

Stichprobenziehung für telefonische Zuwandererumfragen: Einsatzmöglichkeiten der Namenforschung

Humpert, Andreas; Schneiderheinze, Klaus

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Humpert, A., & Schneiderheinze, K. (2000). Stichprobenziehung für telefonische Zuwandererumfragen: Einsatzmöglichkeiten der Namenforschung. *ZUMA Nachrichten*, 24(47), 36-63. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-208052>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

STICHPROBENZIEHUNG FÜR TELEFONISCHE ZUWANDERERUMFRAGEN

EINSATZMÖGLICHKEITEN DER NAMENFORSCHUNG

ANDREAS HUMPERT UND KLAUS SCHNEIDERHEINZE

Im folgenden Beitrag wird ein Konzept für die Zuordnung von Sprachen zu Namen vorgestellt. Das Konzept basiert auf Erkenntnissen der Namenforschung (Onomastik). Für die telefonische Befragung von Zuwanderern können damit Listenauswahlen durchgeführt werden, die nur eingetragene Privathaushalte mit Festnetzanschlüssen enthalten. Für die Fehlerreduktion in diesen Auswahllisten werden für jede Sprache Zuweisungsregeln formuliert, die vor dem historisch-politischen Hintergrund der Zuwanderung gebildet werden. Für diese Regeln werden der Vor- und Nachname, sowie bei Vorhandensein auch der zusätzliche Telefonbucheintrag genutzt. Die Sprachzuweisung und Sammlung dieser Regeln werden anhand von Beispielen erläutert. Erste Listen aus einem Pilotprojekt des Jahres 1999 werden kurz vorgestellt.

The paper describes a procedure of assigning people's names to languages which is based on findings from onomastic research. Using this, we were able to carry out telephone interviews with new immigrants on the basis of telephone directory lists of the names of households with fixed line phones. In order to reduce the error rate of these lists, we defined rules of assignment for every language, taking into account the historical and political background for each group of immigrants. We used these rules for first and last names, as well as for any additional entry in the telephone directory. We illustrate the assignment of languages and the set of rules with examples. The lists developed in a pilot scheme in 1999 are briefly presented.

1. Einleitung

Die gezielte Befragung einer großen Zahl von Personen ausländischer Herkunft wird aufgrund der Vielzahl methodischer Probleme und des hohen Aufwandes (zum Beispiel Stichprobenziehung, muttersprachliche Befragung) nur selten durch-

geführt. Migrationsforschung findet oft auf der Grundlage von empirischen Methoden wie Sekundäranalysen¹, Leitfadeninterviews und Gruppendiskussionen statt. Wenn Primärerhebungen mit hohen Fallzahlen und standardisierten Erhebungsinstrumenten unter Zuwanderern durchgeführt werden, wird oftmals auf der Basis verschiedener Varianten mehrstufiger Auswahlverfahren die Befragung im Haushalt (face-to-face) bevorzugt. So wird zwar methodisch ein zufriedenstellender Qualitätsstandard erreicht, der jedoch zu hohen Kosten und langen Projektlaufzeiten führt. Die zeit- und kostengünstigere CATI-Methode wird bei Befragten ausländischer Herkunft nur sehr selten angewendet, da bislang keine Ansätze für zufriedenstellende Stichprobenkonzepte entwickelt wurden. Durch ein vorgeschaltetes Screening über Zufallsnummeranwahl könnten zwar die methodischen Einwände hinsichtlich einer uneingeschränkten Übertragbarkeit der Befragungsergebnisse auf die Grundgesamtheit (durch Einbeziehung der nichteingetragenen Haushalte) in Ansätzen ausgeräumt werden. Doch die spezifischen Vorteile (kurze Projektlaufzeiten und geringe Kosten) der CATI-Methode würden so beseitigt. Für Befragungen der immerhin über 10 Millionen² Personen ausländischer Herkunft besteht also in zweifacher Hinsicht eine ‚methodische Lücke‘. Die Stichprobenziehung ist entweder enorm zeit- und/oder kostenaufwendig, oder führt *nicht* zu einer verkleinerten Abbildung der Grundgesamtheit. Sowohl die Anforderung der Einhaltung eines ausreichenden Niveaus der Stichprobenqualität, wie die praktischen Erfordernisse einer Projektdurchführung (begrenzte finanzielle oder zeitliche Ressourcen) können also nicht erfüllt werden.

In den folgenden Erläuterungen soll ein Ansatz für die Ziehung von Zuwandererstichproben dargelegt werden, der auf der Zuordnung von Namen zu Sprachen basiert. Die Idee dieses Ansatzes ist nicht neu (vgl. Marhenke 1997: 212f.; Rösch 1998: 109; Salentin 1999: 126ff.) und wurde/wird auch schon mehrfach angewendet. Eine Neuheit ist vielmehr die hier vorgestellte praktische Umsetzung dieses Ansatzes, die durch folgende Punkte gekennzeichnet ist:

- Verknüpfung der Idee mit der Namenforschung (Onomastik),
- die gleichzeitige Verwendung von Vor- *und* Nachnamen, sowie der Partnereinträge (binationale Haushalte),
- die Entwicklung von spezifischen Zuordnungsregeln je Sprache unter Berücksichtigung der historisch-politischen Zuwanderungsbedingungen,

¹ Meist mit Daten aus der amtlichen Statistik oder den wenigen repräsentativen Bevölkerungsumfragen unter Einbeziehung der Zuwanderer.

² Einschließlich Eingebürgerter seit 1981 – Quelle: Statistisches Bundesamt, 31.12.99.

- die Anwendung auf eine größere Zahl von Sprachen³,
- die Bearbeitung des gesamten öffentlichen Telefonverzeichnisses.

Mit der spezifischen Umsetzung dieses namenbasierten Verfahrens soll ein Ansatz für Zuwandererstichproben geliefert werden, der die praktische Durchführbarkeit und die theoretischen Qualitätsanforderungen möglichst unverzerrter Stichproben⁴ gleichermaßen nicht aus den Augen läßt. Dieser Ansatz stellt dennoch keine ideale Lösung dar, weil es sich um eine Listenauswahl handelt. Die Liste ist unvollständig, und es ist bisher ungeklärt, ob sich die ausgeschlossenen Teile der Grundgesamtheit systematisch von den Elementen der Liste unterscheiden. Diese Einschränkung durch die Unvollständigkeit der Auswahlliste (öffentliches Telefonverzeichnis) kann jedoch durch zusätzliche Maßnahmen verringert werden. Zum jetzigen Zeitpunkt kann nur das Verfahren erläutert werden. Empirische Belege für die Effektivität (korrekte Zuordnung der Sprache) des Verfahrens und die Stichprobenqualität aufgrund der Listenauswahl (Nichteingetragene) können in der Entwicklungsphase eines solch komplexen Systems von Zuweisungsregeln⁵ noch nicht geliefert werden. Anhand der Daten eines Pilotprojektes aus dem Jahre 1999 soll vielmehr nur die Vorgehensweise veranschaulicht werden, um einen ersten Eindruck vom Potential des Verfahrens zu bekommen.

Im zweiten Textteil werden zunächst die Probleme und Möglichkeiten einer Zuordnung der Sprache auf Basis der Information des Namens behandelt. Im dritten Teil werden dann diese theoretischen Problemstellungen auf praktische Maßnahmen zur Bildung einer Auswahlgrundlage für telefonische Befragungen von Zuwanderern übertragen. Abschließend soll im vierten Abschnitt ein erster Eindruck von der praktischen Anwendung des Verfahrens auf das gesamte Telefonverzeichnis (1999) vermittelt werden.

2. Überlegungen zur Nutzung der Namenforschung (Onomastik)

Die Namenforschung (Onomastik), hier die Personennamenforschung (Anthroponomastik) befaßt sich seit langem mit Vor- und Nachnamen. Neben Herkunft, Be-

³ Bisher wurden von der Humpert & Schneiderheinze GbR 58 Sprachen bearbeitet.

⁴ Die Problematik der Privathaushalte ohne Eintrag im Telefonbuch kann hier nur kurz angesprochen werden.

⁵ Es mußten circa 200 Millionen Sprachencodes nach über 3.000 Kombinationen aus Vornamen, Nachnamen und Zusatzfeld vergeben werden.

deutung und Etymologie der Namen stehen historische Fragestellungen im Vordergrund. Dies kann zum Beispiel die Ermittlung von Handwerksberufen in einzelnen Orten im Mittelalter oder der frühen Neuzeit sein, die anhand der entsprechenden Berufsbezeichnungen in den Nachnamen möglich ist (vgl. Kunze 1998).

2.1 Die Entstehung der Familiennamen

Die Entstehung von Nachnamen begann in Italien und Frankreich im Hochmittelalter, kurz darauf in den übrigen west- und mitteleuropäischen Staaten. Von einzelnen Gebieten abgesehen (Friesland, Teile Skandiaviens) hatte sich die Kombination von Vor- und Nachnamen bis zum Ende des Mittelalters in diesen Teilen Europas durchgesetzt. Aufgrund des hohen Alters der Nachnamen und der meist bis ins 19. Jahrhundert hinein nicht festgelegten Schreibung kommen viele Namen in verschiedenen Schreibvarianten vor.⁶ Der Fundus besteht in vielen europäischen Sprachen aus mindestens hunderttausend verschiedenen Nachnamen. Hinzu kommen Namen aus Dialekten oder heute kaum noch gesprochenen Sprachen.⁷ Demgegenüber sind Schreibvarianten in Staaten mit umfassenden Sprachreformen vor der Einführung von Nachnamen praktisch nicht vorhanden. Dies gilt zum Beispiel für die Türkei. Das europäische Modell der Kombination von Vor- und Nachnamen ist im Laufe der Zeit auch von vielen außereuropäischen Staaten mit anderen Namentraditionen (z.B. islamischen Staaten) übernommen worden. Dies ist gesetzlich in Algerien bereits 1882, in Tunesien 1925, im Iran 1932, in der Türkei 1926/1934 und in Ägypten 1970 festgelegt worden. In Japan wurden 1868 Familiennamen, die bis dahin nur dem Adel, privilegierten Handwerkern und Künstlern vorbehalten waren, auch der übrigen Bevölkerung gestattet. 1875 wurden sie per Gesetz für jeden zur Pflicht. Die späte Einführung von Familiennamen führte dort zu der vergleichsweise geringen Zahl von knapp 10.000 verschiedenen Nachnamen (vgl. Kluge 1985: 122).

2.2 Differenzierung nach Sprachen

Für den hier relevanten Zweck müssen die Familiennamen notwendig eine Bedingung erfüllen: Aus ihnen muß sich mit hinreichender Genauigkeit auf die jeweilige

⁶ Siehe zum Beispiel: Meier, Maier, Meyer, Mayer, Mair, Mayr, Majer und: Schmidt, Schmid, Schmitt, Schmied, Schmitz.

⁷ Schottisch, Walisisch und Cornisch in Großbritannien, Flämisch, Bretonisch, Provenzalisch und Baskisch in Frankreich, Katalanisch und Baskisch in Spanien, Friesisch und Sorbisch in Deutschland, Sardisch und Friaulisch in Italien.

Sprache schließen lassen, aus der dann eine Staatsangehörigkeit bzw. Herkunft gefolgert wird. Entscheidend ist hier der Anteil der Namen, die in mehr als einer Sprache vorkommen. Hier lassen sich grundsätzlich zwei Fälle unterscheiden:

- a) mehrfaches Vorkommen aufgrund struktureller Verwandtschaft der Sprachen oder Dialekte,
- b) mehrfaches Vorkommen aufgrund zufälliger Übereinstimmung.

Der erste Fall ist zum Beispiel bei Deutsch und Niederländisch, Dänisch und Norwegisch, Katalanisch und Provenzalisch (Okzitanisch) oder Schottisch und Irisch gegeben. Hier läßt sich die Zuordnung oft durch zusätzliche Betrachtung des Vornamens entscheiden.

Der zweite Fall kann im Prinzip zwischen allen Sprachen auftreten, die nicht näher miteinander verwandt sein müssen. Er ist aber im wesentlichen auf kurze Namen beschränkt, bei denen die Wahrscheinlichkeit zufälliger Übereinstimmungen steigt. Auch hier bietet sich die Hinzunahme des Vornamens an, zumal unmittelbar einsichtig ist, daß die Wahrscheinlichkeit gleicher Vornamen bei nicht näher verwandten Sprachen (etwa Italienisch und Japanisch) äußerst gering ist.

2.3 Ermittlung der Namen

Um das mehrfache Vorkommen von Nachnamen möglichst umfassend erkennen zu können und die Genauigkeit der Zuordnung zu den jeweiligen Sprachen zu erhöhen, ist die Zusammenstellung eines Fundus aus möglichst vielen Namen und Sprachen notwendig. Aus den Erträgen der Namensforschung bieten sich zunächst Lexika an, die jeweils bis zu 75.000 Namen enthalten können. Die seit 1950 erscheinende Zeitschrift ‚Onoma‘ (Journal of the International Council of Onomastic Sciences) bietet in ihrem bibliographischen Teil den besten Überblick über die Vielfalt der Forschung. Von Vorteil ist, daß das Alter der Literatur hier keine Rolle spielt. Allerdings ist sie meist in der jeweiligen Landessprache abgefaßt und oft schwierig zu beschaffen. Bislang konnten wir⁸ rund 500.000 Nachnamen (einschließlich Fehlschreibungen und Doppelnamen) in 58 Sprachen und Dialekten zusammentragen und in eine Datenbank eingeben (siehe Tabelle 1).

⁸ Seit Frühjahr 1999 beschäftigen sich die Verfasser freiberuflich mit dem Aufbau einer Auswahlgrundlage in Verbindung mit der Onomastik. Zu Beginn des Jahres 2000 erfolgte die Gründung einer GbR im Bereich Sozial- und Umfrageforschung mit Schwerpunkt Nicht-deutsche und Migration.

Tabelle 1: Anzahl der Nachnamen (Stand: September 2000)*

Sprache	Anzahl	Sprache	Anzahl
Albanisch	1.588	Russisch	5.266
Aschkenasisch (Rußland)	81	Schottisch	3.142
Baskisch	844	Schwedisch	1.344
Bretonisch	562	Schweizerdeutsch	375
Bulgarisch	206	Schweizerfranzösisch	366
Cornisch	188	Serbokroatisch	20.030
Dänisch	583	sonstige slawische Namen/früher.UdSSR.	40.066
Deutsch	199.575	Sorbisch	829
Englisch	16.007	Tschechisch	2.206
Estnisch	15	Tiroler Dialekt (Österreich)	4.135
Finnisch	918	Türkisch	41.014
Flämisch	1.431	Ukrainisch	620
Französisch	11.854	Ungarisch	1.114
Friesisch	2.092	Walisisch	467
Griechisch	18.379	Wallonisch	1.213
Irisch	4.315		
Isländisch	249	<i>außerhalb Europas:</i>	
Italienisch	43.855	afrikanische Namen	660
Kastilisch (Spanisch)	5.510	Amharisch (Äthiopien)	662
Katalanisch	1.261	arabisch-islamische Namen	17.375
Lettisch	30	Armenisch	322
Litauisch	476	Chinesisch	541
Manx (Isle of Man)	50	Farsi (Persisch)	1.840
Neu-Hebräisch (Iwrit)	178	Georgisch	462
Niederländ.	19.192	indische Sprachen (Hindi, Tamil etc.)	4.164
Norwegisch	540	indonesische Sprachen	43
Polnisch	18.053	Japanisch	2.132
Portugiesisch	3.374	Koreanisch	144
Provenzalisch (Okzitanisch)	2.036	Thai	167
Rumänisch	1.008	Vietnamesisch	1.780

*) Einschließlich Fehlschreibungen, Transkriptionsvarianten und Doppelnamen.

Bei einem mutmaßlichen Gesamtbestand von 800.000 bis 900.000 verschiedenen Nachnamenseinträgen von Privathaushalten in den öffentlichen Telefonverzeichnissen Deutschlands entspricht dies etwa der Hälfte der vertretenen Nachnamen. In bezug auf die daran geknüpften Adressen sieht dieses Verhältnis deutlich besser aus,

da häufigere Namen zum weitaus größten Teil erfaßt sind. Mit den genannten 500.000 Nachnamenseinträgen konnten insgesamt 23,3 Millionen Privathaushalte codiert werden.

2.4 Vornamen

Zusätzlich wurden bislang rund 40.000 Vornamen erfaßt und eingegeben. Die Vornamen unterscheiden sich von den Nachnamen vor allem dahingehend, daß bereits wenige hundert Vornamen von der Mehrheit der Personen eines Landes getragen werden. So ist beispielsweise der Vorname Maria in Spanien wie in Lateinamerika der mit Abstand häufigste weibliche Vorname (Homge 1988: 101). Auch in Italien ist Maria der häufigste weibliche Vorname. Zu Beginn der 80er Jahre trugen 8,4 Prozent der Italienerinnen diesen Vornamen, mit 6,2 Prozent war Guiseppa der häufigste männliche Vorname (De Felice 1982: 332f.). Die Vornamen haben gegenüber den Nachnamen grundsätzlich eine geringere Aussagekraft, da mittlerweile in den meisten (europäischen) Staaten die weitgehende Freiheit der Namenswahl gesetzlich garantiert ist. Daher erhält der Vorname seine Bedeutung hier vor allem als *zusätzliches* Entscheidungskriterium in dem Fall, daß der Nachname in zwei oder mehr Sprachen vorkommen kann. Bei eindeutigen Nachnamen kann er auch als *verifizierendes* Kriterium dienen. Aufgrund der Wahlfreiheit der Vornamen kann hier die Zuordnung zu Sprachen nicht vollständig aufrechterhalten werden. Neben denjenigen Vornamen, die sprachlich etwa als deutsche, französische oder schwedische zu betrachten sind, wurden deshalb auch die Vornamen einer Sprache zugeordnet, die im betreffenden Staat häufig sind, die sogenannten Modenamen. Deshalb werden zum Beispiel französische Vornamen (zum Beispiel Denise, Nicole, Simone) und skandinavische Vornamen (zum Beispiel Sven, Lars, Ingrid und Helga) auch als deutsche gewertet. Als Hilfsmittel stehen „Hitlisten“ der am häufigsten vergebenen Vornamen zur Verfügung, die in vielen Staaten jährlich nach Geschlechtern getrennt ermittelt werden.

2.5 Schreibvarianten

Ein Problem stellen Fehler in den Telefonverzeichnissen dar.⁹ Zum einen handelt es sich hier um Schreibfehler bei der Aufnahme der Namen (Buchstabendreher etc.). Soweit diese offensichtlich sind, können sie in die Namenlisten aufgenommen werden. Hinzu kommt vor allem bei Zuwanderern die Vertauschung von Vor- und Nachnamen. Auch dies läßt sich durch einen entsprechenden Abgleich mit den

⁹ Fehler enthalten *alle* Telefonverzeichnisse.

Namenlisten beheben. Aufgrund verschiedener oder falscher Transkriptionen aus nicht-lateinischen Alphabeten entstehen häufig Fehlschreibungen der Nachnamen. Der erste Punkt bezieht sich etwa auf griechische Namen, deren Transkription nach dem Muster des Altgriechischen wie des Neugriechischen erfolgen kann. Dies führt zu unterschiedlichen Ergebnissen, da die Namen allesamt dem Neugriechischen angehören. Außerdem findet sich die Verwendung unterschiedlicher Transkriptionssysteme in verschiedene Sprachen (vor allem ins Deutsche, Englische und Französische). Dies gilt ebenso für Sprachen, die kyrillische Alphabete verwenden. So wird der russische Name „Puschkin“ (deutsche Transkription) im Englischen mit „Pushkin“ und im Französischen mit „Pouchkine“ oder „Puchkine“ wiedergegeben. Alle genannten Varianten finden sich in deutschen Telefonbüchern. Außerdem werden auch wissenschaftliche Transkriptionssysteme verwendet (der genannte russische Dichter schreibt sich dann „Puskin“, da das diakritische Zeichen über dem 's' nicht dargestellt werden kann). Dabei entsteht dann zusätzlich das Problem, daß die gerade von letzteren reichlich verwendeten diakritischen Zeichen von den verwendeten Zeichensätzen bis auf wenige Ausnahmen nicht wiedergegeben werden (können). Gleiches gilt für Namen aus Sprachen mit lateinischen Buchstaben, die nicht im deutschen Alphabet enthalten sind (z.B. Polnisch, Türkisch, Tschechisch, Slowakisch, Kroatisch). Auch die mit ASCII/ANSI-Zeichensätzen darstellbaren Buchstaben des Französischen (é, è, ê, É etc.), skandinavischer Sprachen (å, æ, ø, Å, Æ, Ø) oder des Spanischen und Portugiesischen (ã, ñ, Ã, Ñ) tauchen nur gelegentlich auf. Meist werden sie durch den Grundbuchstaben oder eine Buchstabenkombination (æ wird zu ae) wiedergegeben. Selbst die Umlaute ö, ü, Ö und Ü werden zum Beispiel in türkischen Namen häufig durch oe, ue, Oe und Ue oder den Grundbuchstaben (o und u) ersetzt. Bei arabisch-islamischen und indischen Namen ist die Zahl der Fehlschreibungen und Schreibvarianten so hoch, daß sie die Mehrheit sowohl der Nachnamen wie der Adressen ausmachen. Die Einbeziehung solcher Einträge also ist aufgrund der relativen Häufigkeit¹⁰ notwendig.

2.6 Sprache und Staatsangehörigkeit

Das Verhältnis von Name, Sprache und Staatsangehörigkeit läßt sich aus verschiedenen Perspektiven betrachten. Schon wegen eingebürgerter Zuwanderer kann eine Sprache trotz eindeutigen Hinweis auf *einen* Staat nicht mit der betreffenden Staatsangehörigkeit gleichgesetzt werden. Ein Name kann einer oder mehreren Sprachen

¹⁰ Die griechischen Nachnamen in den Telefonverzeichnissen bestehen mehrheitlich aus derartigen Varianten.

angehören. Namen aus verschiedenen Sprachen können andererseits in einem Staat seit langem heimisch sein. Und eine Sprache kann auf verschiedene Herkunftsstaaten hinweisen.

Einbürgerung

Häufig wird die Zielgruppe bei Befragungen von Zuwanderern durch Personen ohne deutsche Staatsangehörigkeit definiert sein. Für die Konstruktion einer Stichprobe kommt es also darauf an, von der Sprache der Namen möglichst genau auf eine Staatsangehörigkeit schließen zu können. Da eingebürgerte und nicht-eingebürgerte Zuwanderer durch den Namen nicht zu unterscheiden sind, enthält die Auswahlgrundlage zwangsläufig beide Personengruppen. Mit dem Begriff der ‚nationalen Herkunft‘ wäre die Zuordnung also genauer beschrieben.

Für viele Fragestellungen dürfte die Einbeziehung der eingebürgerten Personen sinnvoll sein. Hier ergibt sich also ein Vorteil zum Beispiel gegenüber Melderegisterstichproben, die dies nicht erlauben. Andererseits sind Fragestellungen denkbar, etwa im Rahmen des Wahlverhaltens, die ausschließlich auf Eingebürgerte abzielen. Mit der vorgestellten Methode ist eine Stichprobenziehung möglich, die genau diese Gruppe enthält. Zwar erhöht sich durch die gleichzeitig enthaltenen nicht eingebürgerten Personen der Anteil der Ausfälle, doch ist kaum eine andere Möglichkeit zur gezielten Befragung dieser Spezialgruppe denkbar.

Von 1982 bis 1998 sind in Deutschland rund 2,5 Millionen Personen eingebürgert worden. Davon entfällt allein die Hälfte auf Bürger der früheren Sowjetunion beziehungsweise ihrer Nachfolgestaaten. Die Anteile der Eingebürgerten an der Gesamtzahl der Gruppen einer ‚nationalen Herkunft‘ liegt bei den Türken bei etwa 10,5 Prozent, bei den Österreichern bei 6,5 Prozent und den Italienern bei 2,5 Prozent. Vor allem bei Nicht-EU-Bürgern ist mit einer weiter anhaltenden Einbürgerungstendenz zu rechnen. Die Eingebürgerten führen also bei Befragungen von Personen ausschließlich dieser Staatsangehörigkeiten nicht zu nennenswerten Ausfällen. Die Befragung nur von Eingebürgerten ist bei Anteilen von 2,5 Prozent nur unter hohen Ausfällen zu realisieren. Die Befragung von eingebürgerten Türken erscheint dagegen – gerade angesichts weiter steigender Einbürgerungszahlen - mit vertretbarem Aufwand realisierbar.

Migrationsgeschichte in Deutschland

Der Schluß auf Staatsangehörigkeiten beziehungsweise Herkunft macht die Berücksichtigung der Migrationsgeschichte erforderlich. Befragungen von Nichtdeutschen erstrecken sich bislang vor allem auf Angehörige und Nachfahren aus Staaten (Ita-

lien, Griechenland, Türkei, Spanien, Portugal und früheres Jugoslawien), mit denen ab 1955 Anwerbeabkommen geschlossen worden waren. Unter Zuwanderern verstehen wir hier Personen und ihre Nachkommen, die nach 1945 nach Deutschland gekommen sind. Da es aber seit der Verbreitung der Nachnamen im Spätmittelalter immer wieder teilweise umfangreiche Einwanderungen gegeben hat, müssen zunächst Träger nichtdeutscher Familiennamen, die bereits seit Generationen Deutsche sind - soweit möglich - erkannt werden, um die Nachfahren dieser früheren Immigranten ausschließen zu können. Dies gilt zum Beispiel für die Hugenotten, die im 16. und 17. Jahrhundert in großer Zahl aus Frankreich, teils auch aus Norditalien (Savoyen) gekommen sind. Nach den großen Bevölkerungsverlusten während des 30jährigen Krieges haben sich viele Landesherren um Neuansiedler bemüht. So siedelten sich beispielsweise in Mecklenburg Schweden an, in Süddeutschland Kroaten. Besonders umfangreich war die Einwanderung in Preußen. Die preußischen Könige bemühten sich nicht nur erfolgreich um die Ansiedlung der vertriebenen Hugenotten, sondern zum Beispiel auch um Siedler aus den Niederlanden. Oft haben sich diese Menschen binnen zwei oder drei Generationen assimiliert und tragen seit langem deutsche Vornamen. Zum Teil wurden zwischenzeitlich auch ihre Nachnamen der deutschen Schreibweise angepaßt.

Hier erweisen sich die Ergebnisse der historischen (Namen)Forschung als hilfreich, die es erlauben, einen sicheren Bestand solcher Namen zu verwenden (zum Beispiel Familiennamen hugenottischer oder italienischer Immigranten) (vgl. Zamora 1992). Auch die nicht so weit zurückliegende Immigration von Polen in die Industriezentren des kaiserlichen Deutschland läßt sich am Namen erkennen, da auch deren Nachkommen mittlerweile deutsche Vornamen tragen oder ihre Nachnamen der deutschen Schreibung angepaßt haben. Aus „Szymanski“ wurde so „Schimanski“. Deshalb ist also anzunehmen, daß die Kombination eines polnischen Nachnamen in polnischer Schreibung und eines polnischen Vornamen auf einen Polen verweist, der erst in jüngerer Vergangenheit nach Deutschland gekommen ist, also einen Zuwanderer im oben genannten Sinne. Ein Deutscher, dessen Vorfahren mindestens seit Ende des 19. Jahrhunderts Deutsche waren, kann also unter anderem einen deutschen, einen friesischen, einen niederländischen, einen französischen, einen italienischen, einen sorbischen oder einen polnischen Nachnamen tragen. Im Falle von deutschen und sorbischen Namen ist bei der Kombination mit einem deutschen (oder in Deutschland häufigeren) Vornamen von einer Person mit deutscher Staatsangehörigkeit auszugehen. Unter den friesischen Namen lassen sich ost- und nordfriesische (also deutsche) und westfriesische (also niederländische) Namen unterscheiden sowie eine Gruppe von Namen, die auf beiden Seiten der Ems vorkommt.

Die französischen Namen der nach Deutschland emigrierten Hugenotten (siehe oben) sind zu einem großen Teil bekannt.

Die Kombination eines nichtdeutschen Nachnamen und eines deutschen Vornamen läßt grundsätzlich nur drei Möglichkeiten zu:

- a) die Vorfahren sind bereits seit langem Deutsche,
- b) es handelt sich um Angehörige der 2./3. Migrantengeneration,
- c) bei weiblichen Vornamen kommt die Möglichkeit einer binationalen Ehe hinzu.

Migrationsgeschichte in den Herkunftsstaaten

Neben früheren Einwanderern, deren Nachkommen seit langem Deutsche sind und die lediglich an ihrem Nachnamen noch als solche zu erkennen sind, ist auch die Zusammensetzung der Bevölkerung in den Heimatstaaten relevant. So ist bei albanischen Namen angesichts der geringen Zahl von albanischen Staatsangehörigen in Deutschland (rund 12.000)¹¹ davon auszugehen, daß die Träger dieser Namen eher aus dem Kosovo (also Jugoslawien) stammen. Ein französischer Nachname kann auf einen Franzosen wie auf einen Belgier, einen Schweizer, einen Franko-Kanadier, aber auch auf eine Person aus einer früheren französischen Kolonie verweisen. Auch in diesem Fall lassen sich weitere Informationen nutzen. So läßt sich zum einen wiederum der Vorname heranziehen, was vor allem bei Afrikanern häufig zur Klärung führt. Andererseits gibt es französische Nachnamen, die eindeutig aus bestimmten Regionen stammen, etwa dem französischsprachigen Teil der Schweiz oder Wallonien (Belgien).

Die Bevölkerung der klassischen Einwanderungsstaaten USA, Kanada, Australien, Neuseeland und Südamerikas hat ihre Nachnamen bei der Emigration mitgenommen. Wenn diese bis heute unverändert geblieben sind, sind sie nicht von denen der Ursprungsstaaten unterscheidbar. Sie werden teilweise erkennbar, wenn man den Vornamen hinzuzieht. Ein Teil der Emigranten hat allerdings ihren Namen mehr oder weniger stark angepasst. Aus „Steinweg“ wurde „Steinway“. Ein bei bloßer Betrachtung des veränderten Namens nicht mehr zu erkennender Ursprung liegt vor, wenn aus dem polnischen Namen „Adamski“ „Adams“ oder aus „Jachimowski“ „Jack“ wird (Kaleta 1998/99: 115). Eine große Zahl solcherart veränderter Namen kommt in den Herkunftsstaaten gar nicht vor. Es handelt sich dabei um neue Namen, die damit für das Einwanderungsland charakteristisch sind. Dies gilt beispielsweise für die US-amerikanischen Namen „Chrysler“ (aus Kreisler), „Cron-

¹¹ Quelle: Statistisches Bundesamt; Ausländerzentralregister; Stichtag: 31.12.1999.

kite“ (aus Krankheit) oder „Eisenhower (aus Eisenhouwer). Auch in Lateinamerika gibt es Neubildungen portugiesischer und spanischer Nachnamen. Das Problem der Zuordnung reduziert sich damit um solche Namen, deren Herkunft aus den Einwanderungsstaaten mit Hilfe der Namenforschung geklärt werden kann.

2.7 Kategorisierung am Beispiel italienischer Namen

Zunächst soll anhand der italienischen Nachnamen die Zuordnungsproblematik veranschaulicht werden. Dazu werden Daten aus dem Jahr 1999 genutzt.

Von den bislang aus fünf verschiedenen Quellen zusammengetragenen 43.855 italienischen Nachnamen (einschließlich Fehlschreibungen und Doppelnamen) kommen nach momentanem Kenntnisstand lediglich 1.350 (3,1 Prozent) auch in anderen Sprachen vor (siehe Tabelle 2), davon 386 in mehr als zwei Sprachen. Die Übereinstimmungen bestehen vor allem zu verwandten Sprachen (Kastilisch [Spanisch], Katalanisch, Portugiesisch, Französisch, Provenzalisch). Die Übereinstimmungen mit germanischen Sprachen (Deutsch, Niederländisch, Englisch) gehen überwiegend auf norditalienische Namen zurück, denen die „typischen“ Endungen italienischer Nachnamen fehlen. Von den insgesamt 43.855 gesammelten Namen sind aber lediglich 26.950 (61 Prozent) auch in dem bearbeiteten Telefonverzeichnis vertreten. Als Grund hierfür kommt einerseits in Betracht, daß selbstverständlich eine Zuwanderergruppe einen kleineren Namenpool aufweisen muß als die Bevölkerung des Heimatlandes. Auch sind unter den von der Namenforschung zusammengetragenen Namen viele im Heimatland seltene, teilweise auf kleinräumige Gebiete beschränkte Familiennamen. Von den im Telefonverzeichnis vertretenen italienischen Familiennamen kommen 1.232 (4,6 Prozent) auch in anderen Sprachen vor. Die ausschließlich italienischen Nachnamen repräsentieren 184.137 (80,8 Prozent) von insgesamt 228.008 Adressen mit italienischen Familiennamen, während aus den 1.232 Namen auch anderer Sprachen 43.871 (19,2 Prozent) Adressen resultieren. Unter diesen sind Namen, die auch spanische, katalanische, portugiesische oder französische sein können, besonders häufig vertreten. Hier bestätigt sich also die oben genannte häufigere Überschneidung aufgrund struktureller Verwandtschaft der Sprachen auch für die Adressenebene.

Tabelle 2: Überschneidung der italienischen Nachnamen mit anderen Sprachen (Namen und Adressen)

Sprache	Vorhandene Namen (Stand Juli 2000)		im Telefonbuch vertretene Namen		daraus resultierende Adressen	
	Zahl	Prozent	Zahl	Prozent	Zahl	Prozent
Insgesamt	43.855	100,0	26.950	100,0	228.008	100,0
ausschließlich italienisch	42.505	96,9	25.718	95,4	184.137	80,8
auch Nachname in anderer Sprache	1.350	3,1	1.232	4,6	43.868	19,2
ausgewählte Fälle*:						
italienisch + kastilisch (spanisch)	125	0,29	119	0,44	5.368	2,35
italienisch + kastilisch + portugiesisch	40	0,09	40	0,15	4.039	1,77
italienisch + türkisch	114	0,26	112	0,42	3.032	1,33
italienisch + französisch	113	0,26	99	0,37	2.921	1,28
italienisch + portugiesisch	78	0,18	73	0,27	2.825	1,24
italienisch + englisch	39	0,09	36	0,13	1.507	0,66
italienisch + provenzalisch	32	0,07	26	0,10	1.497	0,66
italienisch + deutsch	185	0,42	149	0,55	1.251	0,55
italienisch + katalanisch	49	0,11	44	0,16	910	0,40
italienisch + albanisch	15	0,03	15	0,06	699	0,31
italienisch + kastilisch + katalanisch	15	0,03	15	0,06	346	0,15
italienisch + katalanisch + portugiesisch	3	0,01	3	0,01	295	0,13
italienisch + katalanisch + provenzalisch	6	0,01	6	0,02	279	0,12
italienisch + baskisch	5	0,01	5	0,02	41	0,02

* Nach Häufigkeit und sprachlicher Verwandtschaft. Insgesamt ergaben sich rund 100 verschiedene Kombinationen, die aber zum größten Teil nur wenige Adressen codieren.

Nimmt man alle verfügbaren Informationen über Vor- und Nachnamen zusammen, lassen sich Kategorien bilden, die sich nach der Wahrscheinlichkeit der Übereinstimmung mit der tatsächlichen Staatsangehörigkeit (beziehungsweise nationalen Herkunft bei Eingebürgerten) ordnen lassen. Solange empirische Überprüfungen fehlen, lassen sich den Kombinationen von Sprachen zugewiesenen Vor- und Nachnamen ‚Treffer-Erwartungen‘ lediglich aufgrund formaler Kriterien zuordnen.

Daher wird nur zwischen hoher, mittlerer und geringer Treffererwartung unterschieden. Formal heißt hier, daß von einer hohen Treffererwartung dann ausgegangen wird, wenn einer Kombination nur eine Staatsangehörigkeit beziehungsweise Herkunft plausibel zugeordnet werden kann. Die Treffererwartung wird als mittel betrachtet, wenn mehr als eine Zuordnung plausibel möglich ist. Gering ist die Treffererwartung, wenn keine plausible Zuordnung möglich ist oder eine plausible Zuordnung nicht die gesuchte ist. Im letzteren Fall entspricht die geringe Erwartung für einen Italiener dann einer hohen Treffererwartung für eine andere Herkunft. Die Plausibilität der Zuordnung entspringt der Zuordnung der Namen zu einer (oder mehrerer) Sprache, die durch die Onomastik abgesichert ist und der Berücksichtigung demographischer Fakten der Migrationsgeschichte. Der grundsätzlich mögliche Fall, daß eine Person einen für seine nationale Herkunft völlig untypischen Namen trägt, wird als nicht plausibel betrachtet und bleibt deshalb unberücksichtigt. Es handelt sich bei dem Gesamtverfahren um ein dynamisches System, daß beim praktischen Einsatz durch Rückmeldungen, unabhängig davon stets aber durch Vervollständigung der Namensdatenbank und der Informationen zur Migrationsgeschichte verbessert wird.

In unserem Beispiel ist im einfachsten Fall ein ausschließlich italienischer Nachname mit einem ausschließlich italienischen Vornamen kombiniert. Hier läßt sich die höchste Wahrscheinlichkeit für einen Haushalt mit mindestens einer Person italienischer Herkunft annehmen, und damit eine entsprechend hohe Treffererwartung. Kommt der italienische Nachname auch in einer anderen Sprache vor, ist aber der Vorname ein ausschließlich italienischer, so ist ebenfalls von einer hohen Treffererwartung auszugehen. Eine Reihe von Vornamen ist gleichermaßen in Italien wie auf der iberischen Halbinsel verbreitet. Dies sind beispielsweise Alberto, Angelo, Antonio, Roberto und Maria. Ist ein solcher Vorname mit einem (ausschließlich) italienischen Nachnamen kombiniert, handelt es sich wiederum mit hoher Wahrscheinlichkeit um einen Italiener. Kann der Nachname aber auch ein spanischer oder portugiesischer sein, sinkt sie auf ein mittleres Maß. Exakte Angaben wären hier erst nach einem entsprechenden Praxistest möglich. Ist ein ausschließlich italienischer Nachname mit einem ausschließlich spanischen oder ausschließlich portugiesischen Vornamen kombiniert, ist eher von einem Lateinamerikaner auszugehen, denn ein gutes Drittel der argentinischen und gut 10 Prozent der brasilianischen Bevölkerung stammen aus Italien (vgl. Zilio 1963). Entsprechend ist bei einem englischen Vornamen von ausgewanderten Italienern (etwa in die USA) auszugehen. Deutsche Vornamen verweisen entweder auf Angehörige der zweiten Generation, auf binationale Ehen (bei weiblichen deutschen Vornamen) oder auf Nachkommen der

bereits vor längerer Zeit eingewanderten Italiener (die als Deutsche zu betrachten sind).

In der Tabelle 3 sind die häufigsten Kombinationen italienischer Nachnamen mit anderen Sprachen und Vornamen mit der beschriebenen Treffererwartung aufgeführt. Die Sprachzuweisung der Vor- und Nachnamen bezieht sich jeweils auf die genannten Sprachen ausschließlich. Bei leeren Feldern war die entsprechende Kombination unter den Adressen nicht vorhanden. Unberücksichtigt bleiben hier Fälle ohne Vornameneintrag oder mit abgekürztem Vornamen, sowie Fälle mit Nachnamen, die mit hier nicht aufgeführten Sprachen kombiniert sind. Auch Fälle mit bislang noch nicht zugewiesenem Vornamen wurden in die folgende Übersicht wegen Platzmangels nicht einbezogen.

In der ersten Zeile „entscheidet“ der ausschließlich italienische Vorname alle Nachnamen, die auch in einer anderen Sprache vorkommen, zugunsten des Italienischen. Daher ist für die ganze Zeile die Treffererwartung hoch. In der ersten Spalte (Nachname ausschließlich italienisch) sinkt die Treffererwartung, wenn für die betreffende Kombination eine plausible Alternative besteht. So kann es sich bei der Kombination eines italienischen Nachnamen mit ausschließlich portugiesischem, spanischem, französischem oder englischem Vornamen um Nachfahren von Emigranten in diese Staaten handeln (Brasilien, Argentinien, Frankreich, USA, Kanada). Die Treffererwartung (für einen Italiener) ist hier also eher gering, da diese Vornamen in Italien selten vergeben werden. Bei deutschen Vornamen kann es sich außer von Nachfahren von vor längerer Zeit eingewanderter Italiener auch um Angehörige der 2. oder 3. Generation oder um eine binationale Ehe handeln. Daher sinkt die Treffererwartung auf ein mittleres Maß. Die Einbeziehung dieser Gruppe in die Stichprobenziehung verhindert den Ausschluß einer wichtigen sozialen Gruppe und damit systematische Verzerrungen.

Bei einer ‚geringen‘ Treffererwartung für die italienische Sprache ergibt sich bei bestimmten Kombinationen die Möglichkeit zur Zuordnung eines anderen Sprachencodes (zum Beispiel des spanischen) mit einer ‚hohen‘ Treffererwartung. Bei fehlendem oder abgekürztem Vornamen (7,6 Prozent der Adressen mit italienischem Nachnamen) entscheidet nur der Nachname. Von diesen zusammen 17.500 Fällen waren immerhin 15.000 mit einem *ausschließlich* italienischen Nachnamen kombiniert, woraus eine hohe Treffererwartung für einen Haushalt mit mindestens einer Person italienischer Herkunft resultiert.

Tabelle 3: Kategorisierung von Treffererwartungen für einen Haushalt mit mindestens einer Person italienischer Herkunft (Kombinationen italienischer Namen mit ausgewählten Sprachen / n = 152.215)

Vorname	Nachname							
	italien.	italien. spanisch	italien. portug.	italien. französ.	italien. englisch	italien. deutsch	italien. türkisch	italien. span. portug.
italienisch	hoch	hoch	hoch	hoch	hoch	hoch	hoch	hoch
spanisch	gering	gering	gering	gering		gering	gering	mittel
portug.	gering	gering	gering	gering	gering			mittel
franz.	gering	gering	gering	gering	gering	gering	gering	gering
englisch	gering	gering	gering	gering	gering	gering	gering	gering
deutsch (m)	mittel	mittel	mittel	mittel	mittel	gering	mittel	mittel
deutsch (w)	mittel	mittel	mittel	mittel	mittel	gering	mittel	mittel
türkisch	gering	gering			gering	gering	gering	gering
italienisch spanisch	hoch	mittel	hoch	hoch	hoch	hoch	hoch	mittel
italienisch portug.	hoch	hoch	mittel	hoch	hoch	hoch	hoch	mittel
italienisch französisch	hoch	hoch	hoch	mittel	hoch	hoch	hoch	mittel
italienisch englisch	hoch	hoch			mittel	hoch		hoch
italienisch deutsch (m)	hoch	hoch	hoch	hoch	hoch		hoch	hoch
italienisch deutsch (w)	hoch	hoch	hoch	hoch	hoch	gering	hoch	hoch
englisch franz.	gering	gering	gering	gering	gering	gering	gering	
englisch deutsch (m)	mittel	mittel	mittel	mittel	gering	gering	gering	mittel
englisch deutsch (w)	mittel	mittel	mittel	mittel	gering	gering	gering	mittel
franz. deutsch (m)	mittel	mittel	mittel	gering	gering	gering	gering	mittel
franz. deutsch (w)	mittel	mittel	mittel	gering	gering	gering	gering	mittel
italienisch spanisch portug.	hoch	mittel	mittel	hoch	hoch	hoch	hoch	mittel
englisch franz. deutsch (m)	mittel	mittel	mittel	mittel	mittel	gering	mittel	mittel
englisch franz. deutsch (w)	mittel	mittel		mittel			mittel	mittel

(w) = weiblich, (m) = männlich; spanisch= kastilisch, katalanisch oder baskisch

Tabelle 4: Kategorisierung von Treffererwartungen¹⁾, absolute Häufigkeiten

Vorname	Nachname							
	italien.	italien. span.	italien. portug.	italien. franz.	italien. englisch	italien. deutsch	italien. türkisch	italien. span. port.
italienisch	72.689	1.164	800	147	301	516	412	793
spanisch	172	162	11	1		4	1	71
portugiesisch	170	40	70	1	1			78
französisch	803	40	7	86	16	19	12	19
englisch	247	19	8	7	30	1	7	9
deutsch (m)	17.354	338	127	982	290	40	730	194
deutsch (w)	11.620	319	125	327	121	26	282	213
türkisch	235	10			5	9	227	1
italienisch spanisch	2.853	75	30	8	5	30	10	38
italienisch portugiesisch	5.047	119	72	16	29	49	31	77
italienisch französisch	1.929	18	30	11	8	2	10	20
italienisch englisch	64	6			3	1		2
italienisch deutsch (m)	381	9	3	24	6		16	7
italienisch deutsch (w)	5.171	147	62	98	36	17	85	128
englisch französisch	84	4	3	3	4	1	2	
englisch deutsch (m)	2.510	72	36	107	39	5	67	49
englisch deutsch (w)	252	12	5	3	2	1	5	7
französisch deutsch (m)	434	8	6	32	6	2	12	9
französisch deutsch (w)	2.762	82	28	68	35	3	55	51
italienisch spanisch portugiesisch	13.300	372	231	39	55	113	57	286
englisch französisch deutsch (m)	1.337	32	11	70	25	2	29	19
englisch französisch deutsch (w)	69	2		3			4	1
Zusammen	139.483	3.050	1.665	2.033	1.017	841	2.054	2.072

1) Für einen Haushalt mit mindestens einer Person italienischer Herkunft (Kombinationen italienischer Namen mit ausgewählten Sprachen / n = 152.215).

(w) = weiblich, (m) = männlich; Nachnamen, die auch kastilische, katalanische oder baskische sein können, sind hier unter „spanisch“ zusammengefasst; leere Zellen sind nicht besetzt (es wurden keine Adressen gefunden).

In Tabelle 3 sind aus Platzgründen lediglich die häufigen Kombinationen aufgeführt. Die meisten Überschneidungen bei den Nachnamen betreffen nur wenige einzelne Namen und oft auch nur wenige Adressen. Das gleiche gilt für die Vornamen. Viele Vornamen aus anderen Sprachen finden sich bei italienischen Nachnamen nur ausnahmsweise. Auch diese sind hier deshalb nicht aufgeführt. In der gesamten Tabelle aller Kombinationen mit italienischen Nachnamen, die etwa 100 Spalten (Nachnamen) und fast ebenso viele Zeilen (Vornamen) enthält, ist dementsprechend die Mehrzahl der Zellen nicht besetzt. Ein großer Teil der verbleibenden Zellen repräsentiert nur wenige oder einzelne Adressen. Tabelle 4 zeigt die absoluten Häufigkeiten. Dabei wird deutlich, daß sich die Mehrzahl der Adressen in der Spalte „ausschließlich italienischer Nachname“ und in der Zeile „ausschließlich italienischer Vorname“ befinden. Größere Häufigkeiten finden sich noch in den übrigen Zeilen mit „auch“ italienischen sowie mit deutschen Vornamen.

Tabelle 5 enthält die Verteilung der Adressen mit italienischem Nachnamen nach den Treffererwartungsgruppen.

Tabelle 5: Treffererwartung für eine Person italienischer Herkunft nach Kategorien der Treffererwartung

Treffererwartung	Häufigkeit	
	absolut	prozentual
hoch	106.856	70,2 %
mittel	42.408	27,9 %
gering	2.951	1,9 %
insgesamt	152.215	100,0 %

Aufgrund der Verteilung wird deutlich, daß eine große Mehrheit der Adressen (rund 70 Prozent) eine höhere Treffererwartung aufweisen. Die Kategorie mit geringer Treffererwartung dürften in der Praxis noch weniger Adressen auf sich vereinen, da für einen Teil dieser Adressen die Zuordnung einer anderen Sprache mit mittlerer oder höherer Treffererwartung möglich ist.

3. Schlußfolgerungen für die Bildung einer Auswahlgrundlage

Aus den bisherigen Erläuterungen sind für die Konstruktion und Qualität einer Auswahlgrundlage folgende Rückschlüsse zu ziehen:

- a) Da die Zuweisung einer Sprache zu einem Namen nicht in allen Fällen auch mit der entsprechenden Staatsangehörigkeit beziehungsweise Herkunft einer Person

gleichbedeutend ist, wäre eine laufende Pflege der Auswahlgrundlage durch Rückmeldung aus den Erhebungen wünschenswert. Die laufende Fehlerbereinigung kann sich auf Namen oder auch Zuweisungsregeln beziehen.

b) Da ein Nachname in verschiedenen Sprachen (Kaplan, Türk, Wang) vorkommen kann, ist eine *parallele* Anwendung möglichst umfangreicher Namenlisten verschiedener Sprachen unerlässlich. Für die Zuordnungssystematik ist also die Information einer mehrfachen oder ausschließlichen Zuordnung des Vor- und Nachnamen zu Sprachen zu ermitteln, um Treffererwartungen angeben zu können.

c) Die Effektivität des vorgestellten Verfahrens wird vom Vorhandensein eines Vornameneintrages im Telefonbuch beeinflusst. Die notwendige Einbeziehung des Vornamen wird durch die Eintragungsgewohnheiten im Telefonbuch Deutschlands gewährleistet. Nur 8,5 Prozent der 23.283.839 im Nachnamen zugeordneten (deutschen und nichtdeutschen) Privathaushalte im Pilotprojekt 1999 (Stand: Mai 1999) hatten keinen oder nur einen abgekürzten Eintrag im Vornamenfeld.

d) Die Zuordnung wird vor dem speziellen Hintergrund der historisch-politischen Bedingungen der Zuwanderung einer Nationalitätengruppe genauer. Außerdem sind Vorgaben durch die Beschränkung oder Verfügbarkeit der Onomastikliteratur zu berücksichtigen. Beide Sachverhalte machen die Bildung einer Auswahlgrundlage, in der alle Nationalitäten entsprechend ihrer tatsächlichen Anteile abgebildet werden, unmöglich. Es lassen sich nur separate Listen für die einzelnen Sprachen bilden, die dann jedoch mit steigendem Umfang eine repräsentative Grundlage für die im Telefonverzeichnis *eingetragenen* Haushalte dieser nationalen Herkunft darstellen. Denn aufgrund fehlender Zusammenhänge zwischen Sozial- oder Einstellungsvariablen und dem jeweiligen Namen sind disproportionale Ausfälle eher *nicht* zu erwarten.

e) Die Eintragungen im Telefonbuch sind durch systematische (Transkription) oder zufällige Schreibfehler beziehungsweise -varianten nicht vollständig durch Namenlisten zu erfassen. Die bloße Eingabe von Namenlisten ist also nicht ausreichend.

f) Die Anwendung des beschriebenen Verfahrens auf das Telefonbuch Deutschlands ist nur lohnend, wenn das Telefonverzeichnis von Rufnummern bereinigt wird, die nicht zur Grundgesamtheit der Privathaushalte mit mindestens einer Person nicht-deutscher Herkunft gehören. Dazu sind vorab die Anschlüsse mit ausschließlicher Faxnutzung, die Nicht-Privathaushalte, die Mobilfunknummern und die deutschen Privathaushalte in der Kategorie mit der höheren Treffererwartung zu bereinigen. Im Rahmen einer Zufallsnummernauswahl können sie wiederum zur Bereinigung einge-

setzt werden. Während die Fax- und Mobilfunkanschlüsse einfach zu identifizieren sind, ist man bei der Bereinigung der Nicht-Privathaushalte wiederum auf den Einsatz von möglichst fehlerfreien *und* umfassenden Listen mit entsprechenden Bezeichnungen (Firmen-, Institutions- und Serviceeintragungen) angewiesen. Im Rahmen des Pilotprojektes 1999 kam eine noch deutlich zu erweiternde Liste mit 2,6 Millionen verschiedenen Bezeichnungen von Nicht-Privathaushalten zum Einsatz, die zum größten Teil für das Nachnamen- und Zweiteintragsfeld Treffer erzielte. Weiterhin muß die Auswahlgrundlage von Elementen bereinigt werden, die zwei- oder mehrmals vertreten sind und somit eine größere Auswahlchance besitzen würden. Dazu sind die doppelten Rufnummern und die ISDN-Anschlüsse ¹² zu zählen. ¹³

g) Die detaillierte Zuordnung ermöglicht sowohl die Kategorisierung nach der Treffererwartung wie auch das Erkennen von Subgruppen mit besonderen Merkmalen (2./3.Generation, binationale Ehen). Für die Bildung einer Auswahlgrundlage ist deshalb nicht allein die hohe Treffererwartung entscheidendes Kriterium, sondern auch die Integration der genannten Subgruppen. Auf diese Weise können systematische Verzerrungen der Auswahlgrundlage reduziert werden.

h) Als Auswahlgrundlage für die telefonische Befragung von Zuwanderern sind die Namenlisten mit folgenden Einschränkungen behaftet:

- weil es sich um Haushaltsstichproben handelt, sind eine nachgelagerte Zufallsauswahl im Haushalt (zum Beispiel ‚Last-Birthday‘ oder Schwedenschlüssel) und mögliche Gewichtungseinsparungen (Transformation) durch unterschiedliche Haushaltsgrößen (von der Heyde 1999: 46) notwendig,
- eine Sprachzuweisung kann auf mehrere Staatsangehörigkeiten hindeuten (zum Beispiel Spanisch: Spanien, spanischsprachiges Amerika, Philippinen),
- auch eingebürgerte Zuwanderer sind in die Auswahlgrundlage einbezogen,
- in der Liste fehlen die nicht im Telefonverzeichnis eingetragenen Haushalte,
- Haushalte ohne Festnetzanschluß (kein Telefonanschluß oder nur Handy) sind ebenfalls nicht in der Auswahlgrundlage enthalten.

Die genannten Bedingungen führen in unterschiedlichem Umfang zu einer Einschränkung der Eignung dieser Auswahlliste als Grundlage für repräsentative Zuwandererstichproben. Der Ausschluß der nicht-eingetragenen Haushalte erscheint besonders problematisch für die Qualität der Stichproben. Denn hinsichtlich der

¹² Es werden Einträge mit identischer Orts-, PLZ-, Straßen-, Hausnummern- und Nachnamensangabe gekennzeichnet.

¹³ Weiterhin erfolgt eine Zuordnung der amtlichen Gemeindegemeindeflüsselnummern..

Telefondichte sind für Zuwanderer nur geringere Ausfallraten zu befürchten (Vgl. Granato 1999: 49). Es ist nicht auszuschließen, daß sich die Haushalte ohne Eintrag im Telefonverzeichnis von der Gesamtheit systematisch unterscheiden. Mögliche Ansätze zur Reduzierung dieser Fehler (Gewichtung über laufende face-to-face-Befragungen von Zuwanderern, Erweiterung der Stichproben um Zufallsnummern) sind vor dem Hintergrund einer möglicherweise voranschreitenden Nicht-Eintragungstendenz noch zu entwickeln.

4. Erste Umsetzung des Verfahrens (Pilotprojekt 1999)

Die Anwendung des geschilderten Verfahrens auf den Datenbestand von Frühjahr 1999 ermöglicht erste Erfahrungen. An dieser Stelle wären erste empirische Belege für die Stichprobenqualität und Effizienz des Verfahrens wünschenswert. Solche Daten liegen bislang jedoch noch nicht vor. Deshalb folgen hier zunächst Erläuterungen, die lediglich einen ersten Einblick beziehungsweise eine Veranschaulichung bieten sollen.

Es wurden zunächst alle 36 Millionen Eintragungen des öffentlichen Telefonbuches (wie in Punkt f von Teil 3 bereits erwähnt) um die Datensätze reduziert, die *nicht* auf Telefonanschlüsse (Festnetz) von Privathaushalt hindeuten. Die verbliebenen Datensätze (circa 28 Millionen Privathaushalte) wurden mit Namenlisten aus 58 Sprachen (einschließlich Deutsch) abgeglichen. Die Autoren erstellten eine Datenbank (Vor- und Nachnamen) mit entsprechenden Verweisen auf die Onomastikliteratur, aus denen sie entnommen wurden. Für jeden Datensatz entstand beim Vor- und Nachnamen eine einfache oder mehrfache Sprachzuordnung. Die Kombination des Vornamen und des Nachnamen mit ihren einfachen bzw. mehrfachen Sprachzuordnungen führen zu den Treffererwartungen (gering/mittel/hoch) für eine nationale Herkunft, die aufgrund der formalen Aspekte zugeordnet wurden. Nur 23 der rund 28 Millionen Datensätze wiesen nach diesem Zuweisungslauf einen Sprachencode mit Treffererwartung auf. Die nicht zugewiesenen 5 Millionen Privathaushalte sind auf unvollständige Namenlisten oder nicht zu identifizierende Eintragungen zurückzuführen. Eine laufende Eingabe von Namen aus der Onomastikliteratur ist deshalb unerlässlich. Hier können nur für ausgewählte Sprachen erste Informationen zum Umfang der Listen und der regionalen Verteilung vorgestellt werden. Dazu wurden die Sprachen unterschiedlich umfangreicher Bevölkerungsgruppen in Deutschland (Deutsch, Türkisch, Italienisch und Japanisch) ausgewählt.

Über den Vergleich zu amtlichen Daten (wegen fehlender Haushaltsgrößen je Nationalität und Bundesland konnten nur Personenzahlen herangezogen werden) werden erste grobe Hinweise und Indizien für Probleme erkennbar. Zunächst ist für alle Sprachen festzustellen, daß im Rahmen dieses ersten Testlaufs nur knapp die Hälfte der geschätzten Grundgesamtheit¹⁴ erreicht wurde. Diese geringe Erfassungsquote kann für den Zeitpunkt des Pilotprojektes 1999 auf folgende Gründe zurückzuführen sein:

- die Zahl der Personen im Ausländerzentralregister (AZR) ist zu hoch (vgl. Lederer/Rau/Rühl 1999: 44),
- Haushalte ohne Festnetzanschluß können nicht erreicht werden,
- Haushalte ohne Eintrag im Telefonverzeichnis können ebenfalls nicht erreicht werden,
- die Unvollständigkeit oder Fehlerhaftigkeit der Namenlisten,
- die ausschließliche Betrachtung der Adressen mit hoher Treffererwartung.

Im Rahmen einer Optimierung des Verfahrens für die Anwendung auf den Datenbestand aus dem Jahr 2000 sind neben der Erweiterung der Namenlisten auch die Zuordnungsgruppen mit niedrigeren Treffererwartungen einzubeziehen. Weiterhin müssen die steigenden Anteile der Nichteingetragenen und der Haushalte ohne Festnetzanschluß¹⁵ in die Erhebungsergebnisse einfließen. Zunächst sollte überprüft werden, ob in laufenden Erhebungen (face-to-face)¹⁶ systematische Unterschiede zwischen eingetragenen und den beiden ausgeschlossenen Gruppen (Haushalte ohne Eintrag oder Festnetzanschluß) bestehen. Bei Unterschieden könnten Gewichtungmaßnahmen der Befragungsergebnisse vorgenommen werden. Die Anwahl von Zufallsnummern ist theoretisch eine zweite Möglichkeit, die ‚Nonpubs‘ (Häder 1996) zu erreichen. Für die spezielle Befragtengruppe der Zuwanderer steht dieser Möglichkeit jedoch der praktische Aspekt einer hohen Ausfallrate entgegen. Dennoch bliebe zu bedenken, daß unabhängig vom genauen Design der Zufallsnummernanwahl nur das geschätzte Fünftel (Follmer 1998: 43) der ‚Nonpubs‘ einer

14 Da es sich um Haushaltslisten einschließlich eingebürgerte Personen handelt, kann eine Schätzung der Grundgesamtheit nur auf der Basis der verfügbaren Zahlen aus dem AZR, der Einbürgerungsstatistik und der Mikrozensus-Angaben zur durchschnittlichen Haushaltsgröße durchgeführt werden.

15 Hier wären nur die Haushalte mit ausschließlich einem bzw. mehreren Mobilfunkanschlüssen von Bedeutung.

16 Hier wäre eine CATI-Erhebung mit Zufallsnummernverfahren erforderlich, um systematische Verzerrungen der face-to-face-Erhebungen (zum Beispiel Interviewereffekte bei aufwendigen Quotenmodellen oder fremdsprachigen Privathaushalten) auszuschließen.

Zuwanderer Stichprobe zu erreichen wäre. Allerdings ist der Umfang der ‚Nonpubs‘ gerade auch für die Sondergruppe Zuwanderer eher unbekannt. Deshalb würde ein Design für die Zufallsnummernwahl, welches vorab für eine Stichprobe eine genaue Quote von ‚Nonpubs‘ vorgibt, auf ungesicherten Annahmen beruhen. Um hier hinsichtlich der Abschätzung dieses Anteils neue Erkenntnisse zu sammeln, müsste also aus dem ‚Universum‘ der Zufallszahlen eine Stichprobe gezogen werden.¹⁷

Der Vergleich der durch die vorgestellte Methode ermittelten Haushalte mit den amtlichen Daten (siehe Anhang, Tabelle 6 bis 9) ergibt für drei Sprachen (Türkisch, Italienisch, Deutsch) nur geringfügige Abweichungen bei der Verteilung nach Bundesländern. Die Liste japanischer Haushalte zeigt hinsichtlich der Verteilung auf die Bundesländer hingegen größere Unterschiede zur Grundgesamtheit (Personen im AZR). Allerdings sind spezifische Bedingungen der Zuwanderung oder sprachliche Überschneidungen, die gerade beim Japanischen nur in geringem Umfang bestehen, eher nicht als Begründung angegeben werden. Vielmehr ist auffallend, daß mit steigendem Umfang der jeweiligen Auswahlliste sich auch die Abweichungen verringern.

Für die Verteilung der über 20 Millionen Adressen der deutschen Privathaushalte mit hoher Treffererwartung sind besonders die vergleichsweise hohen Abweichungen in den fünf neuen Bundesländern (außer Berlin) hervorzuheben. Die amtlichen

¹⁷ Für die Generierung einer solchen Grundgesamtheit aller möglichen Rufnummern hat sich das „Gabler-Häder-Design“ (Gabler/Häder 1998: 69ff.) bewährt. Aus dem Telefonbuch (2/2000) werden aus den eingetragenen Rufnummern (35.454.869) zunächst alle Mobilfunk-, Service- und Sammelanschlüsse gelöscht (2.883.572; es wurden auch Datensätze mit leeren Rufnummerfeldern gelöscht). Von den verbleibenden Nummern (32.571.297) wurden die letzten zwei Ziffern gestrichen. Daraus entstehen Nummernblöcke (716.043), die durch die möglichen zweistelligen Endziffern (00,01,...99) zu einer Gesamtheit möglicher Rufnummern (71.604.300) führen. Hinsichtlich der ‚Bereinigung‘ dieser Zufallsnummernmenge können die über Onomastik identifizierten eingetragenen deutschen Haushalte (circa 25 Millionen), die eingetragenen Faxanschlüsse (circa 2 Millionen) und die eingetragenen Nicht-Privathaushalte (circa 5 Millionen) gelöscht werden. Die ungefähr 39 Millionen verbleibenden Zufallsnummern werden also aus eingetragenen und nicht-ingetragenen Zuwandererhaushalten, nicht-ingetragenen deutschen Privathaushalten, nichteingetragenen Faxanschlüssen oder Nicht-Privathaushalten, sowie nicht belegten Zufallsnummern gebildet. Eine Erhebung auf der Basis dieser bereinigten Rufnummernmenge steht jedoch trotz Einsatz automatischer Anwähler zur Eliminierung nicht-existierender Nummern vor einem nicht zu unterschätzenden Sprachproblem. Die Verteilung der Nummern auf die Interviewer mit der geforderten Heimatsprache stellt ein organisatorisches Problem dar (schnelle Weitergabe an freie Interviewer mit entsprechender Sprachkenntnisse oder Sprachcodierung für späteren Anruf).

Daten weisen für alle fünf Bundesländer einen Bevölkerungsanteil von zusammen 18,5 Prozent auf. Im Telefonbuch konnte jedoch nur ein Anteil von 16 Prozent identifiziert werden. Damit wird ein größerer Teil der Summe der Abweichungsbeträge durch die Untererfassung dieser fünf Bundesländer hervorgerufen. Die höhere Nichteintragungsquote (vgl. Folmer 1998: 50f.) in den neuen Bundesländern könnte eine mögliche Ursache sein.

5. Zusammenfassung

Im Rahmen eines Pilotprojektes aus dem Jahr 1999 wurde ein Konzept entwickelt, daß die Zuordnung von Sprachen zu Namen ermöglicht. Die Grundlage wird durch Erkenntnisse der Namenforschung (Onomastik) gebildet. Die möglichst fehlerreduzierende Durchführung wird durch Zuweisungsregeln abgestützt, welche die maximale Information der Telefonbucheintragen (Vorname, Nachname und Zweitfeldeintrag) nutzen. Die Zuweisungsregeln werden durch die historisch-politische Betrachtung der Zuwanderung jeder einzelnen Sprache verbessert. Das gesamte Konzept wird durch interne Kriterien (Kombination von Vor- und Nachnamen) und Rückmeldungen aus empirischen Erhebungen laufend verbessert und gepflegt. Dennoch sollte hier auch darauf hingewiesen werden, daß es sich vorrangig um ein Konzept für Listenauswahlen handelt. Mit der festzustellenden Unvollständigkeit dieser Liste rückt die Problematik der nicht im Telefonbuch eingetragenen Zuwandererhaushalte in den Blickpunkt der weiteren Entwicklung. Es wäre zu testen, ob Wege der Einbeziehung dieser Telefonhaushalte gefunden werden können. Ob dies durch Gewichtungmaßnahmen oder Varianten der Zufallsnummernwahl geschehen soll, kann hier noch nicht gesagt werden.

Korrespondenzadresse

*Dr. Andreas Humpert, Klaus Schneiderheinze
Humpert & Schneiderheinze GbR
Kortumstr. 117
47057 Duisburg
E-Mail: hs@hs-gbr.de*

Tabelle 6: Verteilung der türkischen Privathaushalte (hohe Treffererwartung - Pilotprojekt 99) und Personenzahl (AZR: 31.12.98) nach Bundesländern

Bundesland	türkische Bevölkerung*	türkische Haushalte im Telverzeichnis	Differenz
01 Schleswig-Holstein	4	7.453 2,2%	+ 0,1%
02 Hamburg	80	11.593 3,4%	- 0,4%
03 Niedersachsen	138	23.985 6,9%	+ 0,3%
04 Bremen	31	5.611 1,6%	± 0,0%
05 Nordrhein-Westfalen	714	110.842 32,1%	- 1,8%
06 Hessen	219	34.650 10,0%	- 0,4%
07 Rheinland-Pfalz	8	13.365 3,9%	+ 0,1%
08 Baden-Württemberg	357	57.809 16,7%	- 0,2%
09 Bayern	268	47.545 13,8%	+ 1,1%
10 Saarland	10	2.397 0,7%	- 0,1%
11 Berlin	143	25.671 7,4%	+ 0,6%
12 Brandenburg	0,1%	1.079 0,3%	+ 0,2%
13 Mecklenburg-Vorpommern	0,1%	513 0,1%	± 0,0%
14 Sachsen	0,2%	1.322 0,4%	+ 0,2%
15 Sachsen-Anhalt	0,1%	849 0,2%	+ 0,1%
16 Thüringen	0,1%	651 0,2%	+ 0,1%
gesamt	2.100	345.335 100%	+ 5,7%

Tabelle 7: Verteilung der italienischen Privathaushalte (hohe Treffererwartung - Pilotprojekt 99) und Personenzahl (AZR: 31.12.98) nach Bundesländern

Bundesland	italienische Bevölkerung*	italienische Haushalte im Telverzeichnis	Differenz
01 Schleswig-Holstein	4	1.657 1,1%	+ 0,4%

02 Hamburg	7	2.151	1,4%	+ 0,1%
	1,3%			
03 Niedersachsen	26	7.885	5,3%	+ 1,0%
	4,3%			
04 Bremen	1	638	0,4%	+ 0,1%
	0,3%			
05 Nordrhein-Westfalen	275	36.096	24,3%	- 0,2%
	24,5%			
06 Hessen	78	17.310	11,7%	- 1,2%
	12,9%			
07 Rheinland-Pfalz	30	7.520	5,1%	+ 0,1%
	5,0%			
08 Baden-Württemberg	186	43.318	29,2%	- 1,3%
	30,5%			
09 Bayern	89	21.111	14,2%	- 0,5%
	14,7%			
10 Saarland	19	4.888	3,3%	+ 0,2%
	3,1%			
11 Berlin	11	3.570	2,4%	+ 0,6%
	1,8%			
12 Brandenburg	1	518	0,3%	+ 0,1%
	0,2%			
13 Mecklenburg-Vorpommern	472	242	0,2%	+ 0,1%
	0,1%			
14 Sachsen	2	681	0,5%	+ 0,1%
	0,4%			
15 Sachsen-Anhalt	915	419	0,3%	+ 0,2%
	0,1%			
16 Thüringen	978	350	0,2%	± 0,0%
	0,2%			
gesamt	612	148.354	99,9%	+ 6,2%
	100%			

Tabelle 8: Verteilung der japanischen Privathaushalte (hohe Treffererwartung -Pilotprojekt 99) und Personenzahl (AZR: 31.12.98) nach Bundesländern

Bundesland	japanische Bevölkerung*	japanische Haushalte im Telverzeichnis	Differenz
01 Schleswig-Holstein	656 2,2%	125 2,3%	+ 0,1%
02 Hamburg	4.3 14,9%	298 5,5%	- 9,4%
03 Niedersachsen	925 3,2%	188 3,5%	+ 0,3%
04 Bremen	378 1,3%	32 0,6%	- 0,7%
05 Nordrhein-Westfalen	9.8 33,7%	2.107 39,0%	+ 5,3%
06 Hessen	4.4 15,3%	868 16,1%	+ 0,8%
07 Rheinland-Pfalz	450 1,5%	96 1,8%	+ 0,3%
08 Baden-Württemberg	2.4 8,5%	551 10,2%	+ 1,7%
09 Bayern	3.3 11,5%	682 12,6%	+ 1,1%
10 Saarland	146 0,5%	29 0,5%	± 0,0%
11 Berlin	1.8 6,2%	338 6,3%	+ 0,1%
12 Brandenburg	59 0,2%	15 0,3%	+ 0,1%
13 Mecklenburg-Vorpommern	22 0,1%	9 0,2%	+ 0,1%
14 Sachsen	198 0,7%	30 0,6%	- 0,1%
15 Sachsen-Anhalt	64 0,2%	17 0,3%	+ 0,1%
16 Thüringen	67 0,2%	16 0,3%	+ 0,1%
gesamt	29.29 100%	5.401 100%	+ 20,3%

* Quelle: Statistisches Bundesamt, Auszählung des Ausländerzentralregisters, Stichtag: 31.12.1998. Aufgrund der geringen Fallzahlen werden in einzelnen Bundesländern keine Mikrozensus-Zahlen zur Haushaltsgröße je Nationalität ausgegeben. Deshalb wurden ersatzweise die Personenzahlen des AZR für den Vergleich der relativen Anteile herangezogen. Eingebürgerte Personen blieben hier unberücksichtigt.

Tabelle 9: Verteilung der deutschen Privathaushalte (hohe Treffererwartung - Pilotprojekt 99) und Personenzahl (Bevölkerungsfortschreibung: 31.12.98) nach Bundesländern

Bundesland	deutsche Bevölkerung*	Deutsche Haushalte im Telverzeichnis	Differenz
01 Schleswig-Holstein	2.623.370 3,5%	766.994 3,7%	+ 0,2%
02 Hamburg	1.381.434 1,9%	442.057 2,1%	+ 0,2%
03 Niedersachsen	7.391.715 9,9%	2.061 10,0%	+ 0,1%
04 Bremen	586.099 0,8%	179.057 0,9%	+ 0,1%
05 Nordrhein-Westfalen	15.980.347 21,4%	4.321 21,0%	- 0,4%
06 Hessen	5.193.394 7,0%	1.581 7,7%	+ 0,7%
07 Rheinland-Pfalz	3.725.804 5,0%	1.051 5,1%	+ 0,1%
08 Baden-Württemberg	9.157.035 12,3%	2.611 12,7%	+ 0,4%

09 Bayern	10.992.548	14,7%	3.221	+ 0,9%
			15,6%	
10 Saarland	994.520	1,3%	290.223	+ 0,1%
11 Berlin	2.916.818	3,9%	797.310	± 0,0%
12 Brandenburg	2.536.045	3,4%	571.555	- 0,6%
13 Mecklenburg-Vorpommern	1.772.887	2,4%	383.460	- 0,5%
14 Sachsen	4.406.790	5,9%	1.111	- 0,5%
			5,4%	
15 Sachsen-Anhalt	2.627.103	3,5%	625.190	- 0,5%
16 Thüringen	2.431.509	3,3%	594.809	- 0,4%
gesamt	74.717.418	100%	20.627.301	100%
			5,7%	

* Quelle: Statistisches Bundesamt, Bevölkerungsfortschreibung, Stichtag: 31.12.1998.

Literatur

De Felice, E., 1982: I nomi degli italiani. Informazioni onomastiche e linguistiche socio-culturali e religiose. Rom: Sarin.

Follmer, R./Smid, M., 1998, Nichteingetragene Telefonnummern: Ergebnisse eines Methodentests. S. 43-57 in: S. Gabler/S. Häder/J.H.P. Hoffmeyer-Zlotnik (Hrsg.): Telefonstichproben in Deutschland. Opladen: Westdeutscher Verlag.

Gabler, S./Häder, S., 1998: Ein neues Stichprobendesign für telefonische Umfragen in Deutschland. S. 69-88 in: S. Gabler/S. Häder/J.H.P. Hoffmeyer-Zlotnik (Hrsg.): Telefonstichproben in Deutschland. Opladen: Westdeutscher Verlag.

Granato, N., 1999, Die Befragung von Arbeitsmigranten: Einwohnermeldeamts-Stichprobe und telefonische Erhebung?. ZUMA-Nachrichten 45: 44-60.

Häder, S., 1996: Wer sind die ‚Nonpubs‘? Zum Problem anonymer Anschlüsse bei Telefonumfragen. ZUMA-Nachrichten 39: 45-68.

Homge, R., 1988: Zur modernen spanischen Vornamengebung. Die Vornamen in Salamanca von 1900 bis 1986. Dissertation Universität Siegen.

Kaleta, Z., 1998/99: The Polish surname as a carrier of moral values and as an ethnic heritage. ONOMA 34: 101-118.

Kluge, I.-L., 1985: Die heutigen japanischen Familiennamen. S. 121-128 in: R. Schützeichel/A. Wendehorst (Hrsg.), Erlanger Familien-Kolloquium. Neustadt a. Aisch: Degener & Co.

Kunze, K., 1998: dtv-Atlas Namenkunde. München: Deutscher Taschenbuch Verlag.

Lederer, H.W./Rau, R./Rühl, S., 1999: Migrationsbericht 1999. Bamberg: Europäisches Forum für Migrationsstudien.

Marhenke, W., 1997: Telefonanschlußdaten als Auswahlgrundlage. S. 207-220 in: S. Gabler/J.H.P. Hoffmeyer-Zlotnik (Hrsg.), *Stichproben in der Umfragepraxis*. Opladen: Westdeutscher Verlag.

Onoma - Journal of the International Council of Onomastic Sciences, 1950ff.

Rösch, G., 1998: Maßnahmen zur Reduktion von Stichprobenfehlern bei telefonischen Bevölkerungsumfragen. S. 101-119 in: S. Gabler/S. Häder/J.H.P. Hoffmeyer-Zlotnik (Hrsg.), *Telefonstichproben in Deutschland*. Opladen: Westdeutscher Verlag.

Salentin, K., 1999: Die Stichprobenziehung bei Zuwandererbefragungen. *ZUMA-Nachrichten* 45: 115-135.

von der Heyde, C., 1999: Techniken und Möglichkeiten der Realisierung von Random-Stichproben. S. 35-60 in: ADM Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V./AG.MA Arbeitsgemeinschaft Mediaanalyse e.V. (Hrsg.): *Stichproben-Verfahren in der Umfrageforschung*. Opladen: Leske + Budrich.

Zamora, J., 1992: Hugenottische Familiennamen im Deutschen. *Beiträge zur Namensforschung Neue Folge*, Beiheft 36. Heidelberg: Carl Winter

Zilio, G., 1963: Appunti di onomastica Rioplatense. Comportamenti grafici e fonetici dei cognomi italiani nel Plata. In: VII Congresso internazionale di Scienze Onomastiche Firenze - 4-8 Aprile 1961. *Atti e Memorie*, Bd. 3: *Memorie della Sezione Antroponymica*. Florenz, S. 227-239.