-contracts peu à peu.'

Lawyers need to understand that (1) unclear, fuzzy data of this type do not mean that language – or the science of linguistics for that matter – is deficient; and (2) that a linguist stating them in court is telling the truth and nothing but the truth.

The second example differs from the above in that it is unclear, not only to non-linguists but also to linguists as well. It concerns lexicon and word formation; that is, the formation and use of compound adjectives in German legal jargon or terminology. Although the fact, as such, is well known in German “*Fachsprachenforschung*” – the “*Fachsprache*” has and uses compounds where the “*Umgangssprache*” uses simple (uncompounded, briefer, less terminological) terms. A full and updated description of the data is not in sight, however, let alone there being a satisfactory description and explanation of this ongoing “linguistic change in progress”. The following remarks point out, only briefly and programmatically, the direction in which such research would have to go.

The first time many Germans – especially university professors, including linguists in the state of North Rhine-Westphalia – came across the compound adjective *lebensälter* was probably when they read or heard of an official letter/decreed issued by the Ministerium für Wissenschaft und Forschung des Landes Nordrhein-Westfalen on 21 January 1993.

The word *lebensälter* occurs in the text four times, as an attributive adjective to nouns denominating persons – (1) *Berufung lebensälterer Professorinnen und Professoren* in the subject heading above the text, (2) *lebensälterer Bewerber für das Amt eines Professors* in the first sentence and paragraph, (3) *Berufung lebensälterer Bewerber* in the second paragraph, and (4) *lebensältere Professoren aus Nordrhein-Westfalen* in the last paragraph. Twice, syntactic paraphrases of the adjective are given in the text: (5) *weil sie in einem höheren Lebensalter stehen*; (6) *steht ein höheres Lebensalter (in der Regel das vollendete 52. Lebensjahr)*.

These can be used as hints as to what *lebensälter* means. It “does not exist” in the German language in general (that is, it is not listed in the DUDEN) and it sounds somewhat strange, to the man in the street. It means something similar to ‘people of a higher age’, actually referring to applicants for a professor’s tenured position in their late forties and early fifties who can be hired and become “*Beamte*” under certain conditions discussed in the text. The text also gives clues as to how this word came into being, how it is to be explained in its formation. The compound adjective *lebensälter* renders the phrase “*ein höheres Lebensalter habend, Bewerber von/mit höherem Lebensalter*” (‘(applicants) of a higher age’) as occurring with the (well-established) noun “*Lebensalter*” in (5) and

(6). It is also easily plausible why it was coined and how it is motivated. It is a short(er) form of a prepositional or other (longer) phrase – an “abbreviatory univerbation” almost – in line with the general attempts at brevity and precision in administration/legal terminology and jargon. It is not unlikely that the adjective *lebensälter* existing in the comparative form is only formed by analogy to the adjective *dienstälter*, meaning something similar to ‘of longer service, having occupational enciennity’ (in a company, or government job). Neither of the “positive forms” (**lebensalt*, **dienstalt*) exist in German to date. So, the analogy pattern would be *lebensälter*: “höheres Lebensalter (habend)” as with *dienstälter*: “höheres Dienstalter (habend)”.

The first point to be made here is that linguists actually need to take this kind of data into consideration and make detailed text linguistic analyses of texts such as the one analyzed above. Who, if not linguists, must do this service for the general public, explaining what this “odd” word means and how and why it was coined? Maybe they could assist lawyers with advice on the benefits and dangers of new word formations.

The second point is that the description of ongoing “linguistic change in progress”, as outlined by Labov (1966), is an important objective of modern linguistics for its own sake, and can be of great relevance for the interdisciplinary dialogue with legal personnel. Furthermore, the data of legal terminology discussed prove to be an almost ideal field to exemplify how linguistic change in progress works. There is no reason why this should remain a field neglected by linguistic scholarly study. Rather, it gives evidence that much more linguistic basic and applied research is needed concerning the structure and status of “new” terminological items and the processes that lead to them. It is also of great importance to make the results available to legal personnel since they may help to make law texts somewhat easier for lay people to comprehend.

The word *lebensälter* was not created in the Ministry’s letter/decreed. The authors in the Ministry of Science of NRW kindly informed me that they did not create it in this text but that it “had been around for some time”. After a systematic inquiry of the history of this word as an item of law terminology, the first officially documented written accounts of the word *lebensälter* that I was able to find are four verdicts of the German Supreme Labor Court (Bundesarbeitsgericht, BAG) of 1986 (two cases), 1990 and 1991, in which the word occurs in the general meaning of ‘having a higher age’, as represented in Table 9.5. The judicial aspects of the cases quoted cannot be discussed at all. Even the linguistic data cannot be described in more detail here. The only aim here is to exemplify early occurrences of the word *lebensälter* and its uses in law texts. Only the immediate linguistic context is quoted.

Table 9.5 Early occurrences of the word *lebensälter*

(a) BAG 7. Senat; Datum: 1986–01–29; Az: 7 AZR 259/84

Orientation: The case was about an employee whose working contract had been terminated by his employer unlawfully.

“Der langjährig beschäftigte und damit in der Regel auch lebensältere Mitarbeiter solle vor entsprechenden Folgen geschützt werden”

(b) BAG 2. Senat; Datum: 1986–03–06; Az: 2 AZR 262/85

Orientation: The case had to do with contractual retirement of airline cockpit crews. A pilot was born only a few days after a cut-off date that was of critical importance for retirement regulations.

“[Ausgleichsregelung] Sie führe auch nicht allein im Verhältnis zu nur wenig lebensälteren Kollegen zu einer niedrigeren Übergangsversorgung des Klägers”

(c) BAG 6. Senat; Datum: 1990–04–26; Az: 6 AZR 589/88

Orientation: A member of a philharmonic choir had been given a double room on a tour, but had booked a single room on his own, since he snored and was afraid of disturbing his colleague. He filed suit to recover his costs, which the employer had declined to pay.

“Drei Doppelzimmer seien einzeln mit jeweils lebensälteren Chormitgliedern belegt worden, bei denen Besonderheiten wie Behinderungen oder ähnliches vorgelegen hätten”

(d) BAG 8. Senat; Datum: 1991–06–13

Orientation: The case was about a female and a male employee of a civilian company working for the German Armed Forces who had applied for (promotion to) the same job. The woman was given the job. The man then took legal action against this decision on the grounds that he was older and had been longer on the job (and a former soldier). He succeeded and she then filed suit against this decision.

“zu dem Ergebnis, allein die Auswahl des in gleichem Maße geeigneten, jedoch dienst- und lebensälteren Bewerbers H, eines ehemaligen Zeitsoldaten und Oberfeldwebels, der sich bereits langjährig in Tätigkeiten einer höheren Vergütungsgruppe als die Klägerin bewährt habe, sei sachgerecht; ... Bei gleicher Eignung der Bewerber sei er lebens- und dienstälter ... ; ... Die Beklagte hat bei ihrer Auswahlentscheidung auf das bei gleicher Eignung höhere Lebens- und Dienstalter des Bewerbers H abgestellt”

I have no way of telling whether these are indeed the earliest occurrences of the word *lebensälter* in legal texts. There may be others. The above are the oldest occurrences in texts/verdicts of the BAG, however. There are no occurrences recorded in its jurisdiction before 1986. It seems that the cases cited in Table 9.5 are among the earliest references of the word. They contain, for the most part, a parallel construction of the nouns *Dienst-* and *Lebensalter* – and, accordingly, of the adjectives *dienstälter*

(which has been in usage in German for a long time) and *lebensälter*, still in its infancy, which help explain its meaning and linguistic explanation.

There can be no doubt that the science of linguistics has to deal with this type of word formation and language change, and offer interdisciplinary explanations of it to the science of law, which uses the word or the term, but does not spend much time explaining its meaning and structure. Perhaps linguists could also help establish guidelines concerning which terms need explanation most urgently, which new formations are necessary for terminological and non-terminological use, and which are not. There are many other examples such as *lebensälter* that are not at all clear. There are also examples that are much less intelligible and which pose a much bigger problem for lawyers, linguists and laymen.

Summary

To sum up, one can say that serious consequences can be derived from lawyers' misconceptions and misunderstandings of language and linguistic matters – and, equally so, linguists' misconceptions and misunderstandings of the law and objectives of lawyers' work. A substantial set of specific needs has been summarized above. The most common indicator of what is to be done is that each science, in addition to its “performative”, “diagnostic”, and “persuasive” jobs, has to offer “propedeutic” educational information across the borders of the disciplines. This is both worthwhile and necessary. The overall device must lead to a new accountability of professional “interdisciplinary realism”. Linguists working as forensic linguists together with lawyers dealing with language matters have to face reality in a wider, deeper and more precise fashion.

Coping with real life and its needs in a truly interdisciplinary perspective may entail some necessary adjustments and surprises for all parties involved. Forensic linguists may realize that what they are doing in an expert opinion presented in court is different from what they are doing the rest of the week, in their classrooms or in their armchairs. Lawyers dealing with language and linguistic matters may be surprised that this object looks different and that it takes more to understand how language works than when they were first exposed to it in secondary school, college or law school. Much more detailed information and a sound analysis of the interaction of experts and non-experts from both (and other) sciences involved seem to be the first step towards improving the spectrum of professional needs and interdisciplinary cooperation.

10

Outlook

Rather than giving a summary of the various issues touched upon in this book, I would like to state some major tasks ahead in forensic linguistics in the future, and some ingredients and consequences of future developments. The perspective is that of a German forensic linguist in 2007.

1. There has been an immense increase in the number of cases in which forensic linguists are asked for help and could be asked for help in legal contexts. There has also been an enormous increase in the number of practitioners acting as consultants and giving expert testimony in forensic linguistics in Germany, and also in research programs in forensic linguistics at German-speaking universities. A *Google* search for the term “Forensische Linguistik” showed some 1,230 hits. The number of hits for the English term “forensic linguistics” is much higher (89,400).

No one, including myself, would have expected this enormous increase in the 1970s, when a number of expert opinions were given in and for German courts, and in the early 1980s, when the first report of this activity appeared (Kniffka 1981). This increase in forensic linguistic activity, may it be welcomed or whatever, undoubtedly needs stronger, more intensive and deeper scientific analysis and attention than it has received to date. It is also of considerable urgency that the techniques and methodologies of forensic linguistics are developed further and receive some basic research coverage in general linguistics. This applies to basic research in the field of applied linguistics as well (see below). An increase in research and practical work also needs to be developed, describing ethical standards and guidelines as to how to distinguish forensic linguistic experts working *lege artis* from charlatans who are not.

Hand in hand with the development in the field of forensic linguistics in recent years, a development in linguistics – particularly corpus

linguistics and computational linguistics – has taken place. This can be considered a blessing for forensic linguistics, especially the analysis of anonymous texts with regard to authorship attribution. For the first time in history, very large corpora are available for various languages, including English and German. Also, using computers, several other successful quantitative analyses have been achieved in a prototypical analysis of textual data in authorship attribution; for example, Chaski (2001).

2. There is a terminological question at issue that has to be dealt with. It is not as critically important as it sometimes seems, if certain facts are taken into consideration. I share the view of Roger Shuy (personal communication, many years ago, my own rephrasing) that forensic linguistics is simply that type of linguistics that happens to deal with language used or analyzed in legal contexts. I entirely agree with this and have therefore not used the term “forensic linguistics” in the report given in Kniffka (1981) and in the following years. Forensic linguistics to me is merely a branch of applied linguistics in the very sense mentioned.

The real life usage of the term “forensic linguistics” and the German term “Forensische Linguistik” (which is not claimed to be identical in denotation and connotation with the English term here) is characterized by an ambiguity: “forensic linguistics” (henceforth FL) is (1) used broadly as a cover term for research and practical work in the entire area of “language and law”, for which many people working in that area would prefer that latter term (“language and law”) only; (2) FL is used as a technical term for linguistic consultancy and expert testimony “applied to all aspects of law” (Geracimos 2006). Most people working in the area would probably agree that this is the most preferable terminology, and would use the term FL in this sense.

Although such ambiguities are to be avoided, generally speaking, I do not believe that these two senses in which the term FL is used are a big disaster. There are precedents in the naming of other fields and subfields of linguistics. The field of “sociolinguistics” has experienced a similar fate in the last 50 years: “sociolinguistics” is used (1) as a cover term for the entire area of “sociolinguistics” and “sociology of language” (in Joshua Fishman’s (1971) sense); and (2) as a technical term for (mainly) micro-research of language in its social context only. Sociolinguistics has survived in both senses, and it looks as if it is going to stay for the future. Similarly, FL may also survive as a broad cover term for “language and law” and in the more specific sense of a branch of applied linguistics referring to linguistic consultancy and expert testimony in legal cases.

I would prefer to use “language and law” as the cover term and “forensic linguistics” in the more specific sense mentioned.

Rather than having a never-ending discussion on these issues, it seems to me of greater concern that truly false and unjustified terminological usages should be avoided. “FL” – at least, the German term “Forensische Linguistik” – is frequently used and understood by linguists as well as laymen as being identical with FL authorship attribution or FL analysis of anonymous authorship. This is certainly inadequate and should be avoided, since authorship attribution is just *one* part of the field.

3. As stated above, computer-aided quantitative analysis of linguistic and textual data is “where the future is” in FL, if I may say so. There is no question in my mind that everything that *can* be measured *should* be measured. At the same time, quantitative analysis is not all there is in and for FL’s future.

First, in an almost trivial sense, quite a few real life authorship attribution cases just do not have enough data to justify a computerized analysis. I have worked on many such cases.

In a second, even more crucial and salient sense, the general statement that there are other aspects of equal importance should be noted. Considerable basic research is to be done, as computational linguists working in the area have shown prototypically. It is a vast task ahead of us to achieve the basic research and the empirical testing of the methods of quantitative analysis in real life cases. In addition, a great deal of basic research in (general and applied) linguistics is necessary. There is an enormous reservoir of experience of linguists that have worked in the field for decades, which should be applied to the notion of idiolectal variation in more general theoretical and practical terms.

4. The most important task from an empirical sociolinguistic and anthropological linguistic perspective is a cross-cultural and cross-language exchange of knowledge relating to differences of facts, including linguistic facts, in different cultural, legal and language systems. It strongly reminds me of William Labov’s sarcastic dictum “Well, talking about linguistic universals – let’s just take any language – say English”. It appears to me that also in FL we have analyzed one, two or three different languages, legal systems, and cultures – and have absolutely no idea of what is going on in the rest of the many legal systems and cultures existing in the world. We are able, entitled and bound to draw empirical generalizations on legal systems, and perhaps activities such as FL, in a universal comparative perspective, if we have undertaken the basic comparative work on different cultures, legal systems and languages.

5. This would entail some modifications of the tasks ahead. We need to perform detailed and in-depth structural comparisons of, for example, the US, the German, and the Islamic, the Hindu and the Papua New Guinean systems, including questions of basic terminology, the linguistic status of legal institutions and all other elements. Also, notions such as “linguistics expert” and “linguistic experts’ testimony” would have to be compared thoroughly in the different cultures. It is difficult enough to account for the main differences, equivalences and non-equivalences of, for example, “Linguistischer Sachverständiger” in German and “linguistic expert witness” in English (in Britain and the US), and “Linguistischer Sachverständigenbeweis” in German and “linguistic expert testimony” in US and British courts. A “witness” is a “Zeuge” in German, a function that any respectable adult person can fulfil, while “an expert witness” is a “Sachverständiger” in German, with a legal status different from other witnesses. If one wishes to translate these differences, which are difficult enough just for English and German, one would see that, apart from the main translation business, there would be numerous difficult questions to be dealt with. From the American end, one could (and would have to) look for translations and notional equivalents of the “Frye Standard”, “Federal Rule of Evidence 702”, the “Daubert Standard”, and other criteria, which would have to be paraphrased in a rather complex and lengthy way to make sense in an entirely different legal system.

6. The main point made in 5 (that we do not know enough and, even more important, that we are not sufficiently well informed as far as data across cultures, legal systems and languages are concerned) can be made for less “exotic” anthropological domains, too. We need a much wider, closer and deeper documentary analysis of our own legal cultures, systems and jargons. As was stated repeatedly some time ago (see Kniffka 1990a; 1996b), we have no reliable data, let alone a full-fledged description and explanation of how many expert opinions, which types and sub-types of expert opinions have been given in legal contexts in the various European and US legal cultures. There are not even rough statistics of the types of cases in which FL expert opinions were or were not requested. Roger Shuy’s dictum (1993a: xxi) that one “can only speculate why the prosecution so infrequently calls on the services of linguists” still holds widely. We can indeed only speculate in many cases. Hard descriptive evidence beyond speculation as to why things are the way they are is urgently needed. Thorough investigation must be performed to discover to which legal systems and cultures this statement applies. As has been pointed out above, this does not seem to

apply to the German situation. From my own experience, I can only suspect that there is not such a sharp difference between the number of expert opinions called for by the prosecution and by the defense. The greatest number of expert opinions by far is requested by the courts, however. I believe the numbers to be almost equal in German criminal cases. I can only speculate on this and have no sound quantitative information. The fact that most FL expert opinions are requested by courts (judges) in Germany I cannot substantiate by sound empirical data. The desideratum stated in Kniffka (1996b) can only be repeated here: that it is of prime importance to acquire solid empirical data on who orders an FL expert opinion when, why, how and for what purpose.

Of equal, or even greater, importance would be a detailed documentation of the goals and outcomes that FL expert opinions have had in criminal and civil cases in courts in Germany, the US, the UK and other countries. This would seem to be the most important data that forensic linguists have to take into account for their future practical and theoretical work. A statistics of this caliber should also include in which (types of) criminal and civil cases linguistic experts were ordered to give (and have actually given) expert testimony in courts, and in which they have not.

A further question would be in which cases experts were actually asked to appear in court, and in which they were only asked to give written expert testimony for a court (which, as far as my own practical experience as an FL expert is concerned, by far outnumbers the former, being some 95 per cent of all cases).

7. An area of FL research of prime importance in the future will be the methodology, including some kind of a “meta-methodology” for FL. This includes a thorough analysis of the examples and the quotation of original data. To give just one example: the metaphors used in FL expert opinions and in the scientific literature composed by forensic linguists seem symptomatic in many respects. Every science, including the natural sciences, makes use of metaphors to exemplify and explain difficult concepts, relations and explanations, in particular if addressed to non-specialists (which is, indeed, the prototypical situation when giving FL expert testimony in court). It seems to me that linguists – in particular, theoretical linguists – are, to say the least, somewhat “allergic” to using metaphors to illustrate complicated matters, partly for a good reason and partly because of some kind of phobia against a tentatively “anecdotal” style and argumentation. On the basis of some three decades of giving expert testimony in and for German courts, I tend to think that such an aversion is totally unjustified for our profession. In the many

cases I have been consulted on, judges and courts were much more often swayed by linguistic data, examples and metaphors than by any theoretically adequate or more theoretical “insider” way of speaking. The job that FL has to do, to quote Roger Shuy (1993a: xvii) again, is explained most instructively by the following metaphor: “Just as physicians are trained to see things in an X-ray that the average person with excellent vision cannot see, so linguists are trained to see and hear structures that are invisible to lay persons.” Linguists have to understand that what convinces a fellow linguist may not convince a judge at all, and vice versa.

8. This leads to the final, and perhaps most important, postulate to be stated in this outlook. Every forensic linguistic book published in the last 15 to 20 years has at least one chapter, if not more, on “legal language”. There are also some (impressive) books devoted to this topic, such as Larry Solan’s “The Language of Judges” (1993), Solan and Tiersma’s “Speaking of Crime. The Language of Criminal Justice” (2005) and Peter Tiersma’s “Legal Language” (1999). The linguistic involvement with and attention to legal texts – in particular, the comprehensibility of legal texts, texts of law, texts of verdicts, the language used in court and so on – amounts to three-digit-numbers as far as books, and four- or even five-digit-numbers as far as articles are concerned, considering only the US and the German situations. This is quite reassuring.

It crosses my mind that not nearly as much linguistic attention is given to the language of forensic linguistics and the language used in FL expert opinions, however, including its comprehensibility. Linguists almost seem to prefer looking at other people’s language to looking at their own language, although it may need even more attention and critical reflection. One of the reasons that there is not nearly as much coverage of the language of linguistic people working in the judiciary may, in part, be explained by the fact that linguists have not made their own business accessible enough to non-linguists.

From my point of view, linguists have to do their homework in this respect, which seems to hold across legal cultures, languages and experts – particularly since the comprehensibility of linguists’ explanations and argumentations is generally of critical importance for the outcome and the effect that a linguistic expert opinion has. The theoretically and methodologically most well designed linguistic expert opinion is worthless if it is not written in a language that an “average

judge” can understand. To quote Roger Shuy once more (1993a: xix):

Appearing as an expert witness may well be the ultimate test of the applied linguist, since we are expected to be technically expert enough to have useful things to tell the jury but, at the same time, effective enough as teachers to be able to communicate technical information in ways that can be of immediate interest and usefulness to a jury.

This holds, *mutatis mutandis*, for the German system as well, in which not a jury but a judge is addressed.

True progress in interdisciplinary exchange in the area of “language and law” will not be gained by hoping or expecting that the other side will, after all, learn more from one’s own field, or by assuming that judges should just become “better linguists”. The reverse seems appropriate; that is, that one cleans one’s own house and makes it more accessible for visitors from other disciplines to find their way around and get a clearer impression of it, even if only for a short visit. Providing this type of interdisciplinary service to other sciences, linguistic work in the area of language and law and, in particular, forensic linguistics, may have a great future as an applied auxiliary science.

Notes

3 Status and Tasks of Forensic Linguistic Authorship Analysis

1. To illustrate the German situation more precisely, the German expressions in Tables 3.1a and 3.1b are supplied with English translations.
2. The English adjective 'linguistic' corresponds to two different adjectives in German: 'sprachlich' (concerning the object language) and 'linguistisch' (concerning the meta-language of linguistics).
3. The situation of corpus linguistics has changed considerably from the time when this paper was written (1987–88) to the present time. The simple fact that there are very large corpora available now for various languages including German is also of great help for FL.

4 “Shibboleths” as Data of Linguistic Behavior

1. 'Social marker' is used in an expansive social science meaning, rather than as a linguistic term as used by Labov.
2. From this, it follows that any use of shibboleths as indicators of a person's country of origin (as a political category), as done by some asylum authorities, is in itself unscientific and nonsensical.

5 Libel, Linguists, and Litigation in Germany

1. A shorter version of this paper was read at the 50th ILA Conference at John Jay College of Criminal Justice, New York, 15–17 April, 2005. I thank all who commented, in particular Roger Shuy, Carole Chaski, Peter Tiersma and Blake Howald.
2. More exactly, §§ 185–200 StGB are concerned with them, which cannot be described here in more detail.
3. I am indebted to Franz Rohrer and Angelika Wöbken of the BKA for supplying me with the unofficial English translation provided by the Federal Ministry of Justice.
4. A closer comparative analysis of German and US (and UK) law is not within the (linguistic) framework of this chapter.
5. I am indebted to Franz Rohrer and Angelika Wöbken of the BKA, who kindly supplied me with an online edition of the German 'Police Crime Statistics' of 2003.
6. The list is incomplete.
7. Many of the cases are minor physical injuries committed by more than one person acting together.
8. On this basis it is difficult to understand that in US courts notions such as 'defamatory *per se*', 'words that are defamatory *per se*' are used as classificatory

concepts. (Cf.:<http://www.wislawjournal.com/archive/2004/0929/slander-0929.html> 14.01.2007)

7 The System and Diagnostic Potential of Orthographic Data in Forensic Linguistic Authorship Attribution

1. Graphemic representation is given in < >.

8 Orthographic Data in Forensic Linguistic Authorship Analysis

1. Note: 'x' represents end of the text before ... /the last word before ...

9 Language and Law: Some Needs

1. In fact, the space available here could easily be used up merely by listing specific problems in linguist-lawyer interaction that have come up in the cases in which I was asked to give expert testimony. But, to give only a list would not be of much avail for practitioners or theoreticians. In addition to the information that misunderstandings are much more numerous, complex and multifold, one would wish to have more detailed information of when, where, why and how they happen and, at least on occasion, some specific consequences, and what one could and should do to (help) avoid them.
2. The blanks stand for personal names, which are omitted here.
3. Ungrammatical form instead of standard German 'fragt'.
4. A more detailed account of the case is given in Kniffka 1994. The point here is that all of this concerns common knowledge of (general) linguistics. It does not pertain to the specific object of the expert testimony given in this case (cf. needs listed Table 9.1: B.3, C.1, C.2, C.3).
5. This, again, may hold the other way around too. Faulty and defective linguistic expert testimony has been described elsewhere in more detail (Kniffka 1993a).
6. I am indebted to the Ministerium für Wissenschaft und Forschung des Landes Nordrhein-Westfalen and to Professor P. Hanau and his staff of the Institut für Arbeits- und Wirtschaftsrecht at the University of Cologne for supplying detailed information.

References

- Agheysi, R. and Fishman, J.A. (1970) 'Language Attitude Studies: A Brief Survey of Methodological Approaches', *Anthropological Linguistics*, 12/5: 137–57.
- Bauman, R. and Sherzer, J. (eds) (1974) *Explorations in the Ethnography of Speaking* (Cambridge: Cambridge UP).
- Beaugrande, R.-A. de and Dressler, W. (1981) *Einführung in die Textlinguistik* (Tübingen: Niemeyer).
- Biber, D. (1989) 'A Typology of English Texts', *Linguistics*, 27: 3–43.
- Biber, D. and Finnegan, E. (1986) 'An Initial Typology of English Text Types', in J. Aarts and W. Meijs (eds), *Corpus Linguistics II* (Amsterdam: Rodopi): 19–46.
- Bierwisch, M. (1970) 'Semantics', in J. Lyons (ed.), *New Horizons in Linguistics* (Harmondsworth: Penguin Books): 166–84.
- Blum, A. (1990) 'Anregung zu mehr Interdisziplinarität bei forensischen linguistischen Untersuchungen', in H. Kniffka, *Texte zu Theorie und Praxis forensischer Linguistik* (Tübingen: Niemeyer): 289–319.
- Bremer, H. (1973) *Der Sachverständige. Seine Stellung im privaten und öffentlichen Recht* (Heidelberg: Verlagsgesellschaft Recht und Wirtschaft mbH).
- Brinker, K. (1988) *Linguistische Textanalyse: Ein Eingriff in Grundbegriffe und Methoden. 2. durchges. u. erg. Aufl.* (Berlin: Schmidt).
- Brinker, K. (1993) *Textlinguistik* (Heidelberg: Groos).
- Brown, P. and Fraser, C. (1979) 'Speech as a Marker of Situation', in K.R. Scherer and H. Giles (eds), *Social Markers in Speech* (Cambridge: Cambridge UP): 33–62.
- Brückner, T. (1989) 'Gibt es einen "sprachlichen Fingerabdruck"? Kritische Anmerkungen zum forensischen Textvergleich', in *Institut für deutsche Sprache. Sprachreport*, 4: 14–16.
- Bundeskriminalamt (1989) 'Forensisch-linguistischer Textvergleich', Symposium am 8. und 9.12.1988 im BKA Wiesbaden, Wiesbaden 1989 (= BKA – Technische Forschung Band II).
- Bundeskriminalamt (2003) *Police Crime Statistics*.
- Chaski, C.E. (1997) 'Who Wrote It? Steps Toward A Science of Authorship Identification', *National Institute of Justice Journal*, September: 15–22.
- Chaski, C.E. (2001) 'Empirical Evaluations of Language-based Author Identification Techniques', *Forensic Linguistics: International Journal of Speech, Language, and the Law*, 8(1): 1–65.
- Chaski, C.E. (2004) 'Recent Validation Results for the Syntactic Analysis Method for Author Identification', International Conference on Language and Law, Cardiff, Wales.
- Cooper, R.L. and Fishman, J.A. (1975) 'The Study of Language Attitudes', *IJSL*, 4, Singles Issue: 5–19.
- Corder, S.P. (1971) 'Idiosyncratic Dialects and Error Analysis', *International Review of Applied Linguistics*, 9: 147–159.
- Coulmas, F. (1985) 'Diskursive Routine im Fremdsprachenunterricht', *Sprache und Literatur in Wissenschaft und Unterricht*, 56: 47–66.

- Coulthard, M. (1993) 'Beginning the study of forensic texts: corpus, concordance, collocation', in M.P. Hoey (ed.), *Data Description Discourse* (London: HarperCollins): 86–97.
- Davis, T. (1996) 'Clues and Opinions: Ways of Looking at Evidence', in H. Kniffka (ed. – in cooperation with Susan Blackwell and Malcolm Coulthard), *Recent Developments in Forensic Linguistics* (Frankfurt a.M.: Lang): 53–73.
- Dippel, K. (1986) *Die Stellung des Sachverständigen im Strafprozess* (R. v. Deckers rechts- und sozialwissenschaftliche Abhandlungen, Bd. 30) (Heidelberg: Decker).
- Doherty, E.T. and Hollien, H. (1978) 'Multiple-factor speaker identification of normal and distorted speech', *Journal of Phonetics*, 6: 1–8.
- Dreher, E. (1977) *Kommentar zum StGB* (München: Beck).
- DUDEN (1980) *Rechtschreibung der deutschen Sprache und Fremdwörter* (Mannheim: Duden).
- DUDEN (1996) *Rechtschreibung der deutschen Sprache. 21. Aufl.* (Mannheim: Duden).
- DUDEN (2000) *Die deutsche Rechtschreibung. 22. Aufl.* (Mannheim: Duden).
- Durant, A. (1996) 'Allusions and Other 'Innuendo' Meanings in Libel Actions: The Value of Semantic and Pragmatic Evidence', *Forensic Linguistics: International Journal of Speech, Language, and the Law*, 3(2): 195–210.
- Eades, D. (2005) 'Applied Linguistics and Language Analysis in Asylum Seeker Cases', *Applied Linguistics*, 26(4): 503–26.
- Eades, D. and Arends, J. (2004) 'Using Language Analysis in the Determination of National Origin of Asylum Seekers: An Introduction', *Speech, Language and the Law*, 11(2): 179–99.
- Feibleman, J.K. (1946) *An Introduction to Peirce's Philosophy Interpreted as a System* (London: Allen & Unwin).
- Fischer, J.L. (1958) 'Social Influences on the Choice of a Linguistic Variant', *Word*, 14: 47–56.
- Fishman, J. (ed.) (1971) *Advances in the Sociology of Language. Basic Concepts, Theories and Problems. Alternative Approaches* (The Hague: Mouton).
- Fleischer, W. and Barz, I. (1992) *Wortbildung der deutschen Gegenwartssprache* (Tübingen: Niemeyer).
- Friedrichs, J. (1973) *Methoden empirischer Sozialforschung* (Reinbek: Rowohlt).
- Fromkin, V.A. (ed.) (1980) *Errors in Linguistic Performance. Slips of the Tongue, Ear, Pen and Hand* (New York: Academic Press).
- Geerds, F. (1983) 'Sachbeweis und Sachverständigenbeweis in Strafsachen aus kriminalistischer Sicht', *Archiv für Kriminalistik*: 172.
- Geracimos, A. (2006) 'CSI: Language Analysis Unit', *Washington Times*, online edition. www.washingtontimes.com/entertainment/20060111-094446-5919r.htm
- Gibbons, J. (1999) 'Language and the Law', *Annual Review of Applied Linguistics*, 19: 156–73.
- Gilbert Law Summaries (1984) *Dictionary of Legal Terms*, 1st edn (Chicago: Harcourt, Brace, Jovanovich Legal and Professional Publications, Inc).
- Giles, H. (1979) 'Ethnicity Markers in Speech', in K.R. Scherer and H. Giles (eds), *Social Markers in Speech* (Cambridge: Cambridge UP): 251–89.
- Giles, H., Scherer, K.R. and Taylor, D.M. (1979) 'Speech Markers in Social Interaction', in K.R. Scherer and H. Giles (eds), *Social Markers in Speech* (Cambridge: Cambridge UP): 343–81.

- Grant, T.D. (2001) 'Reviewing and Revising Stylometric Authorship Attribution for Use in a Forensic Context', Paper presented at the IAFL 5th Biennial Conference, University of Malta.
- Grice, H.P. (1975) 'Logic and Conversation', in P. Cole and J.L. Morgan (eds), *Syntax and Semantics, Vol. 3, Speech Acts* (New York: Academic Press).
- Grotjahn, R. (1979) *Linguistische und statistische Methoden in Metrik und Textwissenschaft* (Bochum: Brockmeyer): Quantitative Linguistics 1.
- Gumperz, J.J. (1964) 'Linguistic and Social Interaction in Two Communities', *American Anthropologist*, 66.2, Part 2: 137–53.
- Gumperz, J. and Hymes, D. (eds) (1972) *Directions in Sociolinguistics. The Ethnography of Communication* (New York: Holt, Rinehart & Winston).
- Halle, M. (1954) 'Why and How Do We Study the Sounds of Speech?', in M. Halle (ed.) (2003), *From Memory to Speech and Back* (Berlin/New York: de Gruyter).
- Haller, B. and Klein, H. (1986) 'Überlegungen zum kriminaltechnischen Sachbeweis und den Möglichkeiten seiner wahrscheinlichkeitstheoretischen Bewertung', *Archiv für Kriminalistik*, 177.
- Hancher, M. (1980) 'Speech Acts and the Law', in R. Shuy and A. Shnukal (eds), *Language Use and the Uses of Language* (Washington, D.C.: Georgetown UP): 245–56.
- Hecker, M.R. (1990) 'Forensische Linguistik und forensische Schriftuntersuchung – Zwei methodisch verwandte Disziplinen?', in H. Kniffka (ed.), *Texte zu Theorie und Praxis forensischer Linguistik* (Tübingen: Niemeyer): 231–45.
- Hecker, M.R. (1997) 'Geschlechtseinflüsse in die Handschrift', in M.R. Hecker (ed.), *6. Symposion für Schriftvergleichung des Bundeskriminalamtes und der Landeskriminalämter*, 17/18 March 1997 (Tagungsband: Wiesbaden): 47–64.
- Hecker, M.R. (2000) *Traktat über den Wissenschaftlichkeitsanspruch der forensischen Schriftvergleichung* (Habilitationsschrift, Leopold-Universität Breslau).
- Heike, G. (1990) 'Stimmenvergleichsgutachten – Probleme und Methoden der forensischen Phonetik am Beispiel einer Fallstudie', in H. Kniffka (ed.), *Texte zu Theorie und Praxis forensischer Linguistik* (Tübingen: Niemeyer): 481–9.
- Herdegen, G. (1975) *Leipziger Kommentar zum Strafgesetzbuch* (§§ 185–94).
- Hirsch, H.-J. (1967) *Ehre und Beleidigung* (Karlsruhe: C.F. Müller).
- Holst, F. (1978) *Morphologie. Einführungspapier mit Arbeitsaufgaben* (Trier: L.A.U.T. (= KLAGÉ 2)).
- Hymes, D. (1962) 'The Ethnography of Speaking', in T. Gladwin and W. Sturtevant (eds), *Anthropology and Human Behavior* (Washington, D.C.: Anthropological Society of Washington): 13–53.
- Hymes, D. (1966) 'On Communicative Competence', Paper originally read at the Research Planning Conference on Language Development among Disadvantaged Children, Yeshiva University, June 1966, Reprinted, in part, in C.J. Brumfit and K. Johnson (eds) (1979), *The Communicative Approach to Teaching* (Oxford: Oxford UP).
- Jakobson, R., Fant, G. and Halle, M. (1951) *Preliminaries to Speech Analysis* (Cambridge, MA: MIT Press).
- James, C. (1998) *Errors in Language Learning and Use: Exploring Error Analysis* (Harlow: Longman).
- Jessnitzer, K. (1978) *Der gerichtliche Sachverständige*, 7., neubearbeitete Auflage (Köln: Heymanns).

- Jöns, D. (1982) 'Der philologische Steckbrief. Über den Einsatz der Philologie bei der Täterermittlung', in *Gesellschaft und Universität. Probleme und Perspektiven*. Festschrift zur 75-Jahr-Feier der Universität Mannheim: 273–87.
- Kastenholz, R. (1998) 'Eine afrikanistische Stellungnahme zur Sprachanalyse zum Zwecke der Herkunftsbestimmung von Asylantragstellern', <http://www.proasyl.de/lit/spranaly/spr2.htm>
- Kastovsky, D. (ed.) (1980) *Perspektiven der lexikalischen Semantik*, Beiträge zum Wuppertaler Semantikkolloquium 2–3 December 1977 (Bonn: Bouvier).
- Kleinknecht, Th. (1977) *Strafprozessordnung (StPO) Kurzkommentar* (München: Beck).
- Kniffka, H. (1980) *Soziolinguistik und empirische Textanalyse. Schlagzeilen- und Leadformulierung in amerikanischen Tageszeitungen*, Habilitationsschrift, 1980 (Tübingen: Niemeyer (= Linguistische Arbeiten 94)).
- Kniffka, H. (1981) 'Der Linguist als Gutachter bei Gericht. Überlegungen und Materialien zu einer "Angewandten Sprachwissenschaft"', in G. Peuser and S. Winter (eds), *Angewandte Sprachwissenschaft. Grundfragen – Bereiche – Methoden* (Bonn: Bouvier): 584–634.
- Kniffka, H. (1989) 'Zum gegenwärtigen Stand und zu den Aufgaben "forensischer" Linguistik', Vortrag auf dem BKA-Symposium am 09.12.1988 in Wiesbaden, in Bundeskriminalamt (ed.), *Forensisch-linguistischer Textvergleich*, Symposium am 8. und 9.12.1988 im BKA Wiesbaden: 202–37 (= BKA – Technische Forschung Band II).
- Kniffka, H. (ed.) (1990a) *Texte zu Theorie und Praxis forensischer Linguistik* (Tübingen: Niemeyer (= Linguistische Arbeiten 249)).
- Kniffka, H. (1990b) 'Einführung: "Forensische Linguistik"', in H. Kniffka (ed.), *Texte zu Theorie und Praxis forensischer Linguistik* (Tübingen: Niemeyer): 1–55.
- Kniffka, H. (1990c) 'Alles ist Tatfrage des Einzelfalls – oder: "Herr X bzw. seine Konkubine"', in H. Kniffka (ed.): *Texte zu Theorie und Praxis forensischer Linguistik* (Tübingen: Niemeyer): 189–200.
- Kniffka, H. (1990d) 'Autorschaftsausschluss. Ein "liquet" und ein "non-liquet"', in H. Kniffka (ed.), *Texte zu Theorie und Praxis forensischer Linguistik* (Tübingen: Niemeyer): 437–56.
- Kniffka, H. (1990e) 'Nochmals: Verrät der Text den Verfasser?', in *Kriminalistik*, 11/90: 604–10.
- Kniffka, H. (1991) 'Schibboleths. Philologische Bestandsaufnahme und Gesichtspunkte zu ihrer soziolinguistischen Analyse', in *Deutsche Sprache*, 2/1991: 159–77.
- Kniffka, H. (1992) 'Sprachwissenschaftliche Hilfe bei der Täterermittlung', in G. Grewendorf (ed.), *Rechtskultur als Sprachkultur. Zur forensischen Funktion der Sprachanalyse* (Frankfurt: Suhrkamp): 157–93.
- Kniffka, H. (1993a) 'Gutachten über ein Fehlgutachten', in *Kriminalistik und forensische Wissenschaften*, 82: 123–42.
- Kniffka, H. (1993b) 'Zur Wertigkeit und zum Erklärungspotential orthographischer Varianten', in M. Hecker and M.P. Leder (eds), *Akten des 1. Kongresses der Deutschen Gesellschaft für Schriftanalyse* (XI. Mannheimer Symposium für Schriftvergleichung). Humboldt-Universität Berlin, 30 March–2 April 1993, Wiesbaden: 1–28.
- Kniffka, H. (1994) 'Understanding Misunderstandings in Court. "La Serva Padrona" – Phenomena and other Mis-Communications', in D. Carson and

- R. Bull, *Expert Evidence. The International Digest of Human Behaviour, Science and Law* (Chichester, West Sussex: SLE Publications), 2.4, March 1994: 164–75.
- Kniffka, H. (1995) 'Lückenkonsrast. Elemente einer "kultur-kontrastiven" Grammatik Deutsch – Arabisch und Deutsch – Chinesisch, in H. Kniffka (ed.), *Elements of Culture-Contrastive Linguistics* (Frankfurt a.M.: Lang): 37–62.
- Kniffka, H. (ed.) (1996a) *Recent Developments in Forensic Linguistics* (in cooperation with Susan Blackwell and Malcolm Coulthard) (Frankfurt a.M.: Lang).
- Kniffka, H. (1996b) 'Editor's Introduction', in H. Kniffka (ed.), *Recent Developments in Forensic Linguistics* (in cooperation with Susan Blackwell and Malcolm Coulthard) (Frankfurt a.M.: Lang): 19–50.
- Kniffka, H. (1996c) 'On Forensic Linguistic "Differential Diagnosis"', in H. Kniffka (ed.), *Recent Developments in Forensic Linguistics* (in cooperation with Susan Blackwell and Malcolm Coulthard) (Frankfurt a.M.: Lang): 75–121.
- Kniffka, H. (1998) 'Forensische Phonetik und forensische Linguistik', in B.J. Kröger, C. Riek and G. Sachse (eds), *Festschrift Georg Heike* (Forum Phonetikum 66), Frankfurt a.M.: 279–91.
- Kniffka, H. (2000a) 'Stand und Aufgaben der "Forensischen Linguistik"', in D. Veronesi (ed.), *Linguistica giuridica italiana e tedesca/Rechtslinguistik des Deutschen und Italienischen* (Padova: Unipress): 29–46.
- Kniffka, H. (2000b) 'Anonymous Authorship Analysis without Comparison Data? A Case Study with Methodological Impact', Paper read at the 4th Biennial Conference of the IAFL at the University of Birmingham, UK, 28 June – 1 July 1999, in *Linguistische Berichte*, 182: 179–98.
- Kniffka, H. (2001) 'Eine Zwischenbilanz aus der Werkstatt eines "forensischen Linguisten": Zur Analyse anonymer Autorschaft', in *Linguistische Berichte*, 185: 75–104.
- Kniffka, H. (2002) 'Sprach- und Kulturkontakt "across the fence(s)". Linguistische Perspektiven', in E. Apeltauer (ed.), *Interkulturelle Kommunikation: Deutschland – Skandinavien – Großbritannien* (Tübingen: Narr): 7–30.
- Kniffka, H. (2003a) 'Zur Wertigkeit und zum Erklärungspotenzial orthographischer Varianten', in M. Baumann and J. Ossner (eds), *Diagnose und Schrift 1: Die Schreiber*, OBST (= Osnabrücker Beiträge zur Sprachtheorie 66): 49–84.
- Kniffka, H. (2003b) 'Texturheber- und Schreiberheberschaftsanalyse. Eigengesetzlichkeit und diagnostisches Potenzial orthographischer Daten im Rahmen einer forensisch-linguistischen Analyse', in M. Baumann and J. Ossner (eds), *Diagnose und Schrift 1: Die Schreiber*, OBST (= Osnabrücker Beiträge zur Sprachtheorie 66): 85–116.
- Koch, P. and Oesterreicher, W. (1994) 'Schriftlichkeit und Sprache', in H. Günther and O. Ludwig (eds), *Schrift und Schriftlichkeit. Writing and Its Use. Ein interdisziplinäres Handbuch internationaler Forschung*, 1, Halbbd (Berlin: de Gruyter) (Handbücher zur Sprach- und Kommunikationswissenschaft 10.1): 586–604.
- Köller, N., Nissen, K., Rieß, M. and Sadorf, E. (2004) 'Probabilistische Schlussfolgerungen in Schriftgutachten: Zur Begründung und Vereinheitlichung von Wahrscheinlichkeitsaussagen in Sachverständigengutachten', *Polizei und Forschung Bd.*, 26 (Neuwied/Rhein: Luchterhand).
- König, R. (ed.) (1967) *Handbuch der empirischen Sozialforschung, Band I* (Stuttgart: Enke).
- Künzel, H. (1987) *Sprechererkennung. Grundzüge forensischer Sprachverarbeitung* (Heidelberg: Kriminalistik Verlag).

- Labov, W. (1966) *The Social Stratification of English in New York City* (Washington, D.C.: Center for Applied Linguistics).
- Labov, W. (1970) 'The Study of Language in its Social Context', *Studium Generale*, 23/1: 30–87.
- Labov, W. (1972a) *Sociolinguistic Patterns* (Philadelphia: Pennsylvania UP).
- Labov, W. (1972b) 'Some Features of the English of Black Americans', in R.W. Bailey and J.L. Robinson (eds), *Varieties of Present-Day English* (New York: MacMillan): 236–55.
- Labov, W. (1972c) 'Some Principles of Linguistics Methodology', *Language in Society*, 1/1: 97–120.
- Labov, W. (1975) 'On the Use of the Present to Explain the Past', in L. Heilmann (ed.), *Proceedings of the 11th International Congress of Linguists* (Bologna: Il Mulino): 825–51.
- Labov, W. (2001) *Principles of Language Change. Volume II: Social Factors* (Oxford: Blackwell).
- Laver, J. and Trudgill, P. (1979) 'Phonetic and Linguistic Markers in Speech', in K.R. Scherer and H. Giles (eds), *Social Markers in Speech* (Cambridge: Cambridge UP): 1–32.
- Levi, J.N. (1982) *Linguistics, Language, and the Law: A Topical Bibliography* (Bloomington: Indiana University Linguistics Club).
- Levi, J.N. (1986) 'Applications of Linguistics to the Language of Legal Interactions', in P.C. Bjarkman and V. Raskin (eds), *The Real-World Linguist: Linguistic Applications in the 1980s* (Norwood, NJ: Ablex): 230–65.
- Levi, J.N. (1994) 'Language as Evidence: The Linguist as Expert Witness in North American Courts', *Forensic Linguistics: International Journal of Speech, Language, and the Law*, 1:1: 1–26.
- Levi, J.N. and Walker, A. (eds) (1990) *Language in the Judicial Process* (New York: Plenum Publishing).
- Maas, U. (1979) 'Auf der Suche nach dem unverbildeten Durchschnittsleser. Ein sprachwissenschaftliches Gutachten in einem politischen Prozess', *Linguistische Berichte*, 61: 52–68.
- McMenamin, G.R. (1993) *Forensic Stylistics* (Amsterdam: Elsevier Science).
- McMenamin, G.R. (2002) *Forensic Linguistics: Advances in Forensic Stylistics* (Boca Raton, FL: CRC Press).
- Michel, L. (1982) *Gerichtliche Schriftvergleichung: Eine Einführung in Grundlagen, Methoden und Praxis* (Berlin: de Gruyter).
- Michel, L. (1990) 'Grundlagen und Methoden der forensischen Schriftuntersuchung', in H. Kniffka (ed.), *Texte zu Theorie und Praxis forensischer Linguistik* (Tübingen: Niemeyer): 217–30.
- Müller, K. (1978) *Der Sachverständige im gerichtlichen Verfahren. Handbuch des Sachverständigenbeweises* (Heidelberg: C.F. Müller).
- Oksaar, E. (1967) 'Sprache als Problem und Werkzeug des Juristen', *Archiv für Rechts- und Sozialphilosophie*, 91–132.
- Preisdanz, H. (1975) *Strafgesetzbuch: Lehrkommentar mit Erläuterungen und Beispielen, ausgewählten Nebengesetzen sowie einem Anhang über Jugendstrafrecht. 29. Auflage* (Berlin: J. Schweitzer).
- Pullum, G.K. (1991) 'The Linguistics of Defamation', in G.K. Pullum, *The Great Eskimo Vocabulary Hoax and Other Irreverent Essays on the Study of Language* (Chicago: Chicago UP): 92–9.

- Rehbein, J. (ed.) (1985) *Interkulturelle Kommunikation* (Tübingen: Narr).
- Rieber, R.W. and Stewart, W.A. (eds) (1990) 'The Language Scientist as Expert in the Legal Setting: Issues in Forensic Linguistics', in *Annals of the New York Academy of Sciences* 606 (New York: The New York Academy of Sciences: 1990).
- Sapir, E. (1927) 'Speech as a Personality Trait', *American Journal of Sociology*, 32, 892–905. Reprinted as 'Speech and Personality', in N. Markel (ed.) (1969) *Psycholinguistics. An Introduction to the Study of Speech and Personality* (Homewood, IL: Dorsey Press): 44–56.
- Saville-Troike, M. (1982) *The Ethnography of Communication: An Introduction* (Oxford: Blackwell).
- Schall, S. and Hehn, W. (1996) 'Das System KISTE im BKA', in H. Kniffka (ed. – in cooperation with Susan Blackwell and Malcolm Coulthard), *Recent Developments in Forensic Linguistics* (Frankfurt a. M.: Lang).
- Scherer, K.R. (1979) 'Personality Markers in Speech', in K.R. Scherer and H. Giles (eds), *Social Markers in Speech* (Cambridge: Cambridge UP): 147–209.
- Scherer, K.R. and Giles, H. (eds) (1979) *Social Markers in Speech* (Cambridge: Cambridge UP).
- Schönke, A. and Schroeder, H. (1976) *Strafgesetzbuch – Kommentar* (§§153–205), bearbeitet von Th. Lenckner (München: Beck).
- Sendler, H. (1986) 'Richter und Sachverständige', *Neue Juristische Wochenschrift*, 39: 2907–15.
- Shuy, R.W. (1982) 'Topic as the Unit of Analysis in a Criminal Law Case', in D. Tannen (ed.), *Analysing Discourse: Text and Talk* (Washington, D.C.: Georgetown UP): 113–26.
- Shuy, R.W. (1984) 'Linguistics and Other Professions', *Annual Review of Anthropology*, 13: 419–45.
- Shuy, R.W. (1986a) 'Language and the Law', *Annual Review of Applied Linguistics*, 7 (Cambridge: Cambridge UP).
- Shuy, R.W. (1986b) 'Context as the Highest Standard of Semantics: A Lawsuit Involving the Meaning of "Accuracy" in Accounting', *Journal of English Linguistics*, 19.2 (October): 295–303.
- Shuy, R.W. (1987) 'Some Contributions of a Linguist in a Criminal Case', in S. Fisher and A. Todd (eds), *Discourse and Institutional Authority* (Norwood, NJ: Ablex).
- Shuy, R.W. (1990) 'Evidence of Cooperation in Conversation: Topic Type in a Solicitation to Murder Case', in R. Rieber and W.A. Stewart (eds), *The Language Scientist as an Expert Witness: Issues in Forensic Linguistics*, in *Annals of the New York Academy of Sciences*, 606: 85–105.
- Shuy, R.W. (1993a) *Language Crimes: The Use and Abuse of Language Evidence in the Courtroom* (Oxford: Blackwell).
- Shuy, R.W. (1993b) 'Using Language Evidence in Money Laundering Trials', in *Money Laundering Update: Legislative, Regulatory, Enforcement Developments*, American Bankers Association and American Bar Association, Washington D.C.: 936–54, Published in *American Speech*, 68.1.
- Shuy, R.W. (1995) 'Dialect as Evidence in Law Cases', *Journal of English Linguistics*, 23, 1/2: 195–208.
- Shuy, R.W. (1998a) *The Language of Confession, Interrogation and Deception* (Thousand Oaks, CA: Sage Publications).

- Shuy, R.W. (1998b) *Bureaucratic Language in Government and Business* (Washington, D.C.: Georgetown UP).
- Shuy, R.W. (2000) 'Forensic Linguistics', in M. Aronoff and J. Rees-Miller (eds), *The Handbook of Linguistics* (Oxford: Blackwell)
- Shuy, R.W. (2001) 'Discourse Analysis in the Legal Context', in D. Schiffrin, D. Tannen and H. Hamilton (eds), *The Handbook of Discourse Analysis* (Oxford: Blackwell): 437–52.
- Shuy, R.W. (2002a) 'To Testify or Not To Testify?', in J. Cotterill (ed.), *Language in the Legal Process* (New York: Palgrave Macmillan): 3–18.
- Shuy, R.W. (2002b) *Linguistic Battles in Trademark Publications* (New York: Palgrave Macmillan).
- Shuy, R.W. (2005) *Creating Language Crimes* (New York: Oxford UP).
- Shuy, R.W. (2006) *Linguistics in the Courtroom: A Practical Guide* (New York: Oxford UP).
- Shuy, R.W. and Fasold, R. (eds) (1973) *Language Attitudes: Current Trends and Prospects* (Washington, D.C.: Georgetown UP).
- Solan, L.M. (1993) *The Language of Judges* (Chicago: Chicago UP).
- Solan, L.M. (1998) 'Linguistic Experts as Semantic Tour Guides', *Forensic Linguistics*, 5.2: 87–106.
- Solan, L.M. and Tiersma, P.M. (2002) 'The Linguist on the Witness Stand: Forensic Linguistics in American Courts', *Language*, 78: 221–39.
- Solan, L.M. and Tiersma, P.M. (2005) *Speaking of Crime: The Language of Criminal Justice* (Chicago: Chicago UP).
- Sornig, K. (1985) 'Intim-Varianten', in J. Rehbein (ed.), *Interkulturelle Kommunikation* (Tübingen: Narr): 175–89.
- Strafgesetzbuch (1994) *Textausgabe mit ausführlichem Sachregister und einer Einführung von Jeschek, H.-H.* (München: DTV).
- Svartvik, J. (1968) *The Evans Statements: A Case for Forensic Linguistics*, Gothenburg Studies in English, 20 (Goteborg: Acta Universitatis Gothoburgensis).
- Tannen, D. and Saville-Troike, M. (1985) *Perspectives on Silence* (Norwood, NJ: Ablex).
- Thomas, H. and Putzo, H. (1977) *Zivilprozessordnung* (München: Beck).
- Tiersma, P.M. (1987) 'The Language of Defamation', *Texas Law Review*, 66:2: 303–50.
- Tiersma, P.M. (1999) *Legal Language* (Chicago: Chicago UP).
- Truss, L. (2003) *Eats, Shoots & Leaves. The Zero Tolerance Approach to Punctuation* (London: Profile Books).
- Wagner, B. (1997) 'Probleme bei der Begutachtung türkischer Handschriften', in M. Hecker (ed.), 6. Symposium für Schriftvergleichung des Bundeskriminalamtes und der Landeskriminalämter, 17/18 March 1997, Tagungsband, Wiesbaden: 65–82.
- Wahrig, G. (1973) *Anleitung zur grammatisch-semantischen Beschreibung lexikalischer Einheiten* (Tübingen: Niemeyer).
- Weinreich, U. (1966) 'Explorations in Semantic Theory', *Current Trends in Linguistics*, 3: 395–477.

Index of Authors

[Numbers appearing in brackets following page numbers suffixed with “n” or “nn” indicate the chapter number followed by the relevant note cue number(s)]

- Agheyisi, Rebecca N., 95
- Barz, Irmhild, 198, 199
- Bauman, Richard, 95
- Biber, Douglas, 163
- Blackwell, Susan A., 15
- Blum, Anita, 30
- Bolinger, Dwight, 139
- Brinker, Klaus, 163
- Bundeskriminalamt, 4, 17, 40, 71, 121, 142
- Chaski, Carole E., 217, 223, 226, 272n(5.1)
- Cooper, Robert L., 95
- Corder, S. Pit, 68, 81, 177, 211
- Davis, Tom, 187
- Dippel, Karlhans, 74
- Doherty, E. Thomas, 43, 44
- DUDEN, 82, 189, 221, 222, 223, 224, 229, 257, 258, 259, 261
- Eades, Diana, 88
- Ervin-Tripp, Susan M., 139
- Fasold, Ralph, 95
- Ferguson, Charles A., 139
- Fishman, Joshua A., 95, 266
- Fleischer, Wolfgang, 198, 199
- French, J. Peter, 3, 15, 17, 64, 87, 208, 250, 260
- Fromkin, Victoria, 178, 198
- Geerds, Friedrich, 74
- Geracimos, Ann, 266
- Giles, Howard, 19, 95, 97
- Grant, Tim, 217
- Grice, H. Paul, 93
- Grotjahn, Rüdiger, 223
- Gumperz, John J., 65, 79, 92, 95, 241
- Hecker, Manfred R., 30, 196, 220
- Heike, Georg, 46
- Herdegen, Gerhard, 143
- Hollien, Harry, 43, 44
- Hymes, Dell H., 65, 79, 92, 95, 100, 130, 241
- James, Carl, 177
- Jessnitzer, Kurt, 243
- Kniffka, Hannes, 2, 15, 16, 17, 28, 32, 33, 34, 51, 70, 71, 77, 91, 100, 126, 129, 149, 166, 180, 186, 192, 196, 209, 211, 223, 230, 241, 245, 268, 269, 273nn(9.4–5)
- Koch, Peter, 91
- Köller, Norbert, 36
- Labov, William, 95, 100, 103, 114, 127, 139, 175, 182, 262, 267, 272n(4.1)
- Levi, Judith N., 2, 3, 16, 28, 31, 166, 230
- McMenamin, Gerald R., 166
- Michel, Lothar, 15, 21, 30, 36, 91, 196, 220
- Nissen, Kai, 36, 46
- Österreicher, Wulf, 91
- Preisendanz, Holger, 138
- Pullum, Geoffrey K., 20

- Rieber, Robert, 2, 230
Rieß, Michael, 36, 46
- Sadorf, Erwin, 36, 46
Sapir, Edward, 19, 91, 92, 94, 108,
185, 218
Saville-Troike, Muriel, 95
Scherer, Klaus R., 19, 95, 97,
101
Schönke, Adolf, 138
Schroeder, Horst, 138
Searle, John R., 139
Sandler, Horst, 34, 35, 243
Shapero, Jess, 15
Sherer, K.R., 19, 95, 97
- Sherzer, Joel, 95
Shuy, Roger W., 2, 3, 16, 17, 20, 28,
31, 34, 42, 95, 115, 120, 139, 149,
166, 230, 266, 268, 270, 271,
272n(5.1)
Solan, Lawrence M., 3, 16, 32, 166,
230, 270
Stewart, William A., 2, 230
- Taylor, Donald M., 19, 95, 97
Tiersma, Peter M., 3, 16, 32, 166, 230,
270, 272n(5.1)
Truss, Lynne, 222
- Walker, Anne G., 16, 28, 31, 166, 230

Subject Index German

(Numbers appearing in **bold** indicate that the pages contain diagrams/figures/illustrations/tables)

- Äußerung, 138, 206, **207**
- Bedeutung, **140**
- Begutachtbarkeit, 5
- Behördengutachten, 8
- Beklagte/r, 144, **263**
- Beleidigung, **118, 121, 122, 124, 138, 143, 145**
- Beleidigungsrecht, 4, 7
- Bezeichnung, wertneutrale, **125**
- Bundeskriminalamt (BKA; German Federal Criminal Police Office), 4, 84, 142
- Durchschnittsleser,
unvoreingenommener/
unbeeinflusster/unverbildeter,
133, 135, 242, 245
- Ermittlungsverfahren, 9, 201,
230
- Fehler
- Flüchtigkeits-, 82, 178
- Sprach-, 174
- Gendarstellung, 139, **140**
- Gericht, 2, 243, **251**
- Gerichtliche Schriftvergleichung, 16,
27, 47, 151
- Gutachten/
Sachverständigengutachten, 5, **140**
- privates, 8
- zum Autorschaftsnachweis, 54,
139, **140**
- zum Verständnisnachweis, 54,
139
- Hauptverhandlung, 201, 230
- Kommentar
- Lehr-, 143
- Praxis-, 143–144
- Kriminalpolizei, 64
- Landeskriminalamt (LKA), 4,
Linguistik, 17
- Forensische, 78, 265, 266,
267
- liquet/non-liquet*, 58, 59, 60, 62, 63,
189, 206, **224**
- Merkmal
- -ausprägungskonfiguration,
180
- -konfiguration, 180
- linguistisches, 75
- non-liquet see liquet*
- Oberlandesgericht (OLG), 19, 56
- Praxisfeld, 32, 52, 84, 238
- Richter, **131**
- Rote Armee Fraktion (RAF), 8,
142
- Sachverständigenbeweis, 243, 244,
268
- “Würdigung des
Sachverständigenbeweises
durch das Gericht”, 7, 47, 134,
243
- Sachverständige/r, 8, 268
- Schibboleth, 19, 86–109
- Schimpfwort (Konkubine), 9, 19,
125–136, 142, 144–148
- Strafprozessordnung (StPO), 8

Tatsachenbehauptung, 121–123

Tatschriften/Tatschreiben (TS), 5, 41,
42, 166, 167, 169–170, 172,
178–179, 182, 183, 184, 185, 188,
189–190

Üble Nachrede, **116, 119, 121, 122**

Verbalinjurie, 12, **125**, 126, 144

Vergleichsschreiben (VS), 5, 167,
173

Widerruf, 4, 58, 139, **140**

Zusammenhang, 137

Subject Index English

[Numbers appearing in brackets following page numbers suffixed with “n” or “nn” indicate the chapter number followed by the relevant note cue number(s)]
(Numbers appearing in **bold** indicate that the pages contain diagrams/figures/illustrations/tables)

- accent, 88
- ambiguity, 173, 183, 184, 189, 266
- analysis
 - authorship, 17, **28**, 39, 41, 69, 71–85, 167, 172, 173, 176, 190, 191, 192–234
 - forensic, 31
 - handwriting, 5, 16, 27, **28**, 30, 36, 37, 38, 47, 74, 84, 151, 166, 170, 186, 246
 - phonetic, 31
- anonymous
 - authorship, 1, 17, **28**, 39, 61, 65, 79–81, 166, 167–170, 172, 173–176, 185, 190, 192, 193, 205, 229, **233**, 234, 254, 255
 - writership, 190, 218, **224**
- argumentation, **125**, **129**, 133, 135, 142, 256, 269, 270
- attention style, 75
- attitude, 31, 80, 93, 95, 96, 106, 143, **214**, 215, 248
- authenticity, 41, 46, 59, 170, 197, 241
- author identification, 241, 247, 254
- authorship
 - analysis, 76, 79–81, 167–172, 173–175, 176–191, 192, 210
 - anonymous, 1, 13, 17, 18, **28**, 39, 61, 65, 79–81, 166, 167–170, 173–176, 185, 189, 190, 192, 193, 205, 218, **233**, 234, 258
 - attribution, 4, 10–14, 18, 22, **28**, 29, 42, 44, 47, 51, 54, 62–70, 78, 79, 80, 82, 89, 162, 166–190, 217, 252, 255, 266, 267
 - disguised, 11, 21, **153**
 - disputed, 4, 58–59, 61
 - identification, 67, 83, 190, 247, 252
 - multiple, 11, 13, 14, 42, 61, 63, 151, **156**, 252
- auxiliary science, 17, 27, 32, 70, 73, 78, 79, 183, 271
- behavior
 - language, 18, 19, 21, **29**, 37, 62, 64, 72, 73, 81, 90, 100, 102, 107, 108, 126, 169, 173–181, **214**, 219, 241
 - linguistic, 19, 36, 39, 48, 58, 62, 65, 67, 68, 69, 80, 82, 83, 86–109, 127, 167, 168, 176, 180, 185, 196, 211, 224
 - orthographic, 45, 82, 167–171, 173, 174, 175, 176–181, 182–190, 208–215, 216–225, **233**, 234
 - semiotic, 190, **239**
- bias, 256
- blackmail, 84, **154**, **156**, 164, 193, 201, 225–233
- capitalization, 82, 185
- case, 4–9, 10–14, 19, 54–62, 63–70, 84, 85, 113, 115, 118–121, 124–132, 143–144, 194–208, **227**, 231, 241, 243, 247–258, 262–263, 267–270
- coherence, 163
- cohesion, 163
- comparison
 - data, 5, 10, 40–42, 59–61, 65, 74, 76, 149, 166, 201–205, 215, 222
 - text, 5, 59, 61, 72, 74, 167, 194, 231, 234
- compound, 104–113, 183, 198, **199**, 250, 258–261

- comprehension, 6, 10
 confession, 61, 63, 142
 context
 - criminal, 61, **181**
 - extra-linguistic/situational, 5, 9, 11, 12, 56, 66, 115, 126, 127, 132, 139, 145–148, 243
 - forensic, 82, 172, 238, 240
 - linguistic, **131**, 132, 145, 146, 148, 177, 223, 262
 continuum
 - of diagnostic potential, 167, 211–212
 - heuristic, **182**, 209, 225
 - of orthographic behavior/data/features, 22, 176, 177, 208–210, 213–224
 - of verbal interaction, 241–242
 co-occurrence, 32, 158, 180, 229
 corpus, 37, 170, 191, 212
 covariation, 100, **179**, 205, 230
 criminal court procedures, 8, 74
 criminology, 84, 88, **162**, 163
 cumulative principle, 77, 234
 data
 - idiosyncratic, 77, 185, 189, 219
 - language, **28**, 39, 65, 70, 74, 170, 240, 244, **249**, 257, 260
 - linguistic, 5, 39, 82, 126, 166, **168**, **182**
 - observable, 80, 174, 221–225, 246
 - orthographic, 22, 161–191, 192–234
 - phonetic, 39, **168**
 - text-external, 79, 184, 188, 190
 - text-internal, 187, 188, 190
 defamation
 - law, 1, 19, 20, 115, 116–125
 - letter, 63, **154**, **160**, **181**, 184, 186, 198, 205, 206, 222, 228, 232, 248, 254
 definition, 9, **28**, 75, 78, 92, 95, 97, 101–107, 108, 126, 130, 158, 163, 164, 165, 192, 219, 228, 232, **233**
 descriptiveness, 31, 82, 91, 95, 132, 169, 176, 184, 185, **212**, 232, 241, 246, 257
 deviation, 60, 66, 67, 82, 94, 171, 173, 175–179, 189, 203, 229
 diagnostic potential, 166–191, 192–213, 230, 232, 234
 dialect, 67, 87, 91, 102, 103, 104, 105, 106
 difference of the legal systems, 2–8, 84–85, 121–125, 133–144, 230, 267–271
 differential diagnosis, 70, 82–83, 171–172, 185
 discourse, **181**
 disguise, 37, 38, 43, 88, 100, 102, 152–153, 154–155
 - of authorship, 11, 21, 44, 69, 158, 172, **182**
 - of writership, 21, 158, 161, 187, 196, 203, 252
 Distinctive Text-Profile-Matrices (DTPM), **156**, 159
 DNA analysis, 32, 136, 142, 247
 documentation
 - of cases, 42, 84–85, **157**, 269
 - of forensic linguistic expert opinions, 6, 14–15, 32, 37, 48, 141
 errors, 44, 46, 60, 66–69, 81, **160**, 170, 171, 173, 175, 198, 203–204, 207, 211, 232
 ethics, professional, 253–256
 ethnography/ethnographic, 61, 79, 92, 98, 99, 124, 175, **182**, 241
 evidence
 - circumstantial, **75**, **119**, 130, 138, 143, 144, 171, 210
 expert
 - analysis, **152**
 - opinion/testimony, 2, 3, 4–20, 22–23, **29–35**, 37–41, 44–48, 51–60, 62–70, 76–77, 79–80, 84–85, 115, **125**, 126, 128, 132–136, 139–141, 144, 149, 150, **152**, 173, 187, 190, 192, 197, 201, 218, **221**, 225, **227**, 230, 234, 238–243, 245, 247, 252, 255, 256, 258, 264–266, 268–270

expert – *continued*

- witness, 51–70, 133, 139–143, 268, 271

extortion

- letter, 11, 13, 63, 65, 161, 178, 180, 181, 187, 201–205, 213

feature

- attitudinal, 92, 93
- behavioral, 196, 239
- binary, 20, 21, 41, 154–158
- concomitant, 37, 68, 69, 95, 179
- configuration, 18, 45, 66–68, 82, 100, 166, 173, 180, 230, 231
- distinctive, 75, 83, 159, 160–162
- dynamic, 22, 45, 66, 68, 69, 230
- group-specific, 19, 108, 185, 216
- identifying, 66, 67, 69, 71, 88, 109, 196, 215
- idiolectal, 108, 219
- of linguistic behavior, 19, 36, 67, 69, 94
- orthographic, 180, 193, 205, 209, 210, 213–216, 217, 218, 220, 225–233
- phonetic, 104
- semantic, 128, 130
- speaker-specific, 36, 90, 100, 101, 174, 175
- text external, 163
- text type-specific, 68, 193
- text-pragmatic, 60
- value configuration, 180, 231

forensic

- handwriting analysis, 5, 7, 16, 27, 28, 30, 38, 84

linguistics, 1, 27–48, 51, 73–74, 78–79, 91, 113–116, 141, 149, 151, 152, 190–192, 230, 234, 246, 247, 254, 255, 265–271

- phonetics, 10, 16, 27, 28, 29, 35, 38–48

sciences, 16, 27–32, 35, 36, 46–48, 73, 169, 247

forensic linguistic

- authorship analysis/attribution, 17, 71–85, 172, 192–234
- cases, 54–61, 85, 194–208, 221

- expert, 4, 7, 35, 150, 151, 152, 162, 170, 194–208, 247, 252, 253, 265

- expert opinion/testimony, 2, 3–9, 14, 17, 20, 32, 35, 60, 62–70, 84–85, 133, 139–141, 144–148, 157, 230, 234

forensic phonetic

- speaker recognition/identification, 27, 39, 42, 107

genre, 36, 98, 100, 184, 185, 212, 213, 214, 219, 230

German legal context, 9–14

graphemic representation, 178, 181, 188, 204, 207

heuristic

- continuum, 182, 209, 225
- taxonomy, 19, 20, 92, 150, 151, 152, 154–155, 193, 209, 225, 238, 239

Hitler diaries, 170, 197

hyphen/hyphenization, 64, 82, 177, 181, 183, 206, 215, 220, 258, 259, 260

hypothesis, 46, 61, 91, 102, 187, 188, 193, 206, 208

idiolect, 17, 19, 29, 63, 69, 70, 72–73, 77, 82, 91, 106, 185, 213, 216, 218–220, 252, 260

idiom, 11, 62, 64, 66, 182, 207

idiosyncrasy coefficient (ISC), 22, 184, 186, 188, 189, 190, 193, 208–215, 216–225

illocutionary force, 114, 129

incriminating

- data, 41, 188
- letter, 5, 21, 154, 156, 162, 170, 181, 197, 203, 205–209, 214, 215, 227, 231, 232, 254
- text, 5, 11, 21, 38, 39, 41, 42, 64, 65, 70, 72, 76, 81, 152–165, 166–167, 181, 194, 209, 212, 216, 221, 222, 224, 225, 230–234

indicator, 61, 69, 75, 76, 95, 177, 184, 189, 200, 211, 214, 229, 230, 272n(4.2)

- insult, 9, 10, 11–12, 113–148, 240, 242, 243, 245–247
- interdisciplinarity, 10, 16, 23, 27–48, 53–54, 83–84, 135, 163–165, 237–240, 264, 271
- item
- linguistic, 81, 90, 98
- jargon, 20, 45, 46, 133, 140, 155, 243, 252, 268
- judge, 7–9, 20, 33–35, 47, 60, 129–131, 133–137, 142–143, 239–247, 252, 255–256, 270–271
- judgment, 42, 43, 70, 117, 121, 122, 123, 124, 126, 145, 226, 229, 245, 248
- judiciary, 17, 23, 29, 31, 33, 34, 35, 40, 46, 56, 73, 129, 133, 135–141, 143, 145, 164–165, 197, 242, 244, 250, 252
- justice, 119
- language
- data, 39, 65, 74, 77, 81, 130, 137, 170, 214, 230, 234, 240, 244, 245, 249, 257, 260
- system, 45, 80, 126, 137, 184, 205, 212, 267
- use, 12, 37, 66, 80, 90, 99, 126, 132, 136, 137, 145, 184, 185, 196, 197, 245–247, 257, 260
- language and law, 28, 29, 32, 53, 79, 238, 239, 266, 267, 271
- law
- common, 164
- criminal, 119, 120, 208
- lawyer, 8, 33, 59, 124, 135, 143, 237–257, 261, 264, 273n(9.1)
- lay people/lay person/laymen, 9, 39, 42–47, 70, 72, 77, 82, 92, 103, 124, 136, 140, 152, 214, 257, 258, 270
- legal
- language, 270
- profession, 253
- training, 135
- legislation, 117, 120
- lexicon, 47, 60, 91, 250, 261
- libel, 7, 10, 11–12, 20, 29, 77, 79, 84, 113–115, 116–118, 122, 123, 125–127, 139, 154, 156–157, 163, 225–229, 230, 231, 232, 242, 254
- linguistic
- behavior, 19, 36, 39, 48, 58, 62, 65, 67, 68, 69, 80–83, 86–109, 127–130, 167, 168, 171, 175, 176, 185, 211
- data, 11, 18, 39, 80, 82, 90, 126, 135, 168, 182, 190, 192, 197, 222, 234, 239, 247, 270
- differential diagnosis, 70, 82–83, 171, 185
- expert opinion/testimony, 4, 5–14, 18, 20, 39, 40, 44–48, 51–62, 64, 65, 68, 72, 77, 80, 84, 85, 89, 125, 128, 129, 132–135, 139, 140, 141, 144–148, 187, 230, 238, 240, 243, 268
- feature, 14, 47, 71, 72, 75, 98, 101, 163, 164, 209, 229
- “fingerprint”, 17
- linguistics
- applied, 3, 16, 18, 23, 28, 32, 40, 47, 52–53, 77, 86, 126, 163, 239, 265
- computational, 219, 266
- corpus, 27, 68, 272n(3.3)
- forensic, 27–48, 51, 73–74, 78–79, 114, 141–142, 149, 167, 218, 227, 230, 234, 246, 247, 254, 265–271
- general, 27, 28, 227, 239, 265
- psycho-, 19, 27, 28, 29, 31, 219
- socio-, 19, 28, 29, 31, 45, 76, 77, 79–80, 86, 95, 99, 115, 124, 126, 127, 138, 145, 215, 228, 266, 267
- systemic (vs. sociolinguistic), 30–31, 192
- text, 11, 168, 173, 184, 188, 239
- theoretical, 18, 29, 52–53, 142
- liquet/non-liquet*, 58, 59, 60, 62, 63, 189, 206, 224
- marker, 19, 66, 75, 76, 95, 97, 99–101, 106, 272n(4.1)
- meaning
- connotative, 114, 127, 129, 145, 242
- in context, 54, 131, 246

- meaning – *continued*
 - denotative, 56, 87, 114, 127, 129, 145, 242
 - word, 5, 89, 130, 232
see also semantics
- medium, 124
- metalinguistics, 226, 271
- metaphor, 27, 81, 269, 270
- method, 18, 34, 35, 52, 65, 69, 82, 83, 98, 100, 136, 215, 239, 246, 267
- methodology/methodological, 19, 22, 27, 30, 31, 33, 36, 44, 45, 46, 52, 58, 62–70, 76–77, 78, 86, 113, 127, 149, 167, 172, 177, 192, 197, 201, 205, 208–216, 217, 228, 230, 238, 239, 269
- mistake, 48, 58, 66, 67, 68, 76, 82, 107, 160, 171, 173, 175, 176, 177, 202, 232, 233
- misunderstanding, 23, 53, 106, 137, 237, 241, 242, 245, 247, 257, 264, 273n(9.1)
- morphology, 60, 83, 91, 184, 238, 257
- multiple authorship, 11, 14, 61, 63, 151, 156, 157, 252
- multiple writership, 151
- needs analysis, 237, 240, 243
- nominalization, 233
- non-native
 - speaker, 67, 105, 178, 194, 206
 - writer, 204
- non-verbal communication, 98, 120
- orthography/orthographic, 160, 169, 175, 177, 182, 185–189
 - behavior, 22, 45, 167, 168, 169, 171, 174–191, 205, 208–224, 229–234
 - data, 22, 150, 166–191, 192–234
 - norm, 179, 183
 - system, 184, 193, 194, 205, 212
- peculiarity status, 175–176
- phonology, 30, 67, 83, 91, 106, 159, 219
- police
 - criminal, 74, 142
- polyvalence/polyvalent, 173, 182, 183, 184, 193, 196
- pragmatics, 92, 134, 138, 142, 145
 precedent, 17, 266
- prescription/prescriptive, 229, 236
- probability scale, 35, 36, 60, 83
- proof, 54, 56, 58, 70, 83, 85, 124, 133, 244
- prosecution, 5, 197, 241, 268, 269
- proverb, 64, 73, 105, 207
- punctuation mark, 180, 181, 186, 215, 216, 218, 227, 229, 233
- quotation/quote, 23, 46, 64, 105, 181, 188, 189, 207, 215, 220, 221, 222, 223
- reading copy, 198, 200
- recording, 39, 41
- register, 45, 63, 182
- reliability, 43, 76, 82, 103, 203, 247
- result, 2, 3, 7, 35, 40, 43–47, 52, 59, 60–63, 69, 70, 77, 83, 129, 134, 168, 190, 200, 224, 247, 255
- rules
 - sociolinguistic, 228
- semantics, 12, 56, 64, 83, 142, 184
see also meaning
- semiotic, 97, 186, 187, 190, 193, 203, 239
- sense, 19, 54, 56, 90, 100, 143–145, 266, 267
- sentence
 - length, 101
- shibboleth, 19, 86–109, 272n(4.2)
- slander, 7, 10, 11–12, 20, 29, 115, 116, 117, 118, 122, 123, 242, 243, 252
- sociolect, 63, 91, 99
- sociolinguistic
 - analysis, 66, 127
 - variation, 76, 82, 83, 99, 115–116, 198, 200–201, 228
- sociolinguistics
 - applied, 126
- speaker-specific
 - behavior, 45, 72, 81–82, 98, 168
 - feature, 36–37, 80, 90, 91, 100, 101, 107, 108, 174, 175
- speech act, 129, 132, 145, 148, 163

- spelling, 64, 159, **160**, **179**, 189,
 203–208, **212**, **214**, 223, **224**, 227,
260
- stereotype, 76, 80, 95, 96, 97, 105,
214, 246
- stigmatism, 174
- style, 22, 69, **75**, 83, 95, 147, **155**,
 175, **182**, 213, **216**, **226**, 239, 243
- swearing, 126, 144
- syndrome, 37, 44, 65, 67, 81, 82, 175,
 176, **179**, 185–188, 191, 215, **239**
- synonym, 28, 127, 129, **131**, 146, 164
- syntax, 37, 47, 60, 64, 67, 83, 91, 184,
 223, 257
- taxonomy
 - heuristic, 19, 20, 92, 150, 151,
152, **154–155**, 193, 209, 225,
 238, **239**
- term/terminology
 - legal, 23, 59, 237, 243, 262
 - linguistic, 71, 74–76, 115, 134,
140, 242, 272n(4.1)
 - technical, **117**, 266
- test/testing, 37, 52, 61, 87, 93, 95,
 102, 108, 128, 130, **239**, **260**, 267
- text type, 19, 21–23, 39, 44, 58, 59,
 76, 81, 82, 100, 142, **156**, 159,
160, **162**, 163, 175, 209–215, **220**,
221, 225–233
- threat, 146, 163, 164, 174
- threatening letter, 5, 13, 142, **156**,
 163, 178, 180, **181**, 194–197, 201,
 218
- trademark
 - case, 46
 - law, **29**, 79, 139, **140**
- transcript, 46, 74, 170, 198, **200**, **251**
- typewriter analysis, 169
- typewriter script analysis, 169
- typist, 10, 13, 20, 41, 42, 149–165,
 167, **168**, 169, 170, 172, 175, 186,
 196, 206, 210
- validity, 38, 59, 76, 82, 102, 103, 136,
 166, **168**, 171, 184, 247, 250
- variation
 - linguistic, 99, 103, **200**
 - orthographic, **214**
 - text type-specific, 171
- verdict, 58, **125**, **131**, 132, 147, 252,
260, 262, **263**
- vocabulary, 59, **226**, 231, 257
- witness
 - expert, 51–70, 133, 268, 271
- written
 - data, 38, 149, 167, **221**
 - language, 91, 106, 169, 170, 171,
 173–182
- writer, 10, 13, 20, 41–43, 65, 132,
 149–165, 167, **168**, 169–172, 175,
 178, **179**, **181–190**, **195**, 196, 198,
200, 201, 203, 204, 206, **207**, 208,
 210, **214**, **216**, 217, 219
- writing
 - system, 186, 193, 204
 - use, 186, 191