

Art von seelichem Erlebnis zusammen), а при винительном падеже в подавляющем большинстве случаев при помощи генитивного управления (*Er verspürte eine Art hoffnungslosen Hungers nach diesen leichtbekleideten Braungebrannten*).

В качестве вторых компонентов данных словосочетаний, оформленных при помощи как генитивного, так и предложного управления представлены в подавляющем большинстве случаев существительные женского рода (*Es ist eine Art sommerlicher Erinnerung*), а в роли вторых компонентов, словосочетаний, организованных на основе согласования, чаще всего встречались существительные мужского и среднего рода (*Sie schienen fortgezogen zu werden von einer Art bräutlichem Schmerz*). Все словосочетания со вторыми компонентами – существительными во множественном числе – оформлены исключительно при помощи генитивного и предложного управления (*eine Art goldfarbener gemusterter Pluderhosen, eine Art von riesigen Rosen*).

Таким образом, можно утверждать, что словосочетания исследуемого типа оформляются примерно в равной степени при помощи генитивного и предложного управления, а также согласования. Следует особо подчеркнуть, что, несмотря на то, что нормативные грамматики современного немецкого языка рекомендуют выбирать согласование как вид подчинительной связи, генитивное и предложное управление не только не исчезает из словосочетаний данного типа, но и остается наиболее часто используемым видом подчинительной связи. По-видимому, реальный узус в способах оформления рассматриваемых словосочетаний еще не устоялся в достаточной мере, чтобы формулировать нормативные рекомендации столь однозначно. Вместе с этим подтверждается гипотеза о том, что синтаксическое оформление данных конструкций определяется грамматическими характеристиками их компонентов.

Литература

1. Габрите Е.К. Немаркированные падежные формы в современном немецком языке: Дисс. ... канд. филол. наук. М., 1971.
2. Кутасина Е.С. Раевский М.В. Еще раз о виде подчинительной связи словосочетаний типа ein Liter Wasser и падежной форме их второго компонента // Вестн. Моск. ун-та. Сер. 9. Филология. 1985. № 2.
3. Blatz Fr. Neuhochdeutsche Grammatik mit Berücksichtigung der historischen Entwicklung der deutschen Sprache. 3. völlig neu bearb. Aufl. Karlsruhe, 1896. Bd. II. Satzlehre.
4. Duden. Richtiges und gutes Deutsch: Wörterbuch der sprachlichen Zweifelsfälle. 5., neu bearb. u. erw. Aufl. Mannheim etc.: Dudenverlag, 2001. (Der Duden in 12 Bd.: Das Standardwerk zur deutschen Sprache; Bd. 9.)
5. Engel U. Syntax der deutschen Gegenwartssprache. Berlin (West), 1977.
6. Mensing O., Erdmann O. Grundzüge der deutschen Syntax nach ihrer geschichtlichen Entwicklung dargestellt. Stuttgart, 1898. 2. Abt.

A.V. Pirozhkov

STATISTISCHE UNTERSUCHUNGEN ZUM BEKANNTHEITSGRAD RUSSISCHER (U.A. SLAWISCHER) ENTLEHNUNGEN IN DER DEUTSCHEN SPRACHE DER GEGENWART UND IHRER ABHÄNGIGKEIT VOM ALTER UND VOM WOHNORT

Universität Potsdam, Deutschland

Die Anwendungsmöglichkeiten mathematischer und statistischer Methoden in der linguistischen Forschung sind umfangreich. In der Anfangsphase ging es zuerst um einfache deskriptive Methoden der Zusammenfassung und Darstellung von Daten und um statistische Verfahren zur Verarbeitung empirischer linguistischer Daten. Danach hat man mehr auf Introspektion zurückgegriffen. Heute ermöglicht Statistik allen empirischen Wissenschaften objektive Entscheidungen über die Brauchbarkeit der überprüften Hypothesen.

Der Teilaspekt der Statistik, der sich mit der Überprüfung von Hypothesen befasst und mit dem wir uns

im Weiteren beschäftigen möchten, wird als analytische Statistik oder Inferenz-(schließende) Statistik bezeichnet [vgl. Bortz 1993: 1-5]. In der Inferenzstatistik formuliert man eine statistische Hypothese: das ist eine präzisierete Annahme über die Häufigkeitsfunktion einer beobachtbaren Zufallsvariablen in der Grundgesamtheit, insbesondere über unbekannte Parameter dieser Häufigkeitsfunktion. Die zu prüfende statistische Hypothese nennt man Nullhypothese und bezeichnet sie mit H_0 . Es wird angenommen, dass die wirkliche Differenz Null ist. So kann man auch *unsere Prämisse*, dass *der Sprachgebrauch entlehnter Wörter des unter-*

suchten Korpus vom Alter und von der regionalen Gegebenheit abhängig ist, als Nullhypothese betrachten.

Weil noch keine Theorie, aus der die Arbeitshypothese abgeleitet werden kann, existiert, wird unsere Untersuchung als ein Erkundungsexperiment bezeichnet. Bevor die Hypothese zu einer praktisch brauchbaren, allgemein gültigen Theorie zusammengefasst werden kann, bedarf es weiterer Untersuchungen, in denen mit inferenzstatistischen Verfahren die Gültigkeit der „erkundeten“ Ziele überprüft wird.

Die Ziele der statistischen Untersuchungen in dieser Arbeit sind folgende:

1) die Überprüfung der Häufigkeit der Bekanntmachung russischer (u.a. slawischer) Entlehnungen in der deutschen Sprache der Gegenwart;

2) Die Abhängigkeit/Unabhängigkeit des Bekanntheitsgrades vom Alter und vom Wohnort der Probanden im Ost- und Westteil Deutschlands.

Gebäuchliche Methode der empirischen Sozialforschung ist die Informantenbefragung. Durch Fragen (verbale Stimuli) werden Antworten (verbale Reaktionen) hervorgerufen. In jeder Befragung manifestiert sich ein Interesse: der aktiv Fragende wünscht Informationen solcher Art, dass ein Bezug zwischen Frage und Antwort erkennbar ist. Das Ziel dieser Arbeit war es, durch die Befragung aufzuklären, ob die vorgeschlagenen Entlehnungen den Informanten bekannt oder unbekannt sind, genauer gesagt, ob die Versuchspersonen („Vpn“) oder Probanden diese Entlehnungen vom Blatt und ohne Kontext verstehen oder nicht. Formales Ziel der Befragung war es, eine möglichst hohe Übereinstimmung in der Kommunikation zu erreichen. Eine solche ist auch Indiz dafür, dass das Ungleichgewicht der Motivation geringer, die Verlässlichkeit der Wiedergabe einer Meinungsstruktur höher ist. Je geringer der Grad der Gemeinsamkeit in der Kommunikation, desto selektiver sind die Reaktionen. Das bedeutet, dass Antworten nur jenen Teil der Erfahrungswelt des Befragten erreichen können, der durch die Struktur der Befragung gegeben ist.

Die Informanten erhielten Fragebögen. Aus den Wörterbüchern von Klien [1954], Duden [1989, 1990, 1994], Wahrig [1986], Klappenbach [1967–1978], Pfeifer [1989], Kluge [1975] wurde der untersuchte Korpus russischer (u.a. slawischer) Entlehnungen zusammengestellt und später durch andere Quellen, wie z.B. Bellmann [1971], Opel'baum [1971] ergänzt. Jedem Probanden wurden 473 Entlehnungen für die Analyse vorgelegt.

Für die Untersuchung wurde die schriftliche Form der Befragung gewählt. Bei der mündlichen Form wäre es zu kompliziert und anstrengend für die Vpn, 473 Wörter durch direktes Hören oder vom Tonband zu rezipieren und sofort in einem gewissen Tempo eine Antwort zu geben. Bei der schriftlichen Form ist es

dagegen einfacher, eine größere Zahl von Informanten (für diese Arbeit waren es 552) zu befragen. Eine schriftliche Befragung kommt nur dann in Frage, wenn der Fragebogen als Frage-Antwort-Medium kaum Schwierigkeiten für die Informanten bereitet.

Der Fragebogen mit den Entlehnungen wurde dem Probanden übergeben und gleichzeitig das Ziel der Untersuchung und die Arbeitsmethode erklärt: Entlehnungen, die der Befragte vom Blatt versteht, muss er ankreuzen und ihm „unbekannte“/„unverständliche“ Wörter unmarkiert lassen.

Für die Befragung wurden grundsätzlich positiv eingestellte Probanden gesucht, d.h. solche, die für die Beantwortung Zeit und Interesse hatten. Die Antworten wurden anonym ausgewertet. Das Alter und der Ort, wo der Informant erzogen wurde oder wohnt, haben für die Untersuchung eine wichtige Rolle gespielt. Beruf, Ausbildung, Eltern und andere sozialen Parameter, wie z.B. die Migration vor und nach der Wende, wurden nicht in die Untersuchung einbezogen, weil es unmöglich ist, alle erhaltenen Daten mit den statistischen Methoden im Rahmen nur einer Arbeit zu bearbeiten. Außerdem war es auch ausgeschlossen, allein, ohne Arbeitsgruppe, Fragebögen für verschiedene Schichten auszuarbeiten und auszuwerten.

Schleswig-Holstein, Baden-Württemberg, Bremen, Hamburg und Niedersachsen repräsentieren in meiner Untersuchung die westlichen Länder. Sachsen, Berlin (Ost) und Brandenburg die östlichen Regionen. Aus jeder Region wurde ungefähr eine gleiche Zahl von Probanden für die Befragung gewählt: Schleswig-Holstein – 88, Baden-Württemberg – 87, Bremen/Hamburg – 101, Niedersachsen – 88, Sachsen – 85, Berlin/Brandenburg – 103. Da der Einfluß der russischen Sprache in den westlichen und östlichen Ländern Deutschlands unterschiedlich war, u. zw. im Westteil kleiner und im Ostteil größer, haben wir 66 % der Befragungen in westlichen Regionen durchgeführt, weil die Qualität und Quantität russischer (u.a. slawischer) Entlehnungen in diesen Regionen bisher nicht so ausführlich wie in den östlichen Ländern analysiert worden ist. Nach der Bearbeitung der Fragebögen wurden für jedes Wort die folgenden Arbeitstabellen entworfen:

Beispiel: Die Tabelle *Statistik* für die Entlehnung *der Irbis*

	16–29	30–59	60–82
Schleswig-Holstein	4 23	8 25	9 19
Baden-Württemberg	3 28	6 21	4 25
Bremen/Hamburg	13 28	11 21	14 14
Niedersachsen	3 23	4 26	8 24

	16–29	30–59	60–82
Sachsen	4 23	3 27	9 19
Berlin/Brandenburg	12 23	19 13	15 21

Die sechs waagerechten Reihen in Spalte 1 stellen die Länder dar und die drei weiteren Spalten zeigen die Altersgrenzen. Jede obere Zahl in der Reihe gibt uns die Information, dass dieses Wort den Befragten bekannt ist. Unten steht die Zahl derer, die das Wort vom Blatt und ohne Kontext nicht verstehen. Beispielsweise ist das vorgeschlagene Wort *der Irbis* in Schleswig-Holstein in der Altersgruppe 16–29 nur für 4 Personen bekannt und für 23 Informanten dieser Untergruppe ist die Entlehnung fremd. Für den sinnvollen Einsatz der Inferenzstatistik in dieser Untersuchung ist es erforderlich, dass die „erkundeten“ Ziele und die Arbeitshypothese im Weiteren an den Beispielen der Entlehnungen des Korpus durch das Chi-Quadrat-Kriterium überprüft und näher betrachtet werden.

Das Merkmal AC (das Alter)

Beispiel 1

Untersucht werden soll, ob der Bekanntheitsgrad eines gegebenen Wortes vom Alter des Respondenten abhängig ist. Die erhaltenen Angaben wurden durch das Chi-Quadrat-Kriterium bearbeitet und im Weiteren analysiert. Bei dem untersuchten Merkmal handelt es sich um eine stetige konstituierliche Variable, wobei das Alter und der Wohnort als Verhältnisskala auftreten.

Wir gehen zum Beispiel vom Wort „die Arbuse“ aus und prüfen, ob die Bekanntmachung des Wortes vom Alter der Respondenten abhängig ist. Dazu wurden folgende Bezeichnungen eingeführt: wenn der Proband das Wort versteht, dann hat das Wort das Merkmal *A*, wenn nicht, das Merkmal „kein *A*“ oder „⊕“, das bedeutet, daß das Merkmal *A* fehlt.

Alle Probanden werden in drei Gruppen eingeordnet. Die Gruppe C_1 – Alter von 16 bis 29, die Gruppe C_2 – Alter von 30 bis 59, die Gruppe C_3 – Alter von 60 bis 82. So können die Untersuchungsergebnisse, die mit dem Wort „die Arbuse“ und dem Alter des Respondenten verbunden sind, in Form einer Tabelle von abhängigen Merkmalen dargestellt werden, wobei das Merkmal *A* die Wortgebräuchlichkeit ist und *C* das Alter der Untersuchungsteilnehmer darstellt.

die Arbuse

<i>CA</i>	<i>A</i>	⊕
C_1	$n_{1,1} = 1$	$n_{1,2} = 186$
C_2	$n_{2,1} = 5$	$n_{2,2} = 179$
C_3	$n_{3,1} = 5$	$n_{3,2} = 176$

Eine Person im Alter von 16 bis 29 versteht das Wort „die Arbuse“, 186 Personen verstehen das Wort nicht. 5 Personen im Alter von 30 bis 59 verstehen dieses Wort und 179 nicht. Im Alter von 60 bis 82 verstehen das Wort 5 Personen und 176 ist es fremd.

Die zu prüfende statistische Hypothese besagt folgendes: Das Merkmal *A* ist vom Merkmal *C* unabhängig, in unserer Situation hängt die Gebräuchlichkeit des Wortes „die Arbuse“ vom Alter des Respondenten nicht ab. Durch die Zahlen aus der Tabelle wird die χ^2 -Statistik nach der Formel 1:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^s \frac{(n_{ij} - np_i \cdot p_j)^2}{np_i \cdot p_j}$$

[Cramer 1976: 478–479], berechnet:

$$\chi^2 = n \left(\sum_{i=1}^3 \sum_{j=1}^2 \frac{n_{ij}^2}{n_i \cdot n_j} - 1 \right) = n \left(\frac{n_{11}^2}{n_1 \cdot n_{.1}} + \frac{n_{12}^2}{n_1 \cdot n_{.2}} + \frac{n_{21}^2}{n_2 \cdot n_{.1}} + \frac{n_{22}^2}{n_2 \cdot n_{.2}} + \frac{n_{31}^2}{n_3 \cdot n_{.1}} + \frac{n_{32}^2}{n_3 \cdot n_{.2}} - 1 \right)$$

Vorhergehend wurden

$$n = \sum_{i=1}^3 \sum_{j=1}^2 n_{ij} = n_{11} + n_{12} + n_{21} + n_{22} + n_{31} + n_{32} = 552$$

und

$$n_1 = n_{11} + n_{12} = 187;$$

$$n_2 = n_{21} + n_{22} = 184;$$

$$n_3 = n_{31} + n_{32} = 181;$$

$$n_{.1} = n_{11} + n_{21} + n_{31} = 11;$$

$$n_{.2} = n_{12} + n_{22} + n_{32} = 541$$

ausgerechnet.

Zur Übereinstimmung mit dem χ^2 -Kriterium ist es notwendig, die Irrtumswahrscheinlichkeit – das Bedeutungs-niveau α – zu wählen. Hypothesen, die unter Zugrundelegung einer Irrtumswahrscheinlichkeit von $\alpha = 0.05 = 5\%$ (Signifikanzniveau) zurückgewiesen werden, stützen Aussagen, von denen man sagt, sie „liegen auf dem 5-Prozent-Niveau der Verlässlichkeit“. Entsprechend bezeichnet man Aussagen, denen eine Irrtumswahrscheinlichkeit $\alpha = 0.01 = 1\%$ zugrunde liegt, als „auf dem 1-Prozent-Niveau der Verlässlichkeit“ gesichert [Claus/Ebner 1974: 186–191].

Außerdem muss man die entsprechende Anzahl der Freiheitsgrade zählen, wie in unserem Beispiel $df = (3-1) \cdot (2-1) = 3$. Dann finden wir durch die Tabelle der χ^2 -Verteilung [Cramer 1976 : 610] den kritischen Wert $\chi_{0,05,2}^2 = 5.991$ über den Schnittpunkt der dritten Zeile und der elften Spalte, der der Irrtumswahrschein-

lichkeit 0.05 entspricht. Weiter wird der berechnete Wert der χ^2 -Statistik mit dem kritischen Wert verglichen.

Beim Prüfen von Hypothesen auf Stichprobenbasis können zwei Arten von Fehlschlüssen eintreten:

– wenn der berechnete Wert der Statistik χ^2 kleiner als χ_0^2 ($\chi^2 < \chi_0^2$) ist, wird eine Hypothese über die Unabhängigkeit der Merkmale auf Grund der Bedeutsamkeit angenommen.

– Wenn χ^2 gleich oder größer als χ_0^2 ($\chi^2 \geq \chi_0^2$) ist, wird die Hypothese über die Unabhängigkeit der Merkmale auf Grund der Bedeutsamkeit abgelehnt.

Wenn $\chi^2 < \chi_0^2$, in unserer Situation $3.079 < 5.991$, bedeutet das, dass eine Nullhypothese angenommen wird, d. h. die Gebräuchlichkeit des Wortes „die Arbeit“ hängt vom Alter der Respondenten nicht ab.

Beispiel 2

Nehmen wir noch ein anderes Beispiel: das Wort „der/das Kolchos“ und prüfen wir, ob der Bekanntheitsgrad des Wortes vom Alter der Respondenten abhängt. Die entsprechende Tabelle der Merkmale sieht so aus:

der/das Kolchos

C/A	A	\oplus
C_1	$n_{11} = 56$	$n_{12} = 131$
C_2	$n_{21} = 82$	$n_{22} = 102$
C_3	$n_{31} = 96$	$n_{32} = 85$

In diesem Fall wird die χ^2 -Statistik nach der oben genannten Formel 1 berechnet, dabei ist

$$n = n_{11} + n_{12} + n_{21} + n_{22} + n_{31} + n_{32} = 552$$

und

$$n_{1.} = n_{11} + n_{12} = 187;$$

$$n_{2.} = n_{21} + n_{22} = 184;$$

$$n_{3.} = n_{31} + n_{32} = 181;$$

$$n_{.1} = n_{11} + n_{21} + n_{31} = 234;$$

$$n_{.2} = n_{12} + n_{22} + n_{32} = 318.$$

Der χ^2 -Wert ist 20,617. Für dieselbe Irrtumswahrscheinlichkeit und die Anzahl der Freiheitsgrade bleibt der kritische Wert derselbe, u. zw. $\chi_0^2 = 5.991$. Vergleichen wir diese zwei Werte: $\chi^2 > \chi_0^2 - 20.617 > 5.991$. Das bedeutet, daß die Nullhypothese abgelehnt wird und der Bekanntheitsgrad dieses Wortes vom Alter des Respondenten abhängt. In diesem Falle wird dieses Abhängigkeitsmaß durch den Kontingenzkoeffizienten der Merkmale charakterisiert. Der Kontingenzkoeffizient informiert über den Grad des Zusammenhangs zwi-

schen zwei Variablen. Sind die Variablen nominalskaliert (entstammen die Merkmale qualitativ verschiedenen Gegenstandsklassen), dann wird ihr Zusammenhang durch einen Kontingenzkoeffizienten ausgedrückt [Claus/Ebner 1974: 115-116]. Nach der Formel [Kendall, Stuart 1973: 747]

$$C = \sqrt{\frac{\chi^2}{n \min(r-1)(m-1)}}$$

die Kramer vorgeschlagen hat, wird der Koeffizient bestimmt. n ist die Zahl der befragten Respondenten ($n = 552$), r ist die Altersgruppenzahl ($r = 3$), und m ist die Zahl der Gruppenmerkmale ($m = 2$). Als Ergebnis erhalten wir, dass $C = 0.193$, was für eine schwache Abhängigkeit dieser Merkmale spricht. Dieser Koeffizient wird von 0 für unabhängige Merkmale bis +1 für funktionell verbundene Merkmale (im Fall der vollen Verbindung) angewendet ($0 \leq C < +1$).

Schlußüberlegungen

Ausgehend von den erhaltenen Daten wollen wir fragen:

1. Wie ist der Bekanntheitsgrad der Entlehnungen in den einzelnen Altersgruppen (1–3)?
2. Gibt es einen statistisch signifikanten Unterschied beim Alter in der Gruppe der Probanden aus dem Osten, d.h. kennen die jüngeren Generationen weniger russische (u.a. slawische) Entlehnungen als die älteren?

Bevor die gegenwärtige Situation der Abhängigkeit/Unabhängigkeit entlehnter Wörter des untersuchten Korpus vom Alter der Vpn im Weiteren konstatiert wird, muss man diese Fragen beantworten. Zu 1) Bei der Arbeitshypothese sind wir davon ausgegangen, dass die Auswirkungen der staatlichen Teilung Deutschlands im Sprachgebrauch russischer (u.a. slawischer) Wörter über mehr als eine Generation zu neuen Unterschieden führten. Die Sprachveränderungen (bei der Untersuchung handelt es sich um den Bekanntheitsgrad entlehnter Wörter im Osten und Westen) werden in einem solchen Zeitraum als Generationsunterschiede im lexikalischen Bereich betrachtet, die man auch in relativ homogenen Sprechergemeinschaft beobachten kann. Natürlich darf man nicht vergessen, dass mit der staatlichen Teilung ab 1949 und gänzlich ab 1961 abrupt die Sprachbeziehungen zwischen den Regionen eingeschränkt wurden. Die Wirkung der russischen Sprache auf die östlichen Regionen und der englischen auf die westlichen wurde mehr oder weniger in den regionalen Generationsunterschieden fixiert. An dieser Stelle muss betont werden, dass die regionale Diversifikation des deutschen Sprachgebietes nicht neu ist. Damit ist es gemeint, dass sie die Elemente aller sogenannter Stammesdialekte (Alemannisch, Bairisch, Fränkisch und Altsächsisch/Niederdeutsch) umfasst,

worauf hier nicht näher eingegangen zu werden braucht.

Ausgehend von den erhaltenen Daten sind 28 % russischer (u.a. slawischer) Entlehnungen vom Alter der Untersuchungsteilnehmer abhängig. Die erste Altersgruppe (16–29) wird durch prozentuale Verhältnis 29 % zu 34 % (genauer 29.31 % zu 33.63 %) – West – Ost – charakterisiert. 33 % im Westen und 44 % im Osten (32.67 % zu 44.04 %) beweisen bei der zweiten Altersgruppe (30–59), dass der russische Spracheinfluss auf diese Generation im Osten stärker als im Westen war. Der Bekanntheitsgrad der Entlehnungen in der Altersgruppe (60–82) ist 36 % zu 48 % (35.46 % zu 47.73 %).

Wir kommen zur zweiten Eingangsfrage, ob es einen statistisch signifikanten Unterschied beim Alter in der Gruppe der Probanden aus dem Osten gibt. Nach der vergleichenden Analyse der erhaltenen Daten ist der Unterschied sichtbar: 34 % – 44 % – 48 %.

Daraus ergibt sich folgende Schlussfolgerung: die Bekanntmachung der zu prüfenden russischen (u.a. slawischen) Entlehnungen hängt bei 28 % der untersuchten Wörter vom Alter der Respondenten ab.

Das Merkmal AB (der Wohnort)

Es wird überprüft, wie die untersuchten Wörter mit dem Wohnort der Informanten verbunden sind, d.h., ob die regionale Gegebenheit irgendeine Rolle für den Bekanntheitsgrad dieser Wörter in der Rede spielt.

Das Merkmal A ist mit dem Bekanntheitsgrad des Wortes verknüpft und das Merkmal B hat sechs Bedeutungen: B₁ – Schleswig-Holstein, B₂ – Baden-Württemberg, B₃ – Bremen/ Hamburg, B₄ – Niedersachsen, B₅ – Sachsen, B₆ – Ost-Berlin/Brandenburg.

B₁, B₂, B₃, B₄ stehen für die westlichen Länder. B₅ und B₆ repräsentieren die östlichen Länder. Bremen und Hamburg auf der eine Seite, und Berlin und Brandenburg auf der anderen Seite, bilden in dieser Studie eine Einheit. Das hat keine politischen Gründe. Zuerst haben wir eine Befragung von Informanten durchgeführt und danach diese Länder (Hamburg/Bremen, Berlin/Brandenburg) zusammengefasst.

Beispiel 1

Im ersten Beispiel schauen wir uns das Wort „der Sputnik“ an. In der folgenden Tabelle sind die entsprechenden Werte zusammengestellt.

der Sputnik

B/A	A	⊕
B ₁	n ₁₁ = 86	n ₁₂ = 2
B ₂	n ₂₁ = 76	n ₂₂ = 11
B ₃	n ₃₁ = 95	n ₃₂ = 6
B ₄	n ₄₁ = 84	n ₄₂ = 4

B/A	A	⊕
B ₅	n ₅₁ = 84	n ₅₂ = 1
B ₆	n ₆₁ = 100	n ₆₂ = 3

Ausgehend von den Daten in der Tabelle kann man folgern, dass 86 Personen im Alter von 16 bis 82 aus Schleswig-Holstein das Wort „Sputnik“ verstehen und dass für 2 dieses Wort unbekannt ist. Im Baden-Württemberg verstehen das Wort 76 Personen und 11 nicht. Für 95 Personen im Alter von 16 bis 82 aus Bremen und Hamburg ist dieses Wort verständlich, für 6 ist das Wort unverständlich. 84 Respondenten aus Niedersachsen und dieselbe Zahl aus Sachsen verstehen das Wort. 4 Respondenten aus Niedersachsen und nur ein befragter Informant aus Sachsen verstehen „der Sputnik“ nicht. In Berlin und Brandenburg ist das Wort für 100 Respondenten verständlich und für 3 unverständlich.

Die Nullhypothese lautet: das Merkmal A hat mit dem Merkmal B nichts zu tun, in unserem Beispiel ist der Bekanntheitsgrad des Wortes mit dem Wohnort der Respondenten nicht verknüpft.

Nach Cramer [1976: 479] wird die χ²-Formel angewendet. Vorhergehend wurden alle notwendigen Werte für die arithmetische Grundrechenart berechnet, u. zw.

$$n = \sum_{i=1}^6 \sum_{j=1}^2 n_{ij} = n_{11} + n_{12} + n_{21} + n_{22} + n_{31} + n_{32} + n_{41} + n_{42} + n_{51} + n_{52} + n_{61} + n_{62} = 552.$$

$$n_1 = n_{11} + n_{12} = 88;$$

$$n_2 = n_{21} + n_{22} = 88;$$

$$n_3 = n_{31} + n_{32} = 11;$$

$$n_4 = n_{41} + n_{42} = 88;$$

$$n_5 = n_{51} + n_{52} = 85;$$

$$n_6 = n_{61} + n_{62} = 103;$$

$$n_{.1} = n_{11} + n_{21} + n_{31} + n_{41} + n_{51} + n_{61} = 525;$$

$$n_{.2} = n_{12} + n_{22} + n_{32} + n_{42} + n_{52} + n_{62} = 27.$$

Wie es schon gesagt wurde, sieht die Hauptstatistik für die Merkmale A und B so aus:

$$\chi^2 = \left(\sum_{i=1}^6 \sum_{j=1}^2 \frac{n_{ij}^2}{n_i n_j} - 1 \right) n.$$

Die entsprechende Anzahl der Freiheitsgrade ist 5 df_{AB} = (6–1)(2–1). Danach finden wir durch die Tabelle der χ²-Verteilung der Irrtumswahrscheinlichkeit von α = 0,05 des kritischen Wertes χ_{0,05,5} = 11.070 = χ_{α,f} = χ₀².

Jetzt wird der berechnete Wert der χ²-Statistik mit dem kritischen Wert verglichen.

Als Kriterium bleibt:

- wenn $\chi^2 < \chi_0^2$, dann Annahme von H_0 ;
- wenn $\chi^2 \geq \chi_0^2$, dann Zurückweisung von H_0 .

In unserem Beispiel $\chi^2 = 16.186$ ist größer als $\chi_0^2 = 11.070$ ($\chi^2 > \chi_0^2$), das bedeutet, dass der Bekanntheitsgrad des Wortes vom Wohnort der Respondenten abhängig ist. In diesem Falle kann die Abhängigkeit durch den Koeffizienten von Kramer charakterisiert werden

$$C = \sqrt{\frac{\chi^2}{n \min(r-1)(m-1)}}$$

wo n die Zahl der befragten Informanten ist ($n = 552$), r als die Zahl der Länder auftritt ($r = 6$), und m die Zahl der Gruppenmerkmale ist ($m = 2$). Als Ergebnis erhalten wir, dass C 0.171 gleich ist. Der Koeffizient zeigt die schwache Abhängigkeit.

Beispiel 2

Als weiteres Beispiel wählen wir das Wort *die Balalaika*.

Die entsprechende Tabelle hat folgendes Aussehen:

die Balalaika

B/A	A	\oplus
B_1	$n_{11} = 70$	$n_{12} = 18$
B_2	$n_{21} = 64$	$n_{22} = 23$
B_3	$n_{31} = 87$	$n_{32} = 14$
B_4	$n_{41} = 70$	$n_{42} = 18$
B_5	$n_{51} = 69$	$n_{52} = 16$
B_6	$n_{61} = 72$	$n_{62} = 31$

Die χ^2 -Statistik wird durch die Formel 1 erhalten, dabei gilt

$$n = n_{11} + n_{12} + n_{21} + n_{22} + n_{31} + n_{32} + n_{41} + n_{42} + n_{51} + n_{52} + n_{61} + n_{62} = 552.$$

$$n_1 = n_{11} + n_{12} = 88;$$

$$n_2 = n_{21} + n_{22} = 87;$$

$$n_3 = n_{31} + n_{32} = 101;$$

$$n_4 = n_{41} + n_{42} = 88;$$

$$n_5 = n_{51} + n_{52} = 85;$$

$$n_6 = n_{61} + n_{62} = 103;$$

$$n_{.1} = n_{11} + n_{21} + n_{31} + n_{41} + n_{51} + n_{61} = 432;$$

$$n_{.2} = n_{12} + n_{22} + n_{32} + n_{42} + n_{52} + n_{62} = 120.$$

Die entsprechende Anzahl der Freiheitsgrade ist $df_{AB} = (6-1)(2-1) = 5$. Der kritische Wert χ_0^2 ist derselbe geblieben und ist 11.070 gleich. $\chi^2 = 9.637$. Der χ^2 -Wert wird mit dem χ_0^2 -Wert verglichen. Daraus folgt, dass χ^2 (9.637) kleiner als χ_0^2 (11.070) ist. Das bedeutet, dass die Nullhypothese angenommen wird und die Merkmale unabhängig sind.

Nach der Bearbeitung der Ergebnisse für die Merkmale *AB*, kann man feststellen, dass für 148 Wörter ($\approx 31\%$) die Nullhypothese angenommen wird und die Merkmale unabhängig sind, weil $\chi^2 < \chi_{\alpha, f}^2$. Für die anderen 325 Wörter kann man dagegen nicht sagen, dass die Merkmale unabhängig sind, weil $\chi^2 > \chi_{\alpha, f}^2$, und dann kann man die Kontingenzkoeffizienten – C berechnen, die die Abhängigkeit bestimmen.

Gegenwärtig wird die Frage nach russischen Wörtern im Deutschen kaum noch diskutiert. Bereits Bielfeldt [1982: IX] hatte betont: „Eine sowohl alles umfassende und zugleich geschlossene Monographie über die slawischen, russischen Wörter im Deutschen aller Zeiten, geographischen Bereiche und Sprachschichten gibt es nicht und wird es vielleicht nie geben“. Wünschenswert und nötig scheint mir aber, eine weitere interdisziplinäre Zusammenarbeit von Germanisten und Slavisten auf dem Gebiet der Entlehnungen, die den Einfluss der Russizismen (u.a. slawischer Integrate) auf die deutsche Sprache genauer untersucht, und zwar in allen Kommunikationsbereichen. Wichtig wären auch Forschungsarbeiten zu den Stufen der Interferenz, zu den Fragen der Eindeutschung slawischer Integrate und zur Aufgabe, wieweit Russizismen in gemeinsprachlichen Texten die Kommunikation beeinflussen. Zur Zeit scheint in der Folge bestimmter politischer Entscheidungen auch ein anderer Problembereich untersuchungswert, u. zw. die Fragen, die mit dem Zuzug von Deutschen aus Russland in die BRD verbunden sind.

Literaturverzeichnis

1. G. Bellmann. Slavoteonica. Lexikalische Untersuchungen zum slawisch-deutschen Sprachkontakt im Ostmitteleutschen. Berlin, New York 1971.
2. H. H. Bielfeldt. Die slawischen Wörter im Deutschen. Ausgewählte Schriften 1950–1978. Leipzig, 1982.
3. J. Bortz. Statistik für Sprachwissenschaftler. Berlin, Heidelberg, New York, London 1993.
4. G. Claus, H. Ebner. Grundlagen der Statistik für Psychologen, Pädagogen und Soziologen. Berlin, 1974.
5. H. Cramer. Matematičeskie metodi statistiki. Moskva, 1976.

6. M. Kendall/A. Stuart. Statisticheskie vyvody i svjazi. Moskva, 1973.
7. E. Opel'baum. Vostocnoslavjanskije lexiceskie elementy v nemeckom jazyke. Kiev, 1971.

Wörterbücher

1. H. Klien. Fremdwörterbuch. Leipzig, 1954.
2. Duden. Deutsches Universal Wörterbuch A–Z. Mannheim; Wien; Zürich, 1989.
3. Duden. Fremdwörterbuch. Der Duden in 10 Bänder. Bd. 5. Mannheim; Wien; Zürich, 1990.
4. Duden. Das große Fremdwörterbuch. Mannheim, Leipzig, 1994.
5. R. Klappenbach/W. Steiniz. 1.Bd.-1967, 2.Bd.-1968, 3.Bd.-1970, 4.Bd.-1975, 5.Bd.-1976, 6.Bd.-1978: Wörterbuch der deutschen Gegenwartssprache. Berlin.
6. F. Kluge. Etymologisches Wörterbuch der deutschen Sprache. Berlin, New York, 1975.
7. W. Pfeifer. Etymologisches Wörterbuch des Deutschen. Berlin, 1989.
8. G. Wahrig. Deutsches Wörterbuch. München, 1986.