

# **Quantitative Studien zum Esperanto unter besonderer Berücksichtigung der Wortbekanntheit**

Vom Fachbereich II: Sprach-, Literatur- und Medienwissenschaften  
der Universität Trier genehmigte

Dissertation

zur Erlangung des akademischen Grades

Doktor der Philosophie

vorgelegt von

Andreas Kück

Erster Gutachter:

Professor Dr. Reinhard Köhler

Zweiter Gutachter:

Professor Dr. Peter Grzybek

Tag der mündlichen Prüfung:

23. April 2009

Rechtenfleth, 2009

## **Vorwort**

An dieser Stelle möchte ich all jenen Menschen danken, die mich bei dieser Arbeit unterstützt haben. Ich danke Herrn Professor Dr. Reinhard Köhler für seine Bereitschaft, diese Arbeit zu betreuen, und für seine wertvollen und hilfreichen Ratschläge. Herrn Professor Dr. Peter Grzybek danke ich für die Übernahme des Zweitgutachtens. Den Esperanto-Sprechern, die sich an der Umfrage zur Lösung von Aufgabenstellungen dieser Arbeit beteiligt haben, danke ich ebenfalls. Meinen Eltern danke ich dafür, dass sie die Voraussetzungen dafür geschaffen haben, dass ich die erforderliche Ausbildung erhielt, ohne die diese Arbeit nicht möglich gewesen wäre. Meinem Großvater danke ich dafür, dass er mir, als ich noch ein Kind war, von der Sprache Esperanto erzählte und so den Grundstein für diese Arbeit legte.

Die hier publizierte Arbeit entspricht meiner am 03.11.2008 dem Fachbereich II der Universität Trier vorgelegten Dissertation.

Rechtenfleth, Mai 2009

Andreas Kück

## Inhaltsverzeichnis

1	Einleitung.....	8
1.1	Ziel und Relevanz dieser Arbeit (Übersicht).....	8
1.2	Ziel der phonemstatistischen Untersuchungen.....	9
1.3	Ziel der Untersuchungen zur Morphologie und Sprachtypologie.....	9
1.4	Ziel der Untersuchungen zur Syntax.....	9
1.5	Ziel der Untersuchungen zur Lexik.....	10
1.6	Ziel der Untersuchungen zur Semantik.....	10
1.7	Ziel der Untersuchungen zur Worthäufigkeit.....	10
1.8	Ziel der Untersuchungen zur Wortbekanntheit.....	10
2	Ausgangssituation (Forschungsstand).....	13
2.1	Grundlegendes zu Esperanto und seiner sprachwissenschaftlichen Erforschung.....	13
2.2	Phonemstatistik.....	14
2.3	Morphologie und Sprachtypologie.....	19
2.3.1	Freie lexikalische und freie grammatische Morpheme.....	19
2.3.2	Gebundene lexikalische Morpheme.....	19
2.3.3	Gebundene grammatische Morpheme.....	20
2.3.4	Wortbildung durch Derivation.....	21
2.3.5	Wortbildung durch Komposition.....	22
2.3.6	Suppletivismus, unikale, diskontinuierliche und Portmanteau- Morpheme.....	22
2.3.7	Spezielle Arten der Komposition.....	22
2.3.8	Grundlegende Wortart des lexikalischen Morphems.....	23
2.3.9	Korrelativwörter.....	23
2.3.10	Häufigkeitsverteilung von Morphemstrukturen.....	25
2.3.11	Aussagen zur sprachtypologische Einordnung.....	26
2.4	Syntax.....	27
2.4.1	Einführendes zur Syntax.....	27
2.4.2	Struktur eines Esperanto-Satzes.....	27
2.5	Lexikon des Esperanto.....	30
2.6	Lexikalische Semantik.....	33

2.6.1	Einführendes zur lexikalischen Semantik des Esperanto.....	33
2.6.2	Polysemie.....	34
2.6.3	Homonymie und Paronymie.....	35
2.6.4	Synonymie.....	36
2.6.5	Antonymie, Meronymie, Hyponymie und Hyperonymie.....	38
2.6.6	Farben.....	39
2.6.7	Bedeutungsverengung, -erweiterung und Konnotationen.....	41
2.6.8	Kollokationen.....	42
2.7	Worthäufigkeit.....	43
2.7.1	Offizielle Basis-Wortwurzelsammlung (BRO).....	43
2.7.2	Andere Worthäufigkeitslisten.....	43
2.8	Zusammenhang zwischen Bekanntheit und Häufigkeit von Wörtern und Phraseologismen.....	45
3	Weiterführende quantitative Studien zum Esperanto: Phonologie, Morphologie, Syntax, Lexik, lexikalische Semantik, Worthäufigkeit.	47
3.1	Phonemhäufigkeit.....	47
3.1.1	Ziel der Untersuchung zur Phonemhäufigkeit.....	47
3.1.2	Anwendung verschiedener Modelle zur Beschreibung der Phonemhäufigkeit.....	47
3.1.3	Wertung von Affrikaten.....	50
3.1.4	Zusammenfassung: Untersuchung zur Phonemhäufigkeit.....	50
3.2	Sprachtypologische Einordnung.....	52
3.2.1	Ziel der Untersuchung zur sprachtypologischen Einordnung.....	52
3.2.2	Isolierende Merkmale.....	52
3.2.3	Synthetische Merkmale.....	55
3.2.4	Inkorporierende und polysynthetische Merkmale.....	58
3.2.5	Zusammenfassung: Untersuchung zur sprachtypologischen Einordnung.....	58
3.3	Detaillierte Analyse zur Häufigkeitsverteilung von Morphemstrukturen des Esperanto.....	59
3.4	Rechts- und Linkseinbettung von Phrasen und Relativsätzen.....	60
3.4.1	Ziel der Untersuchung zur Rechts- und Linkseinbettung von Phrasen und Relativsätzen.....	60
3.4.2	Flexionsphrase und Einbettung von Subjekt-Nominalphrasen.....	60

3.4.3	Einbettung von Objekt-Nominalphrasen.....	62
3.4.4	Einbettung von Adjektivphrasen.....	62
3.4.5	Einbettung von Präpositionalphrasen.....	65
3.4.6	Einbettung von Adverbphrasen.....	68
3.4.7	Einbettung von Negationen.....	70
3.4.8	Einbettung von Relativsätzen.....	72
3.4.9	Zusammenfassung: Untersuchung zur Rechts- und Linkseinbettung von Phrasen und Relativsätzen.....	75
3.5	Zahl der grundlegenden und offiziellen Lexeme.....	76
3.6	Polysemie.....	77
3.7	Zusammenhang zwischen offizieller Basis-Wortwurzelsammlung (BRO) und Häufigkeiten nach dem Dietze-Korpus.....	79
4	Weiterführende quantitative Studien zum Esperanto: Wortbekanntheit.....	82
4.1	Annahmen zur Wortbekanntheit.....	82
4.2	Vergleichbare Aufgabenstellung in einem anderen Fachgebiet.....	82
4.3	Quantitative Beschreibung der Häufigkeit.....	83
4.4	Quantitative Beschreibung der Bekanntheit.....	84
4.4.1	Angestrebte Aussageform zur Bekanntheit.....	84
4.4.2	Eine mögliche Vorgehensweise.....	84
4.4.3	Logistische Regression nach Berens.....	87
4.4.4	Vertiefung der mathematischen Grundlagen.....	95
4.4.5	Finden der maximalen Mutmaßlichkeit.....	98
4.5	Entwurf einer Befragung und Ausführung.....	100
4.6	Antworten in der Befragung.....	106
4.7	Auswertung der fünf Fragebogen.....	108
4.8	Ähnlichkeitsanalyse.....	109
4.9	Gesamtauswertung.....	112
4.10	Auswertung nach einzelnen Sprachfamilien und -zweigen.....	112
4.10.1	Motivation.....	112
4.10.2	Muttersprache aus nicht-indoeuropäischer Sprachfamilie.....	113
4.10.3	Muttersprache aus indoeuropäischer Sprachfamilie.....	114
4.10.4	Muttersprache Chinesisch oder Japanisch.....	114
4.10.5	Muttersprache Japanisch.....	115

4.10.6	Muttersprache aus uralischer Sprachfamilie.....	116
4.10.7	Muttersprache aus dem baltischen oder slawischen Sprachzweig.....	117
4.10.8	Muttersprache aus dem germanischen oder romanischen Sprachzweig.....	118
4.10.9	Muttersprache aus dem romanischen Sprachzweig.....	119
4.10.10	Zusammenfassung und Kommentierung der Auswertungen.....	121
4.11	Selbsteinschätzung des eigenen Esperanto-Anwendungsvermögens.....	122
4.12	Prüfung der Modell-Tauglichkeit.....	124
4.12.1	Devianz-Test.....	124
4.12.2	Vergleich mit einem alternativen Modell.....	124
5	Zusammenfassung.....	126
5.1	Allgemeines zu dieser Arbeit.....	126
5.2	Phonologie.....	127
5.3	Morphologie und Sprachtypologie.....	128
5.4	Syntax.....	129
5.5	Lexik.....	130
5.6	Semantik.....	130
5.7	Worthäufigkeit.....	130
5.8	Wortbekanntheit.....	130
5.9	Ausblick.....	132
	Anhang A: Anpassung des Yule-Modells (monophonematische Wertung von Affrikaten).....	134
	Anhang B: Anpassung des Yule-Modells (biphonematische Wertung von Affrikaten).....	135
	Anhang C: Anpassung des Zipfschen Gesetzes an die Häufigkeitsverteilung von Morphemstrukturen des Esperanto.....	136
	Anhang D: Länge von Lexemen und Polysemie.....	137
	Anhang E: HTML-Quellcode des Fragebogens A.....	141
	Anhang F: Einzelheiten der Befragungsergebnisse.....	143
	Anhang G: Umsetzung des Auswerte-Algorithmus.....	177
	Verwendete Abkürzungen und Symbole.....	186

Abbildungsverzeichnis.....	192
Tabellenverzeichnis.....	197
Literaturverzeichnis.....	199
Stichwortverzeichnis.....	210

# 1 Einleitung

## 1.1 Ziel und Relevanz dieser Arbeit (Übersicht)

Es sollen quantitative Studien zur Plansprache Esperanto zu ausgewählten Fragestellungen der Phonologie, Morphologie, Syntax, Lexik, Semantik, Worthäufigkeit und – als Schwerpunkt – zur Wortbekanntheit durchgeführt werden. Zur Untersuchung des Zusammenhangs zwischen Wortbekanntheit und Worthäufigkeit sowie der Selbsteinschätzung des eigenen Esperanto-Anwendungsvermögens durch seine Sprecher soll das Verfahren der logistischen Regression ausführlich dargestellt und als eine innovative Option auch zur Untersuchung analoger Fragestellungen angeboten werden.

Die Relevanz des Schwerpunkt-Themas Wortbekanntheit besteht auf zwei Ebenen:

- (a) *Bezogen auf Esperanto*: Die Untersuchung liefert unmittelbar statistisch fundierte Erkenntnisse darüber, inwieweit das Lexikon des Esperanto von dessen Sprechern beherrscht wird. Zudem werden Erkenntnisse darüber gewonnen, inwieweit die Tatsache, dass das Esperanto-Lexikon größtenteils aus germanischen und romanischen Sprachen abgeleitet wurde (also gewissermaßen „europäischlastig“<sup>1</sup> ist), nachteilig ist für Sprecher mit einer Muttersprache aus anderen Sprachfamilien und -zweigen. Weitere Erkenntnisse werden bezüglich der Selbsteinschätzung des eigenen Esperanto-Anwendungsvermögens durch seine Sprecher gewonnen.
- (b) *Allgemein*: Das Modell der logistischen Regression (mit Berücksichtigung der statistischen Sicherheit) soll der quantitativen Linguistik als innovative Option zur Datenanalyse (und entsprechender Auslegung von Erhebungen) angeboten werden.

---

<sup>1</sup> Sikosek (2003<sup>2</sup>: 65): „Interessierte, die sich über die Europäischkeit des Esperanto beklagen, wissen gelegentlich eine 'Lösung': Eine internationale Sprache soll die ganze Welt repräsentieren und Elemente – man meint meist Wortwurzeln – aus allen Sprachen haben.“



## **1.2 Ziel der phonemstatistischen Untersuchungen**

Ziel ist, das Inventar der Phoneme des Esperanto und ihre Häufigkeitsverteilung darzustellen. An Letztere sollen verschiedene Modellfunktionen angepasst werden, die sich bereits bei Phonem-Häufigkeitsverteilungen anderer Sprachen bewährt haben: Tuldava (1988); Zipf (1929, 1935, 1949); Yule (1924); Naranan, Balasubrahmanyam (1992a, b, 2000); Altmann (1993); Sigurd (1968); Good (1969); Altmann, Popescu<sup>2</sup>. Damit soll für die Plansprache Esperanto die in Strauss, Altmann, Best (2007) aufgestellte Hypothese überprüft werden, ob die geordnete Reihenfolge der Phonemhäufigkeiten einer regulären Wahrscheinlichkeitsfunktion oder regulären monoton abnehmenden Funktion folgt.

## **1.3 Ziel der Untersuchungen zur Morphologie und Sprachtypologie**

Ziel ist, die Morphologie des Esperanto darzustellen. Mittels quantitativer Methoden soll das Resümee von Sikosek (2003<sup>2</sup>: 204) bestätigt oder widerlegt werden, demzufolge nach Meinung der meisten Autoren und in passendster Verallgemeinerung Esperanto agglutinierend ist und einige Merkmale einer isolierenden Sprache aufweist.

Des Weiteren soll die Modellierbarkeit der Häufigkeitsverteilung der Morphemstrukturen des Esperanto mittels des Zipfschen Gesetzes detailliert dargestellt werden.

## **1.4 Ziel der Untersuchungen zur Syntax**

Die Rechts- und Linkseinbettung verschiedenartiger Phrasen und Relativsätze im Esperanto soll dargestellt werden. Dazu sollen einige Beispielsätze in X-Bar-Theorie analysiert werden. Beachtenswert ist, dass Esperanto als Plansprache zwar eine präskriptive Grundlage in Zamenhof (1991<sup>10</sup>) hat, es darin jedoch keine unmittelbaren expliziten Regeln zur Einbettung von Phrasen und Relativsätzen gibt. Etwaige diesbezügliche Regeln oder Empfehlungen können nur mittelbar aus Beispielsätzen hergeleitet werden.

---

<sup>2</sup> Dem Autor der vorliegenden Arbeit wurde in persönlicher Korrespondenz mitgeteilt, dass eine Veröffentlichung in Vorbereitung sei, die dieses (neue) Modell beinhaltet.

### **1.5 Ziel der Untersuchungen zur Lexik**

Die Modellierbarkeit der Zahl der grundlegenden und oficialisierten Lexeme des Esperanto soll untersucht werden.

### **1.6 Ziel der Untersuchungen zur Semantik**

Charakteristische Merkmale der lexikalischen Semantik der Plansprache Esperanto sollen dargestellt werden. Eine Besonderheit des Esperanto besteht darin, dass die Zuordnung von Bedeutungen zu den einzelnen Lexemen, die weitestgehend aposteriorisch aus Lexemen von Ethnosprachen abgeleitet wurden, zumindest am Anfang zunächst präskriptiv durch die Grundlagentext *Fundamento de Esperanto* (Zamenhof 1991<sup>10</sup>) erfolgte. Die bei anderen Sprachen festgestellte Gesetzmäßigkeit zwischen der Länge von Lexemen und Polysemie soll auch für Esperanto untersucht werden.

### **1.7 Ziel der Untersuchungen zur Worthäufigkeit**

Die Ergebnisse bereits vorhandener Untersuchungen zur Worthäufigkeit sollen miteinander hinsichtlich etwaiger Zusammenhänge untersucht werden.

### **1.8 Ziel der Untersuchungen zur Wortbekanntheit**

Es sollen untersucht werden:

- (a) der Zusammenhang zwischen Bekanntheit und Häufigkeit von Esperanto-Wörtern
- und
- (b) die Selbsteinschätzung des eigenen Esperanto-Anwendungsvermögens durch seine Sprecher<sup>3</sup>.

Dazu soll eine statistische Erhebung in Form einer Befragung erfolgen und mit dem mathematischen Verfahren der logistischen Regression ausgewertet werden. Dabei soll aus Daten dichotomer Ausprägung (z. B. „Wort bekannt“ oder „Wort nicht bekannt“) eine graduelle Ausprägung in Abhängigkeit von einer Variablen ermittelt werden, nämlich (a) die Wortbekanntheit in Abhängigkeit von

---

<sup>3</sup> Diese Einschätzung bezieht sich nicht nur auf die Wortbekanntheit. Aus praktischen Gründen werden diesbezügliche Untersuchungen in der vorliegenden Arbeit jedoch unter dem Thema Wortbekanntheit behandelt. Die statistische Erhebung und die Auswertung erfolgten im selben Rahmen.

der Worthäufigkeit und (b) die Selbsteinschätzung des eigenen Esperanto-Anwendungsvermögens in Abhängigkeit von der Esperanto-Lernzeit. Dabei soll auch das Häufigkeitsmerkmal der Wörter mit zuverlässiger Bekanntheit bestimmt werden. Für solche Wörter gilt:

Ihr Bekanntheitsgrad ist 90% bei einer statistischen Sicherheit von 95%. Das bedeutet: Bei mindestens 95 von 100 Texten sind je mindestens 90% dieser Wörter solchen Esperanto-Sprechern bekannt, die nach eigener Einschätzung Esperanto ausreichend gut verwenden.

Weiters soll ein analoges Merkmal bezüglich der Selbsteinschätzung des eigenen Esperanto-Anwendungsvermögens durch seine Sprecher ermittelt werden: die Jahre, bis der Sprecher glaubt, Esperanto zuverlässig anwenden zu können (also 90% bei 95% Vertrauensgrad).

Es handelt sich hierbei um ein logistisches Regressionsmodell. Es war zuvor in der Werkstoffprüftechnik zur Bestimmung von Riss-Auffindwahrscheinlichkeiten (bei vorgegebener statistischer Sicherheit) aus dichotomen Stichprobendaten („Riss gefunden“/„Riss nicht gefunden“) angewendet worden (Berens 1989). In innovativer Weise soll dieses Modell auf die o. g. dichotomen Fragestellungen der Sprachwissenschaft übertragen werden.

Die nachfolgend beschriebenen Haupthypothesen sollen auf Grundlage der Untersuchung getestet werden (es sei angemerkt, dass die Haupthypothesen (a) und (b) dabei ausschließlich für solche Esperanto-Sprecher gelten, die ihr Vermögen, Esperanto anzuwenden, als ausreichend gut einschätzen; die Haupthypothese (c) gilt auch für solche Esperanto-Sprecher, die ihr Vermögen als nicht ausreichend gut einschätzen):

- (a) Die Häufigkeit der zuverlässig bekannten Wörter ist nicht zufällig, sondern vorhersagbar. Sie entspricht genau einer wohl definierten Häufigkeitsgruppe, und zwar der Häufigkeitsgruppe Sieben auf einer neunstufigen Skala von Häufigkeitsgruppen der von der Esperanto-Akademie entwickelten Statistik („offizielle Basis-Wortwurzelsammlung“, BRO, Akademio de Esperanto [Hrsg.] 2007<sup>2</sup>: 23ff.).
- (b) Sei A die Häufigkeit der Esperanto-Wörter, die Menschen mit einer Muttersprache aus dem romanischen oder germanischen Sprachzweig

zuverlässig bekannt sind. Sei B die Häufigkeit der Esperanto-Wörter, die Menschen mit Muttersprache Chinesisch, Japanisch oder aus der finno-ugrischen Sprachfamilie zuverlässig bekannt sind. Es besteht kein signifikanter Unterschied zwischen A und B. Das heißt: Die naheliegende Vermutung, dass Sprecher mit einer Muttersprache aus dem romanischen oder germanischen Sprachzweig einen Vorteil haben und seltenere Esperanto-Wörter vergleichsweise besser kennen als Sprecher mit anderen Muttersprachen, soll durch die vorgesehene empirische Überprüfung widerlegt werden.

- (c) Bei einer statistischen Sicherheit (Vertrauensgrad) von 95% ist die Selbsteinschätzung des eigenen Esperanto-Anwendungsvermögens durch seine Sprecher auch nach 100 Lernjahren weniger als 90%.

## 2 Ausgangssituation (Forschungsstand)

### 2.1 Grundlegendes zu Esperanto und seiner sprachwissenschaftlichen Erforschung

Esperanto ist eine von Zamenhof veröffentlichte Welthilfssprache<sup>4</sup> (Zamenhof 1887) und gilt heute als die bisher erfolgreichste Plansprache (Sikosek 2003<sup>2</sup>: 284, Wendt 1977: 356). Ihr Wortschatz beruht weitgehend auf Wörtern, die Zamenhof gezielt aus dem Wortschatz germanischer und romanischer Sprachen ableitete (Zamenhof 1992<sup>18</sup>: 284). Grundlage des Esperanto ist das Werk „Fundamento de Esperanto“ („Grundlage des Esperanto“, auch kurz „Fundamento“ genannt; Zamenhof 1991<sup>10</sup>). Es wurde von den Delegierten der Welt-Esperantistenschaft beim ersten Esperanto-Weltkongress in Boulogne-sur-Mer, Frankreich, 1905 als Esperanto-Norm angenommen (im Rahmen der „Bulonja Deklaracio“, Zamenhof 1991<sup>10</sup>: 33ff.). Im „Antaŭparolo“ („Vorrede“, „Vorwort“; in Esperanto) wird der Zweck des „Fundamento“ definiert. Seine Unveränderlichkeit wird festgeschrieben (es darf noch „einmal für immer“ verändert werden, nachdem Esperanto weltweit akzeptiert worden ist und die Regierungen der wichtigsten Staaten einen Ausschuss eingerichtet haben, der über eine eventuelle letzte Änderung berät und sie ggf. beschließt). Bis dahin darf eine sprachnormierende Institution neue Wörter nach sorgfältiger Prüfung offiziell dem Grundwortschatz hinzufügen (in Form von „Oficialaj Aldonoj“, offiziellen Beifügungen). Die sprachnormierende Institution war zunächst das „Lingva Komitato“ (der „Sprachausschuss“) mit dessen Unterinstitution „Akademio de Esperanto“ („Esperanto-Akademie“) und – nach der Vereinigung beider unter dem Namen „Akademio de Esperanto“ – Letztere.

Das „Fundamento“ enthält folgende unmittelbar sprachnormierende Teile:

- (1) Die Grammatik: Sie wird in 16 Regeln in den Sprachen Französisch, Englisch, Deutsch, Russisch und Polnisch dargestellt.
- (2) „Ekzercaro“ („Übungssammlung“; in Esperanto, mit Wortübersetzungen und Erläuterungen in den vorgenannten Ethnosprachen). In einigen der

---

<sup>4</sup> Laut Wendt (1977: 355) versteht man unter Welthilfssprachen im Allgemeinen künstliche Sprachen.

Übungssätze werden auch sprachliche und z. T. über den Stoff der 16 Regeln hinausgehende Regelungen vermittelt (z. B. die Anwendung von Kasus und Präpositionen).

- (3) „Universala Vortaro“ („Allgemeines Wörterbuch“): Wörterverzeichnis Esperanto-Französisch-Englisch-Deutsch-Russisch-Polnisch (2769 Wörter).

In Zamenhof (1991<sup>10</sup>) werden die zehn Wortarten der traditionellen Grammatik verwendet (Substantiv, Adjektiv, Verb, Adverb, Präposition, Konjunktion, Pronomen, Artikel, Numerale, Interjektion).

Grundlegende Darstellungen des Esperanto (als Sprache) finden sich u. a. in Atanasov (1983), Janton (1993<sup>2</sup>), Manders (1947), Manders (1950), Nüssel (2000), Tonkin, Fettes (1996) und Tonkin (Hrsg.) (1997). Wichtige Arbeiten zu linguistischen Aspekten des Esperanto verfassten u. a. Wells (1978) und Wood (1982). Einzelne sprachwissenschaftliche Teilgebiete des Esperanto sind u. a. in folgenden Arbeiten dargestellt: Grammatik: Gledhill (2000<sup>2</sup>), Heil (1999), Kalocsay (1938<sup>2</sup>), Wennergren (2006); Wortbildung: Blanke (1981); Wortstellung: Jansen (2007); Morphologie: Bak (1991); Phraseologie: Fiedler (1999), Fiedler (2002). Eine Bibliographie universitärer Diplomarbeiten und Dissertationen zu Esperanto und Interlinguistik findet sich in Symoens (1989) und Symoens (1995).

## 2.2 Phonemstatistik

Esperanto hat 28 Phoneme (23 Konsonanten und fünf Vokale). Jedes Phonem wird durch ein ihm eindeutig zugeordnetes Graphem verschriftet; man spricht daher von einer phonemischen Schrift (Bußmann 2002<sup>3</sup>: 513). Grapheme sind die 22 (Grund-)Buchstaben a ... p, r ... v, z des lateinischen Alphabets zuzüglich sechs dieser Buchstaben mit je einem diakritischen Zeichen (Zirkumflex oder Brevis)<sup>5</sup>. In Zamenhof (1991<sup>10</sup>: 51; 57; 63; 69; 75) erfolgt die Zuordnung der Phoneme zu den Graphemen lediglich durch die Angabe vergleichbarer ethnosprachiger Phone, was in Einzelfällen die Bildung von Allophenen

5 In Zamenhof (1991<sup>10</sup>: 63) festgelegte Ersatzdarstellungen sind a) für Zirkumflex-Buchstaben (ĉ, ĝ, ĥ, ĵ, ŝ): Kombination des Grundbuchstabens und h (also ch, gh, hh, jh, sh) und b) für den Brevis-Buchstaben (ŭ): u. Bei Ersatzschreibweise liegt also keine phonemische Schrift vor. In der vorliegenden Arbeit wird (besonders außerhalb der Ausführungen zur Phonologie) zumeist die Ersatzschreibweise verwendet.

fördert(e). In Kalocsay, Waringhien (1980<sup>4</sup>: 39ff.), Wells (1978: 15ff.) und Wennergren (2006: 21ff.) wird das Phoneminventar des Esperanto mit den heute üblichen Fachausdrücken der Phonologie erläutert. Es ist in Tabelle 1 mittels der Grapheme des Esperanto zusammengefasst dargestellt.

Artikulationsart		Artikulationsort							
		bi-labial	labio-dental	den-tal	alve-olar	post-alve-olar	palatal	velar	glottal
plosiv	sth.	b			d			g	
	stl.	p			t			k	
frikativ	sth.		v		z	ĵ			
	stl.		f		s	ŝ		ĥ	h
affrika-tiv	sth.					ĝ			
	stl.				c	ĉ			
nasal		m			n				
lateral					l				
intermittierend					r				
approximantisch (halbvokalisches)	sth.						j	ŭ <sup>A)</sup>	
	stl.								
							vorne	hinten (rund)	
geschl.							i	u	
halb-offen							e	o	
offen							a		

<sup>A)</sup>Dient zur Bildung von Diphthongen wie *aŭ* und *eŭ*.

*Tabelle 1: Phoneminventar des Esperanto (erarbeitet unter Rückgriff auf Kalocsay, Waringhien 1980<sup>4</sup>: 39ff., Wells 1978: 15ff. und Wennergren 2006: 21ff. mit Artikulationsort-Bezeichnungen in Anlehnung an die von International Phonetic Association 2008 zusammengestellten)*

Wegen der phonemischen Schrift des Esperanto können Häufigkeitsuntersuchungen von Phonemen durch Häufigkeitsuntersuchungen von Buchstaben durchgeführt werden. Neuere phonemstatistische Untersuchungen liefern Dominte (2001) und Haszpra (2001). Dominte (2001) gibt nicht nur die

erhaltene Phonem-Häufigkeitsrangliste<sup>6</sup> an, sondern erklärt sie auch durch die Häufigkeit von Wörtern und Morphemen. Weiterhin kommt er zu dem Ergebnis, dass Esperanto mit einem durchschnittlichen Vokal-zu-Konsonant-Verhältnis von 43 zu 57 als relativ vokalische Sprache betrachtet werden kann (mit einem Platz zwischen den vorwiegend vokalischen Sprachen Finnisch, Türkisch, Italienisch, Neugriechisch, Serbisch, Kroatisch, Rumänisch, Litauisch, Latein, Slowakisch, Französisch und den eher konsonantischen Sprachen Ungarisch, Schwedisch, Tschechisch, Deutsch und amerikanisches Englisch). Weiterhin geht er auf das Problem der Betrachtung von drei Esperanto-Buchstaben ein, die Affrikaten repräsentieren und von Sadler (1959) und Harry (1967) als je zwei Phoneme angesehen werden:  $c = t + s$ ;  $\hat{c} = t + \hat{s}$ ;  $\hat{g} = d + \hat{j}$  (eine Betrachtungsweise, der er nicht zustimmt, was durch Wennergren 2006: 26 gestützt wird). Weiterführendes dazu findet sich in Kapitel 3.1.3.

Haszpra (2001) legt eine Phonem-Häufigkeitsrangliste vor, die er aus einem Korpus von 496.196 Graphemen aus Haszpra (1998) erhalten hat. Weiterhin resümiert er die Ergebnisse früherer Phonem- und Buchstaben-Häufigkeitsuntersuchungen wie die von Jung (1926), Stancliff (1933), Sadler (1959) und Harry (1967). Wegen der Aktualität und des genannten Korpus-Umfangs wurde in dieser Arbeit die von Haszpra (2001) vorgelegte Häufigkeitsrangliste als Grundlage für die Erprobung von Modellen (Kapitel 3.1) herangezogen. Tabelle 2 zeigt diese Phonem-Häufigkeitsrangliste.

---

6 Er nennt zwar die Quellen des untersuchten Korpus, jedoch nicht dessen Umfang.



Rang x	Phonem	Abs. Häufigkeit	Rel. Häufigkeit $f_x$ in Prozent	Rang x	Phonem	Abs. Häufigkeit	Rel. Häufigkeit $f_x$ in Prozent
1	a	62471	12,59	15	j	13447	2,71
2	e	46444	9,36	16	v	9279	1,87
3	i	44608	8,99	17	g	6351	1,28
4	o	43665	8,80	18	b	5706	1,15
5	n	38654	7,79	19	f	5409	1,09
6	l	30963	6,24	20	c	4218	0,85
7	r	29623	5,97	21	ĝ	3424	0,69
8	s	29325	5,91	22	ĉ	3325	0,67
9	t	27291	5,50	23	z	2580	0,52
10	k	20939	4,22	24	ŭ	2531	0,51
11	u	16672	3,36	25	h	2481	0,50
12	m	15184	3,06	26	ŝ	1737	0,35
13	d	15084	3,04	27	ĵ	744	0,15
14	p	13993	2,82	28	ĥ	50	0,01

*Tabelle 2: Phonem-Häufigkeitsrangliste des Esperanto (erarbeitet unter Rückgriff auf die Ergebnisse von Haszpra 2001)*

Bemerkungswert ist, dass auf den ersten vier Rängen vier der fünf Vokale liegen. Die Besetzung dieser Ränge (wie übrigens auch die der meisten anderen Ränge) ist weitestgehend in Einklang mit den Ergebnissen der Untersuchungen von Dominte (2001), Jung (1926), Stancliff (1933), Sadler (1959) und Harry (1967). Wie nach Dominte (2001) kommt man auch nach Haszpra (2001) auf ein durchschnittliches Vokal-zu-Konsonant-Verhältnis von 43 zu 57, was als ein Indiz für die Vergleichbarkeit beider Studien gewertet werden kann.

Des Weiteren ist interessant, dass in keiner der genannten Untersuchungen Ansätze für eine mathematische Modellierung der empirisch ermittelten relativen Phonemhäufigkeiten identifiziert werden konnten.

Strauss, Altmann, Best (2007) stellten die Hypothese auf, dass die geordnete Reihenfolge der Phonemhäufigkeiten einer regulären Wahrscheinlichkeitsfunktion oder regulären monoton abnehmenden Funktion

entspricht. In den Gleichungen 1 bis 7 werden entsprechende mathematische Modelle zur Beschreibung der relativen Phonemhäufigkeit<sup>7</sup>  $y$  in Abhängigkeit vom Rang  $x$  (in der nach absteigender relativer Häufigkeit geordneten Liste der Phoneme) in Anlehnung an Strauss, Altmann, Best (2007) vorgestellt. Gleichung 8 zeigt ein Modell, dessen Veröffentlichung in Vorbereitung ist.

$$\text{Tuldava (1988): } y = a + b \ln x \quad (1)$$

$$\text{Zipf (1929, 1935, 1949): } y = Ax^{-b} \quad (2)$$

$$\text{Yule (1924): } y = ax^{-b} d^x \quad (3)$$

$$\text{Naranan, Balasubrahmanyam (1992a, b, 2000): } y = Cx^{-a_1} e^{-a_2/x} \quad (4)$$

$$\text{Altmann (1993): } y = \frac{\binom{b+x}{x-1}}{\binom{a+x}{x-1}} f_1 \quad (5)$$

$$\text{Sigurd (1968): } y = f_1 q^{x-1} \quad (6)$$

$$\text{Good (1969): } y = \frac{100}{n} \sum_{i=x}^n \frac{1}{i} \quad (7)$$

$$\text{Altmann, Popescu}^8: y = ae^{-bx} + ce^{-dx} + \dots \quad (8)$$

$y$ :	mit einem Modell berechnete relative Phonemhäufigkeit (in %)
$x$ :	Rang
$f_i$ :	relative Häufigkeit des empirisch ermittelten häufigsten Phonems
$n$ :	Anzahl der Phoneme
$i$ :	Summationsindex
$a, b, c, d, a_1, a_2, A, C, q$ :	Parameter

In der Studie von Tambovtsev, Martindale (2007), in der 95 Sprachen (darunter nicht Esperanto) untersucht wurden, liefert das Yule-Modell die beste Anpassung an die Phonemhäufigkeitsverteilung .

In Kapitel 3.1 werden diese Modelle auf die Phonemhäufigkeit des Esperanto angewendet, um die o. g. Hypothese zur Phonemhäufigkeit zu testen.

<sup>7</sup> In dieser Arbeit werden die empirisch ermittelte relative Phonemhäufigkeit mit  $f_x$  und die durch ein Modell berechnete relative Phonemhäufigkeit (in %) mit  $y$  bezeichnet.

<sup>8</sup> Siehe Fußnote 2.

## 2.3 Morphologie und Sprachtypologie

### 2.3.1 Freie lexikalische und freie grammatische Morpheme

Berücksichtigt man als Wortarten, die eine lexikalische Bedeutung (d. h. einen begrifflichen Gehalt) haben können, nur Substantive, Verben, Adjektive, Adverbien, lokale und temporale Präpositionen<sup>9</sup>, gibt es im Esperanto nur (verhältnismäßig wenige) Präpositionen und Adverbien, die als freie lexikalische Morpheme auftreten können, z. B. /tuj/<sup>10</sup> (= sofort). Zu allen oben noch nicht genannten Wortarten gibt es freie grammatische Morpheme; Beispiele: die weder lokale noch temporale Präposition /kun/ (= mit); die Konjunktion /kaj/ (= und); das Pronomen /li/ (= er); der Artikel /la/ (= der, die, das); das Numerale /du/ (= zwei); die Interjektion /ha/ (= ha, ach).

### 2.3.2 Gebundene lexikalische Morpheme

Es gibt mehrere Tausend gebundene lexikalische Morpheme in Form von gebundenen Wurzeln, z. B. /patr/ (Vermittlung der Idee: Vater, väterlich). Sie treten nur in Verbindung mit mindestens einem weiteren Morphem auf, z. B. mit dem Substantiv-Morphem /o/ oder Adjektiv-Morphem /a/, und werden erst dadurch zu einem Wort: /patr/o/ (= Vater) bzw. /patr/a/ (= väterlich).

Es gibt weiterhin einige Dutzend gebundene lexikalische Morpheme in Form von autosemantischen Affixen, z. B. /on/ (das Morphem, durch das das Bruchzahl-Morphem realisiert wird). Es kann Affigierungen (Affix-Anbindungen) mit sowohl freien als auch gebundenen lexikalischen Morphemen geben. Beispielsweise kann es die Affigierung des freien lexikalischen Morphems /du/ (= zwei) mit dem gebundenen Bruchzahl-Morphem (realisiert durch /on/) und dem Adjektiv-Morphem (realisiert durch /a/) geben: /du/on/a/ (= halb). Beispiele für Affigierungen an gebundene lexikalische Morpheme finden sich in Kapitel 2.3.4.

<sup>9</sup> Lokale und temporale Präpositionen haben nach Bußmann (2002<sup>3</sup>: 402) eine lexikalische Bedeutung.

<sup>10</sup> In der vorliegenden Arbeit wird ein Morphem durch die es bildenden Schriftzeichen zwischen Schrägstrichen unter Verwendung der Ersatzschreibweise (s. Fußnote 5) dargestellt. Der Einfachheit halber wurde analog so auch bei Beispielen der deutschen Sprache verfahren (unter Verzicht auf eine spezielle Lautschrift).

### 2.3.3 Gebundene grammatische Morpheme

Folgende gebundene grammatische Morpheme werden nach Zamenhof (1991<sup>10</sup>: 63ff.) zum Ausdrücken von (im Wesentlichen) grammatischen Funktionen verwendet (es werden die entsprechenden Morpheme angegeben, durch die die jeweiligen Morpheme realisiert werden):

- (a) Substantiv-Morphem: Menge der Allomorphe {/o/, /' /}, z. B. /patr/o/ oder /patr/'/ (= Vater); freie Variation<sup>11</sup>
- (b) Artikel-Morphem<sup>12</sup>: Menge der Allomorphe {/la/, /l' /}, freie Variation nur nach vokalisch endenden Präpositionen<sup>13</sup> (z. B. /de/ /l' / [= von dem, von der]), sonst normalerweise /la/ obligatorisch für „guten Stil“
- (c) Adjektiv/Ordinalzahl-Morphem: /a/, z. B. /san/a/ (= gesund [Adjektiv]), /unu/a/ (= erster [Ordinalzahl])
- (d) Adverb-Morphem: Menge der Allomorphe {/e/, Nullmorph} in komplementärer Distribution: Nullmorph nur bei den in Kapitel 2.3.1 erwähnten freien lexikalischen (Adverb-)Morphemen möglich und üblich, z. B. /tuj/Nullmorph/ (= sofort), aber obligatorisches /e/ bei /san/e/ (= gesund [Adverb]).
- (e) Plural-Morphem: /j/, z. B. patr/o/j/ (= Väter), /san/a/j/ (= gesunde)
- (f) Nominativ-Morphem: Nullmorph, z. B. /patr/o/Nullmorph/ (= Vater [Nominativ])
- (g) Akkusativ-Morphem<sup>14</sup>: /n/, z. B. /patr/o/n/ (= Vater [Akkusativ])

11 Das ist durch die in Zamenhof (1991<sup>10</sup>: 67) gegebene Regel 16 gerechtfertigt: „Die Endung des Substantivs und des Artikels kann ausgelassen werden, indem man dieselbe durch einen Apostroph ersetzt ...“ Weiterhin: Bei aus fremden Sprachen ins Esperanto übernommenen Eigennamen darf das Substantiv-Morphem auch durch das Nullmorph realisiert werden, z. B. /zamenhof/Nullmorph/ (Zamenhof 1991<sup>10</sup>: 49), /vashington/Nullmorph/ (Zamenhof 1991<sup>10</sup>: 92).

12 Im Esperanto gibt es nur den bestimmten Artikel.

13 Vergleiche Fußnote 11. Auf die optionale Verwendung von /l' / (statt /la/) nur nach vokalisch endenden Präpositionen wird in Zamenhof (1991<sup>10</sup>: 113) hingewiesen. Da es jedoch die diesbezüglich weniger restriktive Anweisung im Grammatik-Teil gibt, wird eine Verwendung von /l' / nach nicht vokalisch endenden Präpositionen in Zamenhof (1925: 47) nur als „Sünde gegen die Güte des Stils“ gesehen.

14 Das Akkusativ-Morphem ist außerdem zur Angabe einer Bewegung in Richtung auf ein Ziel obligatorisch (als „Richtungsakkusativ“), wenn das durch andere Morpheme (wie beispielsweise Präpositionen, die bereits lexikalisch eine Richtung vermitteln) nicht deutlich wird; vgl. Zamenhof (1991<sup>10</sup>: 55) und Zamenhof (1991<sup>10</sup>: 114).

- (h) Morphem aller Kasus, die weder Nominativ noch Akkusativ sind:  
Nullmorph plus Verwendung von Präpositionen, z. B. /de/ /la/  
/patr/o/Nullmorph/ (= des Vaters), /a/ /la/ /patr/o/Nullmorph/ (= dem  
Vater)
- (i) Infinitiv-Morphem: /i/, z. B. /far/i/ (= machen)
- (j) Imperativ/Optativ-Morphem: /u/, z. B. /far/u/ (= mach, du mögest  
machen, macht, ihr möget machen)
- (k) Konditional-Morphem: /us/, z. B. /far/us/ (= würde machen, würdest  
machen usw.)
- (l) Indikativ-Präsens-Aktiv-Morphem: /as/, z. B. /far/as/ (= mache, machst  
usw.)
- (m) Indikativ-Imperfekt/Präteritum-Aktiv-Morphem: /is/, z. B. /far/is/  
(= machte, machtest usw.)
- (n) Indikativ-Futur-Aktiv-Morphem: /os/, z. B. /far/os/ (= werde machen, wirst  
machen usw.)
- (o) Partizip-Präsens-Aktiv-Morphem: /ant/, z. B. /far/ant/a/ (= machend)
- (p) Partizip-Perfekt-Aktiv-Morphem: /int/, z. B. /far/int/a/ (= gemacht habend)
- (q) Partizip-Futur-Aktiv-Morphem: /ont/, z. B. /far/ont/a/ (= sich anschickend  
zu machen)
- (r) Partizip-Präsens-Passiv-Morphem: /at/, z. B. /far/at/a/ (= gemacht  
[werdend])
- (s) Partizip-Perfekt-Passiv-Morphem: /it/, z. B. /far/it/a/ (= gemacht [worden])
- (t) Partizip-Futur-Passiv-Morphem: /ot/, z. B. /far/ot/a/ (= [davor stehend]  
gemacht [zu werden])

### 2.3.4 Wortbildung durch Derivation

Die in Kapitel 2.3.2 erwähnten autosemantischen Affixe werden zur Wortbildung durch Derivation auch an gebundene lexikalische Morpheme affigiert. So verwendet man z. B. das durch /mal/ realisierte „Gegensatz zu“-Morphem, um aus /san/a/ (= gesund) /mal/san/a/ (= krank) zu bilden. Weiter werden

beispielsweise gebildet: /mal/san/ul/o/ (= Kranker) mit dem durch /ul/ realisierten „Person mit Eigenschaft“-Morphem und /mal/san/ul/ej/o/ (= Krankenhaus) mit dem durch /ej/ realisierten „Ort für“-Morphem.

### 2.3.5 Wortbildung durch Komposition

Wortbildung durch Verkettung von ausschließlich freien lexikalischen Morphemen ist im Esperanto selten. Sie tritt z. B. bei Numeralien wie /du/dek/ (= zwanzig) auf.

Zumeist werden gebundene grammatische und gebundene lexikalische Morpheme verkettet, z. B. /vapor/ship/o/ (= Dampfschiff).

Dabei sind gebundene grammatische Morpheme des in der linearen Reihenfolge nicht am weitesten rechts stehenden gebundenen lexikalischen Morphems fakultativ (und werden besonders bei Konsonantenhäufung verwendet, um Aussprache oder Verständnis zu erleichtern), z. B. bei /angl/a/lingv/a/ (= englischsprachig; gebildet aus [/angl/a/ + /lingv/o/] + /a/), das in freier Variation zu /angl/lingv/a/ verwendet wird.

Gebundene grammatische Morpheme des in der linearen Reihenfolge nicht am weitesten rechts stehenden gebundenen lexikalischen Morphems sind allerdings obligatorisch, wenn sich ohne sie eine andere Bedeutung ergeben würde, z. B. ist /unu/a/tag/a/ (= ersttägig) zum Unterschied von /unu/tag/a/ (= eintägig) zu verwenden (vgl. Zamenhof 1991<sup>10</sup>: 113).

Wortbildung erfolgt auch unter Zuhilfenahme von Präpositionen, z. B. /kongres/o/ /de/ /esperant/ist/o/j/ (= Esperantistenkongress; Zamenhof 1991<sup>10</sup>: 49).

### 2.3.6 Suppletivismus, unikale, diskontinuierliche und Portmanteau-Morpheme

Diskontinuierliche Morpheme treten im Esperanto vereinzelt auf, z. B. /nek/ ... /nek/ (= weder ... noch). Die anderen o. g. Phänomene gibt es im Esperanto nicht.

### 2.3.7 Spezielle Arten der Komposition

Im Esperanto gibt es (vereinzelt):

- (a) Akronymie, z. B. UEA (Universala Esperanto-Asocio = Esperanto-Weltbund)
- (b) Clipping, z. B. /aut/o/ als Kurzform von /automobil/o/ (= Automobil)
- (c) Reduplikation, z. B. /plen/plen/a/ (= überfüllt)
- (d) Konversion, z. B. /amend/o/ (= Ergänzen [als substantivierte Tätigkeit] oder Ergänzung [als Ergebnis des Ergänzens])

Dem Autor ist nicht bekannt, dass Blending im Esperanto vorkommt.

### 2.3.8 Grundlegende Wortart des lexikalischen Morphems

Im Esperanto ist die grundlegende Wortart des lexikalischen Morphems nur aus seiner Definition oder seinem etwaigen (traditionellen) Gebrauch erkennbar (z. B. durch Zamenhof 1991<sup>10</sup>). So ist beispielsweise die grundlegende Wortart von /komb/ ein Verb (kämmen), wohingegen die grundlegende Wortart von /bros/ ein Substantiv ist (Bürste). Ohne diese Kenntnis weiß man nicht, dass Kamm nicht /komb/o/, sondern /komb/il/o/ (mit dem durch /il/ realisierten „Werkzeug für“-Morphem), und Bürste nicht /bros/il/o/, sondern /bros/o/ heißt.

### 2.3.9 Korrelativwörter

Eine erwähnenswerte Besonderheit des Esperanto sind seine „korelativaj vortoj“ (= Korrelativwörter): In Zamenhof (1991<sup>10</sup>: 116-117) sind 45 spezielle Wörter definiert, die unter sich in einer gewissen Wechselbeziehung stehen (daher ihre Bezeichnung). Dies wird durch die Realisierung ihrer Morpheme durch spezielle (leicht einprägsame und regelmäßig verwendete) Morpheme deutlich (s. Tabelle 3). Ihnen allen gemeinsam ist das Morph /i/, durch das das Korrelativitäts-Morphem<sup>15</sup> realisiert wird.

Beispielhaft wird das Korrelativwort /t/i/u/ (= jener) erläutert. Es besteht (wie die anderen 44 Korrelativwörter) aus drei Morphemen, von denen jedes durch ein Morph realisiert wird:

- (a) Das Morph /i/ realisiert das Korrelativitäts-Morphem, d. h. mit /i/ wird die Zugehörigkeit zu den 45 Korrelativwörtern ausgedrückt.

<sup>15</sup> Die in Tabelle 3 aufgeführten Morpheme und deren Realisierungen durch Morpheme sind (streng genommen) auf das System der 45 Korrelativwörter beschränkt.

- (b) Das Suffix /u/ realisiert das Identitäts-Morphem, d. h. es wird ausgedrückt, dass es um die Auswahl aus mehreren Personen/Sachen geht (/u/ allein drückt aber noch nicht aus, ob es sich bei der Auswahl um eine bestimmte Person/Sache, eine unbestimmte Person/Sache, keine Person/Sache, eine hinterfragte Person/Sache oder um alle in Frage kommenden Personen/Sachen handelt; dies wird durch das Präfix vermittelt, siehe (c)).
- (c) Das Präfix /t/ realisiert das Bestimmtheits-Morphem, d. h. es wird ausgedrückt, dass eine bestimmte Person/Sache bezeichnet wird.



Morphem der/des ...	Unbestimmtheit, realisiert durch Nullmorph als Präfix	Bestimmt- heit, realisiert durch Präfix /t/	Frage, realisiert durch Präfix /k/	Allumfasst- heit, realisiert durch Präfix / ch/	Vernei- nung, realisiert durch Präfix /nen/
Eigenschaft, realisiert durch Suffix /a/	/Nullmorph/i/a/ irgendeine Art von	/t/i/a/ solcher, derartiger	/k/i/a/ attributiv: was für einer; prädikativ: wie	/ch/i/a/ jeglicher	/nen/i/a/ keinerlei
Klassifi- zierung, realisiert durch Suffix /o/	/Nullmorph/i/o/ (irgend)etwas	/t/i/o/ jenes, das (da)	/k/i/o/ was	/ch/i/o/ alles	/nen/i/o/ nichts
Identität, realisiert durch Suffix /u/	/Nullmorph/i/u/ (irgend)einer, (irgend)jemand	/t/i/u/ jener, der (da)	/k/i/u/ welcher, wer	/ch/i/u/ jeder	/nen/i/u/ keiner, niemand
Zugehörig- keit, realisiert durch Suffix /es/	/Nullmorph/i/es/ (irgend)jemandes	/t/i/es/ dessen	/k/i/es/ wessen	/ch/i/es/ jedermanns	/nen/i/es/ niemandes
Ort, realisiert durch Suffix /e/	/Nullmorph/i/e/ irgendwo	/t/i/e/ dort	/k/i/e/ wo	/ch/i/e/ überall	/nen/i/e/ nirgendwo
Zeit, realisiert durch Suffix /am/	/Nullmorph/i/am/ irgendwann	/t/i/am/ dann	/k/i/am/ wann	/ch/i/am/ immer	/nen/i/am/ nie
Art u. Weise / Grad, realisiert durch Suffix /el/	/Nullmorph/i/el/ irgendwie	/t/i/el/ so	/k/i/el/ wie	/ch/i/el/ auf jede Weise	/nen/i/el/ keineswegs
Grund, realisiert durch Suffix /al/	/Nullmorph/i/al/ aus irgendeinem Grund	/t/i/al/ darum	/k/i/al/ warum	/ch/i/al/ aus jedem Grund	/nen/i/al/ aus keinem Grund
Menge, realisiert durch Suffix /om/	/Nullmorph/i/om/ etwas	/t/i/om/ so viel	/k/i/om/ wie viel	/ch/i/om/ alles	/nen/i/om/ nichts

*Tabelle 3: Korrelativwörter des Esperanto; das ihnen allen gemeinsame Korrelativitäts-Morphem wird durch das Morph /i/ realisiert*

### 2.3.10 Häufigkeitsverteilung von Morphemstrukturen

Blahuš (2008: 24) listet die 20 häufigsten Morphemstrukturen des Esperanto (Korpus: 33000 Morphemstrukturen in: Waringhien 2005); s. Tabelle 4. Er weist auf das Zipfsche Gesetz hin, macht jedoch keine detaillierte Analyse. Diese findet sich in Kapitel 3.3.

Frequency	Structure
12.970	*
6.005	*/*
1.386	*/o/*
800	*/aĵ
795	*/ad
642	*/ist
631	*/ig
619	*/ec
505	*/iĝ
417	*/il
400	*/ul
376	*/ej
359	*/et
318	*/ism
317	*/*/ig
311	mal/
254	*/*/iĝ
220	*/uj
215	*/ar
206	*/*/il

*Tabelle 4: Die 20 häufigsten Morphemstrukturen im Esperanto; Stern (\*): lexikalisches Morphem, das kein autosemantisches Affix ist (Blahuš 2008: 24)*

### 2.3.11 Aussagen zur sprachtypologische Einordnung

In Sikosek (2003<sup>2</sup>: 204) wird resümiert, dass nach Meinung der meisten Autoren<sup>16</sup> und in passendster Verallgemeinerung Esperanto agglutinierend ist und einige Merkmale einer isolierenden Sprache aufweist.

In Wendt (1977: 356) wird Esperanto als analytisch agglutinierend charakterisiert (wobei agglutinierend, isolierend, polysynthetisch und flektierend als die vier Haupttypen genannt werden).

Wells (1978: 33) errechnete den Index der Agglutination A/J zu 1,00 und bezeichnet Esperanto als maximal agglutinierend<sup>17</sup>.

<sup>16</sup> Autoren: Bausani (1970: 126), Blanke (1986<sup>2</sup>: 31), Piron (1986: 31-32), Stamatiadis (1986) und Wells (1989<sup>2</sup>: 33).

<sup>17</sup> Wells (1978: 33) weist darauf hin, dass ein A/J < 1,00 sich dann hätte ergeben können,

Quantitative Methoden zur morphologischen Sprachtypologie finden sich in Greenberg (1960), Altmann, Lehfeldt (1973), Kempgen, Lehfeldt (2004) und Hoffmann (2005). Näheres dazu wird in Kapitel 3.2 ausgeführt.

## 2.4 Syntax

### 2.4.1 Einführendes zur Syntax

An komplexeren und systematischeren Untersuchungen zur Syntax der Plansprachen besteht laut Sakaguchi (1998: 191) in der Interlinguistik „ein deutlicher Mangel“. Sie verweist (bezüglich Esperanto) auf Kalocsay, Waringhien (1980<sup>4</sup>) und Wells (1978) als einige wenige, die zu diesem Thema veröffentlichten. Als neuere Arbeit zur Syntax des Esperanto können Gledhill (2000<sup>2</sup>) und Jansen (2007), Letztere mit Schwerpunkt Wortstellung, genannt werden. Da es sich bei Esperanto um eine Plansprache mit präskriptiver Grundlage handelt (vgl. Zamenhof 1991<sup>10</sup>: 45), ist einerseits im Gegensatz zu natürlichen Sprachen die Akzeptabilität bzw. Grammatikalität von Sätzen normalerweise nicht über die Kompetenz eines Muttersprachlers im Sinne Chomskys beurteilbar. Andererseits ist nicht jede grammatische oder syntaktische Regel ausdrücklich formuliert, sondern nur mittelbar aus Texten der Grundlage zu erschließen (vgl. Wennergren 2006: 633). So heißt es beispielsweise in Waringhien, Kalocsay (1980<sup>4</sup>: 362) einleitend im Kapitel zur Wortstellung, dass im Esperanto die Wortstellung frei sei, es jedoch seltene Fälle gebe, in denen man sich an eine feste Wortstellung halten solle (was dann in dem betreffenden Kapitel weiter ausgeführt wird). Ähnliche Aussagen finden sich in Wells (1978: 41) und Gledhill (2000<sup>2</sup>: 87).

### 2.4.2 Struktur eines Esperanto-Satzes

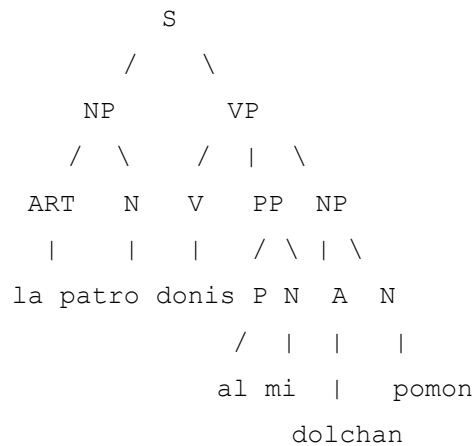
Die Struktur eines beliebig gewählten einfachen deklarativen Esperanto-Satzes wird durch ein Beispiel in einem Phrasenstrukturdiagramm dargestellt (Abbildung 1).

---

wenn sich in der Esperanto-Textprobe Koseform-Morpheme befunden hätten, die durch /chj/ für männliche und /nj/ für weibliche Personen realisiert werden: Diese Morpheme werden an lexikalische Morpheme angefügt, die dabei auf einen Buchstaben bis zu sechs Buchstaben verkürzt werden, womit zwar eine Morph-Junktur, jedoch keine agglutinierende Konstruktion vorliegt, z. B. /vilhelm/o/ (= Wilhelm), /vi/chj/o/ (= Willi) oder auch /vil/chj/o/ (= Willi); vgl. Zamenhof (1991<sup>10</sup>: 129-130).

*la patro donis al mi dolchan pomon* (Zamenhof 1991<sup>10</sup>: 88)

der Vater gab mir [wörtl.: zu ich] einen süßen Apfel



*Abbildung 1: Herkömmliches Phrasenstrukturdiagramm eines Esperanto-Satzes*

Es gelten also folgende Ersetzungsregeln:

S -> NP VP

NP -> ART N

ART -> la

N -> patro

VP -> V PP NP

V -> donis

PP -> P N

P -> al

N -> mi

NP -> A N

A -> dolchan

N -> pomon

Wegen bestimmter Nachteile der konventionellen Phrasenstrukturregeln (Möglichkeit unsinniger Strukturen; Nicht-Formulierbarkeit bestimmter

linguistisch plausibler Analysen; Redundanz) wurde die X-Bar-Theorie entwickelt (vorgeschlagen in Chomsky 1970 und weiterentwickelt in Jackendoff 1977 zur Identifizierung syntaktischer Merkmale, die [nach einem Postulat] allen Sprachen gemeinsam sind). Abbildung 2 zeigt den Beispielsatz von Abbildung 1 in X-Bar-Theorie. Damit sollen in der vorliegenden Arbeit auch weitere Analysen erfolgen (in Kapitel 3.4). Motiv: In der X-Bar-Theorie gibt es, ausgehend von einem Nicht-Terminal (NP, VP usw.), höchstens zwei Verzweigungen. Daher sind Links- und Rechtseinbettungen sofort erkennbar.

*la patro donis al mi dolchan pomon* (Zamenhof 1991<sup>10</sup>: 88)

der Vater gab mir [wörtl.: zu ich] einen süßen Apfel

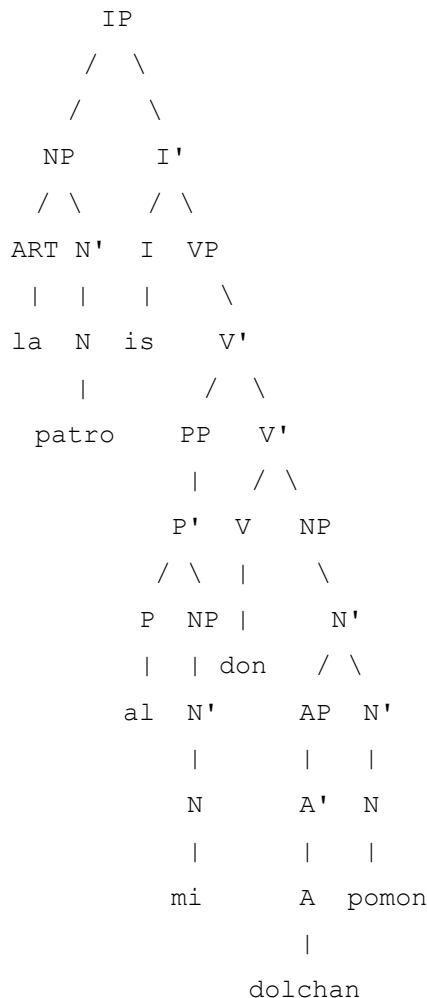


Abbildung 2: Der Esperanto-Satz von Abbildung 1 in X-Bar-Konvention

Abbildung 3 zeigt eine typische Gesamtgrundstruktur einer Phrase XP in X-Bar-Theorie<sup>18</sup>.

<sup>18</sup> In dieser Arbeit werden nur drei Komplexitätsebenen verwendet (XP; X'; X). Laut

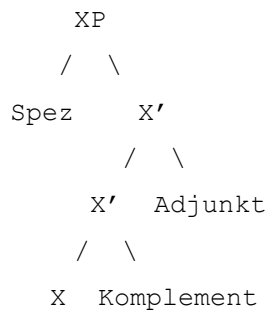


Abbildung 3: Gesamtgrundstruktur einer Phrase

Abbildung 4 zeigt beispielhaft eine Nominalphrase (hier ohne Adjunkte und ohne Komplemente).

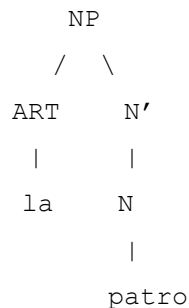


Abbildung 4: Nominalphrase (Beispiel)

Der Spezifizierer ist hier der (bestimmte) Artikel *la*. Obgleich es keine ausdrückliche Regel dazu gibt, ist laut Wells (1978: 47) die Linkseinbettung des Artikels im Esperanto fixiert. Eine Rechtseinbettung (bezüglich des spezifizierten Nomens) ist laut Wells (1978: 47) zwischen Personennamen und Ordnungszahl oder Adjektiv möglich (z. B. „*Luizo la Dekkvara* [Ludwig der Vierzehnte]“).

In Kapitel 3.4 werden die Rechts- und Linkseinbettung von verschiedenartigen Phrasen und Relativsätzen im Esperanto in X-Bar-Theorie dargestellt und quantitativ analysiert.

## 2.5 Lexikon des Esperanto

Man kann das Lexikon des Esperanto in drei Kategorien von Wörtern einteilen:

- (a) „Grundlegende Wörter“: Lexeme, die in Zamenhof (1991<sup>10</sup>) erscheinen;  
Anzahl: 2769 (nach *Akademia Vortaro*, Wennergren 2007b). Diese

---

Bußmann (2002<sup>3</sup>: 759) ist in der Literatur umstritten, welchen numerischen Index die höchste Komplexitätsebene definiert.

Lexeme wurden von Zamenhof gezielt aus Lexemen germanischer und romanischer Sprachen abgeleitet (Zamenhof 1992<sup>18</sup>: 284).

- (b) „Offizielle Wörter“: Lexeme, die durch eine Entscheidung der Esperanto-Akademie (oder deren Vorläufer, des Sprachausschusses) dem „Universala Vortaro“, das ist ein Teil von Zamenhof (1991<sup>10</sup>), hinzugefügt wurden; Anzahl: 2168 (nach *Akademio Vortaro*, Wennergren 2007b).
- (c) „Neue Wörter“: Hierzu steht im (institutionalisierenden) Vorwort zu Zamenhof (1991<sup>10</sup>): „Die Sprache durch neue Wörter *bereichern* kann man bereits jetzt, indem man sich mit jenen Personen beratschlägt, die als die kompetentesten in unserer Sprache angesehen werden, und indem man dafür sorgt, dass alle jene Wörter in derselben Form verwenden, aber diese Wörter dürfen nur empfohlen werden, nicht aufgedrängt ...“ Die Anzahl der in diesem Sinne neuen Esperanto-Wörter liegt vermutlich bei etlichen Tausenden, wenn man renommierte Esperanto-Wörterbücher zugrunde legt wie z. B. Krause (1999) und Waringhien (2002). Internationalismen im Sinne von Regel 15 der Esperanto-Grammatik (Zamenhof 1991<sup>10</sup>: 67) grenzt Zamenhof (1925: 24) jedoch von den neuen Wörtern ab: Erstere gehören zu Esperanto „schon seit dem ersten Augenblick seiner Entstehung“. (Neu wäre demnach lediglich ihre Ingebrauchnahme.)

Im Vorwort zu Zamenhof (1991<sup>10</sup>) ist weiterhin beschrieben, unter welchen Bedingungen und auf welche Weise neue Lexeme zu offiziellen Wörtern werden können: „Erst irgendwann später, wenn der größte Teil der neuen Wörter bereits gänzlich ausgereift sein wird, wird sie eine Art autorisierte Institution in das offizielle Wörterbuch als Beifügung zum ‚Fundamento‘ einführen“. Bis jetzt erfolgten derartige Offizialisierungen mehrmals von 1909 bis 2007, zumeist in Form von „Oficialaj Aldonoj“, also „offiziellen Beifügungen“: Cart (1909); Cart (1919); Grosjean-Maupin (1922); Grosjean-Maupin (1929); Bastien, Rollet de l'Isle (1934); Bailey, Rollet de l'Isle (1935); Akademio de Esperanto (Hrsg.) (1958); Akademio de Esperanto (Hrsg.) (2007<sup>2</sup>); Wennergren (2007a). Es gibt fast 5000 Wörter, die „grundlegend“ oder „offiziell“ sind (Übersicht: Tabelle 5); weitere Einzelheiten sind in Kück (2008) beschrieben.

Offizialisierungs-Dokument o. offizial. Wurzel-Morphem	Jahr der Offizialisierung	Wortzahl	Definitionen <sup>A</sup> in Esperanto	Anzahl der Übersetzungen in die Sprachen										
				Frz.	Engl.	Dt.	Poln.	Russ.	Span.	Ital.	Port.	Katalan.	La.	
Fundamento	1905	2769 <sup>B</sup>		2628	2628	2628	2628	2628						
1. Beifügung	1909	806 <sup>C</sup> (805) <sup>D</sup>	1	805 <sup>C</sup> (804) <sup>D</sup>	805 <sup>C</sup> (804) <sup>D</sup>	805 <sup>C</sup> (804) <sup>D</sup>	805 <sup>C</sup> (804) <sup>D</sup>	805 <sup>C</sup> (804) <sup>D</sup>						
2. Beifügung	1919	587 <sup>E</sup>	3	110 <sup>F</sup>	110 <sup>F</sup>	110 <sup>F</sup>								
3. Beifügung	1921 <sup>G</sup>	203 <sup>H</sup>	3	203 <sup>H</sup>	203 <sup>H</sup>	203 <sup>H</sup>	134		203 <sup>H</sup>	203 <sup>H</sup>	189 <sup>H</sup>			
„analog“	1923	1	?	?	?	?	?	?	?	?	?	?	?	?
„mis“	1929	1	?	?	?	?	?	?	?	?	?	?	?	?
4. Beifügung	1929	118	23	118	118	118			118	118	118	118		
5. Beifügung	1934	8		8	8	8	8		8	8				
6. Beifügung	1935	21		21	21	21	21		21	21				
„end“	1953	1	?	?	?	?	?	?	?	?	?	?	?	?
7. Beifügung	1958	31	16 <sup>I</sup>											
8. Beifügung	1974 <sup>J</sup>	200	154	154	154	154			154	154	154			6
9. Beifügung	2007	209	209											23
Summe		4955 (4954) <sup>K</sup>	409 <sup>I</sup>	4047 (4046) <sup>K</sup>	4047 (4046) <sup>K</sup>	4047 (4046) <sup>K</sup>	3596 (3595) <sup>K</sup>	3433 (3432) <sup>K</sup>	504	504	461	118		29

<sup>A</sup> Als Definitionen werden hierbei solche Beschreibungen angesehen, die auch im *Akademia Vortaro* (Wennergren 2007b) wiedergegeben sind (darin als „oficialaj difinoj“ bezeichnet) zuzüglich der Definitionen von *ISM'* und der Wörter der fünften und sechsten Beifügung. (Die Definitionen dieser Wörter fehlten im *Akademia Vortaro* [Wennergren 2007b] trotz des Hinweises von Kück 2007.)

<sup>B</sup> einschließlich 11 gebundener grammatischer Morphe, weil sie in den Aufstellungen des *Akademia Vortaro* (Wennergren 2007b) und *Akademio de Esperanto* (Hrsg.) (1958) ebenfalls mitgezählt werden

<sup>C</sup> zuzüglich 55 „Fundamento“-Wörter

<sup>D</sup> Wenn man *juristo* als „Fundamento“-Wort betrachtet, gilt die Zahl in Klammern (zzgl. 56 „Fundamento“-Wörter).

<sup>E</sup> zuzüglich 4 „Fundamento“-Wörter

<sup>F</sup> zuzüglich eines „Fundamento“-Wortes

<sup>G</sup> Nach dem *Akademia Vortaro* (Wennergren 2007b); das vom Autor der vorliegenden Arbeit gefundene Dokument (Grosjean-Maupin 1922) ist von 1922.

<sup>H</sup> zuzüglich eines „Fundamento“-Wortes

<sup>I</sup> zuzüglich der Definitionen von 81 „Fundamento“-Wörtern

<sup>J</sup> nach dem *Akademia Vortaro* (Wennergren 2007b)

<sup>K</sup> Wenn man *juristo* als Fundamento-Wort betrachtet, gilt die Zahl in Klammern.

?: nicht bekannt

Tabelle 5: „Fundamento“- und offizielle Esperanto-Wörter; Anzahl (Übersicht)

Abbildung 5 zeigt die Zahl  $n$  der grundlegenden und offiziellen Wörter gegen  $i$ ;  $i$  ist der Rang in der Offizialisierungsfolge, d. h. die Nummer der Beifügung plus 1, wobei Zamenhof (1991<sup>10</sup>) als „nullte Beifügung“ betrachtet wird, also  $i = 1$ . Nach Kück (2008: 19ff) ist die Zahl  $n$  modellierbar durch die (einparametrische Zipf-)Modellfunktion  $n(i) = n(1)/i^{1,706}$  mit  $n(1) = 2769$ . Im Kapitel 3.5 wird die Modellierbarkeit mit einer anderen Funktion untersucht.



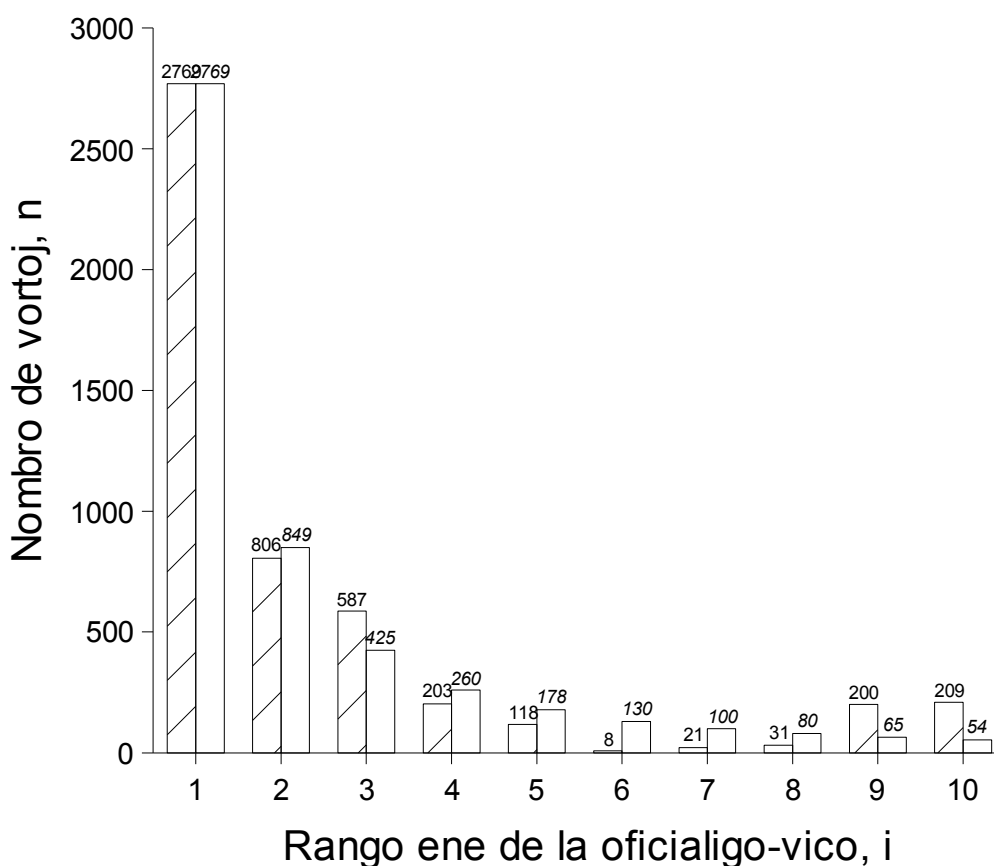


Abbildung 5: Anzahl der Wörter gegen den Rang in der Offizialisierungsfolge; schraffiert: tatsächliche Zahlen; ungeschraffiert: nach Modellfunktion (aus: Kück 2008: 19)

## 2.6 Lexikalische Semantik

### 2.6.1 Einführendes zur lexikalischen Semantik des Esperanto

Nach Sakaguchi (1998: 198) gehört die Semantik zu den am wenigsten erforschten Bereichen der Interlinguistik. Eine Besonderheit des Esperanto besteht darin, dass die Zuordnung von Bedeutungen zu den einzelnen Lexemen, die weitestgehend aposteriorisch<sup>19</sup> aus Lexemen von Ethnosprachen abgeleitet wurden, zumindest am Anfang zunächst präskriptiv erfolgte – und zwar durch die Grundlagerschrift von Zamenhof (1991<sup>10</sup>). Darin erfolgt die

<sup>19</sup> In der Interlinguistik unterscheidet man zwischen aposteriorischen und apriorischen Plansprachen und -projekten: Erstere wurden/werden auf der Grundlage von ethnischen Sprachen konstruiert, Letztere haben mit solchen keine Ähnlichkeit (vgl. Sikosek 2003<sup>2</sup>: 262). Zu Überlegungen Zamenhofs, eine apriorische Plansprache zu entwickeln, siehe Zamenhof (1925: 86).

Bedeutungszuordnung durch die Angabe von ethnosprachigen Äquivalenten<sup>20</sup> oder durch Esperanto-sprachige Paraphrasen. Durch die Sprecherschaft wurden und werden in der Folgezeit die einzelnen Lexeme entweder akzeptiert oder nicht, wobei im erstgenannten Fall eine Bedeutungsverfestigung oder eine Bedeutungsmodifikation eintreten kann. Unter Bezugnahme auf Duc Goninaz (1987a) und Duc Goninaz (1987b) verweist Sakaguchi (1998: 199) darauf, dass die Bedeutung der Lexeme stark durch die entsprechende Bedeutung in den europäischen Ethnosprachen geprägt wurde. Eine der einseitigen Übernahme europäischer Bedeutungen entgegenwirkende Entwicklung stellte sich nach Sakaguchi (1998: 199) im Zuge der Ausbreitung der Esperanto-Sprechergemeinschaft auf nichteuropäische Länder (Iran, China, Japan) ein.

Im Folgenden werden charakteristische Merkmale der sinnrelationalen Semantik (Kapitel 2.6.2 bis 2.6.5) und der Wortfeldsemantik (Kapitel 2.6.6 bis 2.6.8) des Esperanto dargestellt.

### 2.6.2 Polysemie

Polysemie lässt sich auch in Esperanto finden (s. Beispiele 2.6.2-1 und 2.6.2-2):

Beispiel 2.6.2-1: *plumo*

1. (Vogel-)Feder
2. (Schreib-)Feder

Beispiel 2.6.2-2: *de*

1. von (Possessiv)
2. von (Herkunft)
3. von, durch (Agens im Passiv)
4. an, aus, durch (Grund, Ursache)
5. seit

(und nach Waringhien 2002: 218 noch weitere 8 Bedeutungen)

In Sakaguchi (1998: 202) wird die Polysemie von *de* als „ein ernsthafter Konstruktionsmangel“ angesehen und darauf hingewiesen, dass in Plansprachen im Allgemeinen angestrebt wird, Polysemie (und auch Homonymie) zu vermeiden.

---

<sup>20</sup> In Französisch, Englisch, Deutsch, Russisch, Polnisch.

Für einige Sprachen wurden gesetzmäßige (quantitative) Zusammenhänge zwischen der Länge von Lexemen (ausgedrückt durch die Zahl der Silben) und Polysemie (ausgedrückt durch die Anzahl der Bedeutungen) gefunden, beispielsweise für Maori in Köhler (1999). (Grundlegendes zu der Gesetzmäßigkeit zwischen Lexem-Länge und Polysemie findet sich in Altmann, Beóthy, Best 1982 sowie in Köhler 1990: 3ff.) Da diesbezügliche Untersuchungen für Esperanto vom Autor der vorliegenden Arbeit nicht ausfindig gemacht werden konnten, wird in Kapitel 3.6 eine dahingehende Untersuchung erläutert.

### 2.6.3 Homonymie und Paronymie

Im Esperanto gibt es Beispiele für echte und morphologische Homonymie (Beispiele 2.6.3-1 bzw. 2.6.3-2):

Echte Homonymie:

- Beispiel 2.6.3-1: *boro*
1. Bohrung
  2. Bor (chemisches Element)

Morphologische Homonymie (Isomerie):

- Beispiel 2.6.3-2: *etago*
1. Stockwerk
  2. Kleinkind-Alter (als Agglutination aus  
*/et/agh/o/* mit: */et/*: Diminutiv-Morphem;  
*/agh/*: Wurzel-Morphem für „Alter“;  
*/o/*: Substantiv-Morphem)

Nach Sakaguchi (1998: 202) nimmt das Phänomen der Homonymie im Esperanto mit der Zeit zu. Sie nennt keine Gründe dafür. Eine naheliegende Erklärung könnte wie folgt formuliert werden: Mit der Zeit wurden einige Lexeme ins Esperanto eingeführt, weil sie nach Ansicht der Esperanto-Sprecher „internationale Wörter“ und damit „schon esperantisch“ (Cart 1919: 4) sind, d. h. ihre in mehreren Ethnosprachen äquivalenten Lexeme sind unter sich ähnlich und bilden die Grundlage für das äquivalente Esperanto-Lexem, selbst wenn damit eine Isomerie entsteht (Beispiel 2.6.3-3):

- Beispiel 2.6.3-3: *nomado*
1. Nomade („offiziell“ als „internationales Wort“ eingeführt, das gemäß Regel 15

der Esperanto-Grammatik

[Zamenhof 1991<sup>10</sup>: 67] „esperantisch“

ist [Cart 1919: 5ff.]

2. Nennen (als Agglutination aus

*/nom/ad/o/* mit folgenden

„grundlegenden“ Morphemen:

*/nom/*: Wurzel-Morphem für „nennen“;

*/ad/*: Morphem zum Ausdruck der

allgemeinen und abstrakten Idee der durch

das Wurzel-Morphem vermittelten

Tätigkeit; */o/*: Substantiv-Morphem)

In anderen Fällen dagegen wird im Esperanto offensichtlich ein Paronym verwendet, um ein drohendes Homonym gezielt zu vermeiden<sup>21</sup> (Beispiel 2.6.3-4 ist aus Wells 1978: 64):

Beispiel 2.6.3-4:	Stück	1. <i>ero</i>
	Ära	2. <i>erao</i>

#### 2.6.4 Synonymie

Laut Wells (1978: 62) sind Beispiele für Synonyme im Esperanto rivalisierende

(a) Bezeichnungen für geographische Begriffe oder Menschen

Beispiel 2.6.4-1: Kreta	1. <i>Kandio</i>
-------------------------	------------------

2. <i>Kreto</i>
-----------------

Beispiel 2.6.4-2: Odysseus	1. <i>Odiseo</i>
----------------------------	------------------

2. <i>Uliso</i>
-----------------

(b) Lexeme, von denen sich noch keines im Sprachgebrauch als vorherrschend durchgesetzt hat

Beispiel 2.6.4-3: Rasen	1. <i>gazono</i>
-------------------------	------------------

2. <i>razeno</i>
------------------

Beispiel 2.6.4-4: Strand	1. <i>plagho</i>
--------------------------	------------------

2. <i>strando</i>
-------------------

Laut Wells (1978: 62) kann es in einigen Fällen zu einer beginnenden

<sup>21</sup> Vorbilder für Paronym-Bildung zur Homonym-Vermeidung liefert bereits Zamenhof (1991<sup>10</sup>). Beispielsweise ist das Esperanto-Wort für „Planet“ nicht etwa *planeto* (das wäre auch als „Plänchen“ lesbar), sondern *planedo*.

Bedeutungsmodifikation kommen: *plagho* beginnt, die Bedeutung eines organisierten Badeortes für Feriengäste anzunehmen und somit ein Hyponym von *strando* zu werden. Das belegt er zwar nicht mit Quellen, kann jedoch mit Waringhien (2002: 881 und 1087) gestützt werden. Das ist ein Beispiel für eine Tendenz, ein Synonym durch Zuordnung einer speziellen Bedeutung „loszuwerden“.

Im Esperanto ist jedoch auch die gegenläufige Tendenz zu beobachten, nämlich die Verwendung von Neologismen. Zamenhof (1991<sup>10</sup>: 47-48) unterscheidet dabei zwischen einer *neuen Form* (für ein grundlegendes/offizielles Simplex; sie bedarf, falls sie kein Internationalismus im Sinne von Regel 15 der Esperanto-Grammatik ist (Zamenhof 1991<sup>10</sup>: 67), einer Offizialisierung) und einem *neuen Wort* (für eine Zusammensetzung aus grundlegenden/offiziellen Simplizia).  
Motive für die Verwendung von Neologismen:

Beispiel 2.6.4-5: linker *maldekstra*  
neu: *liva* (Motiv: Vermeidung von  
Schwerfälligkeit, z. B. in poetischen Texten;  
s. auch Kapitel 2.6.5)

Beispiel 2.6.4-6: Frikativ *frota sono* (wörtlich: reibender Laut)  
neu: *frikativo* (Motiv: Fachausdruck)  
Wells (1978: 62) würde *frota sono* eher  
gegenüber Laien und *frikativo* eher gegenüber  
seinen Phonetik-Kollegen verwenden

(weitere Synonyme sind in den Beispielen 2.6.5-2 bis 2.6.5-6 in Kapitel 2.6.5).

Nach Gledhill (2000<sup>2</sup>: 119) und Hagège (1985) werden Wörter, die analytisch aus grundlegenden/offiziellen Lexemen gebildet werden, besonders von chinesischen und japanischen Esperanto-Sprechern den Neologismen vorgezogen. Eine naheliegende Erklärung könnte wie folgt formuliert werden: Um analytisch gebildete Wörter zu beherrschen sind weniger Lexeme erforderlich als bei der ausschließlichen Verwendung von Simplizia (sobald Neologismen im „mentalen Lexikon eingetragen“ sind, sind sie keine Neologismen mehr). Warum kommt aber dieser Vorteil bei chinesischen und japanischen Esperanto-Sprechern stärker zum Tragen? Weil viele Neologismen

Ähnlichkeit mit äquivalenten Lexemen in Sprachen der indoeuropäischen Sprachfamilie haben (s. Beispiel 2.6.4-6) und für deren Sprecher offensichtlich weniger lexikalischen „Ballast“ darstellen als für Sprecher einer Sprache aus einer anderen Sprachfamilie.

### 2.6.5 Antonymie, Meronymie, Hyponymie und Hyperonymie

Antonymie (Gegensatzrelation) wird im Esperanto morphologisch durch das Präfix */mal/* ausgedrückt:

Beispiel 2.6.5-1:    klein            *malgranda*  
                           groß             *granda*

Offensichtlich mit der Motivation, weniger schwerfällige Wörter und eine größere Variationsmöglichkeit für literarische Zwecke zu bekommen (Gledhill 2000<sup>2</sup>: 119), ersetzen einige Esperanto-Sprecher morphologische Antonyme durch Neologismen:

Beispiel 2.6.5-2:    dumm            *stulta* statt *malsagha*

Beispiel 2.6.5-3:    linker            *liva* statt *maldekstra*

Beispiel 2.6.4-4:    selten            *rara* statt *malofta*

Beispiel 2.6.4-5:    traurig          *trista* statt *malgaja*

Beispiel 2.6.5-6:    kurz             *kurta* statt *mallonga*

Wells (1978: 64) unterscheidet Antonymie von der Sinnrelation Konversion und nennt als Beispiele u. a. *acheti – vendi* (kaufen – verkaufen) und *instruisto – lernanto* (Lehrer – Schüler).

Für Meronyme (Teil-Ganzes-Beziehungen) werden im Esperanto in vielen Fällen eigenständige Lexeme verwendet (wie z. B. auch im Deutschen):

Beispiel 2.6.5-7:

		korpo (Körper)			
	/		\	\	
kapo (Kopf)		trunko (Rumpf)	brako (Arm)		usw.

In einigen Fällen werden meronymische Beziehungen morphologisch gebildet, und zwar durch das Sammelbegriffs-Morphem */ar/* und das Teil-eines-Ganzen-Morphem */er/*:

Beispiel 2.6.5-8:	Stufe	/shtup/o/
	Treppe	/shtup/ar/o
Beispiel 2.6.5-9:	Baum	/arb/o/
	Wald	/arb/ar/o/
Beispiel 2.6.5-10:	Taste	/klav/o/
	Tastatur	/klav/ar/o/
Beispiel 2.6.5-11:	Sand	/sabl/o/
	Sandkorn	/sabl/er/o/

Interessanterweise gibt es hierzu offensichtlich keine Analogie im Deutschen.

Für Hyponyme werden im Esperanto (wie z. B. auch im Deutschen) eigenständige Lexeme verwendet:

Beispiel 2.6.5-12: *rozo*            Rose; Hyponym von *floro* (Blume)

Solches gilt auch für Hyperonyme:

Beispiel 2.6.5-13: *besto*            Tier; Hyperonym von *simio* (Affe)

Bezüglich Verwandtschaftsbezeichnungen bemerkt Wells (1978: 64), dass im Esperanto ursprünglich ein (Einzahl-)Wort fehlt, das Hyperonym sowohl zu „Ehemann“ (*edzo*) als auch zu „Ehefrau“ (*edzino*) ist (ebenso bezüglich „Bruder/Schwester“ und „Vater/Mutter“). Er verweist auf den beginnenden Gebrauch des zumindest morphologisch akzeptablen *geedzo*, was so viel wie „Ehepartner“ (maskulin oder feminin Singular) bedeutet.

## 2.6.6 Farben

Farbadjektive bilden ein Wortfeld. Diesem gegenüber steht ein als „konzeptuelles Feld“ bezeichneter Phänomenbereich (im Beispiel das Farbkontinuum<sup>22</sup>). Wells (1978: 65) verweist einerseits auf die potentielle Willkür, mit der das Farbkontinuum in den verschiedenen Sprachen in einzelne Bereiche eingeteilt ist, die durch Farbwörter beschrieben sind (deren Anzahl übrigens von der betreffenden Sprache abhängt). In der Tat werden beispielsweise im Walisischen Gras und andere Pflanzen mit einem

<sup>22</sup> Physikalisch ist das Spektrum der Farben ein Kontinuum: Die Wellenlängen des sichtbaren Lichts gehen ohne Unterbrechung vom roten ins blaue Extrem.

Farbadjektiv beschrieben („glas“), dessen Äquivalent in anderen Fällen in vielen anderen europäischen Sprachen nicht „grün“, sondern „blau“ bedeutet.

Andererseits beschreibt Wells (1978: 65) ein Experiment (Berlin, Kay 1969), das folgendes Ergebnis hatte: Verschiedene Angehörige derselben Sprachgemeinschaft kennzeichneten

- (a) zwar die einzelnen Farbbereichsgrenzen durchaus unterschiedlich (und für sich selbst auch mit einer gewissen Unsicherheit),
- (b) aber den „Punkt“ im Farbkontinuum, der eine bestimmte, in der Sprache durch ein Farbwort bezeichnete Farbe besonders treffend wiedergibt (d. h. den „Prototyp“), in signifikanter Übereinstimmung (in Bezug auf dieselbe Sprache).

Laut Wells (1978: 65) lassen sich die Basis-Farbwörter<sup>23</sup> aller Sprachen in einer bestimmten universellen Hierarchie des Vorkommens anordnen:

***schwarz, weiß, rot, gelb, grün, blau, braun, grau, purpurn, rosa, orange***

Das heißt: Wenn in einer Sprache nur fünf Basis-Farbwörter vorkommen, entsprechen diese den fünf erstgenannten in der obigen Liste. Wenn es nur zwei Basis-Farbwörter gibt, dann sind sie die Äquivalente von „schwarz“ und „weiß“ usw.

Nach Einschätzung von Wells (1978: 66) weist Englisch elf und Esperanto acht Basis-Farbwörter auf (Letztere entsprechen den acht erstgenannten in der obigen Liste). Damit bestätigt also auch das Esperanto das oben beschriebene Universal.

---

<sup>23</sup> In einer Sprache ist ein Basis-Farbwort ein solches Farbwort, das aus nur einem morphologisch nicht zusammengesetzten Wort (Simplex) besteht (also nicht z. B. *helrugha* [hellrot]), kein Hyponym bezüglich eines anderen Farbwortes ist (also nicht z. B. *skarлата* [scharlachrot]), keiner Anwendungsbeschränkung unterliegt (also nicht z. B. *blonda* [beschränkt auf Haare]) und von allen Sprechern der betreffenden Sprache verstanden wird (also nicht z. B. *punca* [hochrot]).



## 2.6.7 Bedeutungsverengung, -erweiterung und Konnotationen

Ein Vergleich der Lexeme von Wortfeldern zu verschiedenen Zeitpunkten lässt auch beim Esperanto Phänomene wie Bedeutungsverengung, -erweiterung und Ausbildung von Konnotationen (Bedeutungskomponenten) erkennen (die folgenden Beispiele 2.6.7-1 bis 2.6.7-5 sind aus Sakaguchi 1998: 199ff):

Beispiel 2.6.7-1: *ami*

ursprünglich: 1. lieben, ein starkes Gefühl der  
Zuneigung empfinden

2. mögen, gern haben

jetzt: nur 1. Bedeutung; Wiedergabe  
der 2. Bedeutung durch *shati*

Beispiel 2.6.7-2: *facila*

ursprünglich: 1. leicht, mühelos

2. von geringem Gewicht

jetzt: nur 1. Bedeutung; Wiedergabe der 2. Be-  
deutung durch  
*malpeza* oder das neue Lexem *leghera*

Beispiel 2.6.7-3: *shtofa*

ursprünglich: 1. Stoff (aus Fasern hergestelltes  
Gewebe)

2. Material, Rohstoff, Werkstoff

3. organisches Gewebe, Zellgewebe

jetzt: nur 1. Bedeutung; Wiedergabe der 2. Be-  
deutung durch *materialo* und der 3. Be-  
deutung durch *histo*

Beispiel 2.6.7-4: *baki*

ursprünglich: backen (Brot u. Ä.)

jetzt: Bedeutungserweiterung auf „backen“ im  
übertragenen Sinne, z. B.  
*nove bakita esperantisto* (ein  
frischgebackener Esperantist)

Beispiel 2.6.7-5: *kongreso*

ursprünglich: Kongress (fachliche [Arbeits-])

jetzt: Zusammenkunft)  
 assoziiert mit Konnotation „Unterhaltung,  
 Heiterkeit, Amüsement“ (wegen dieser  
 Komponente von Esperanto-  
 Kongressen)

### 2.6.8 Kollokationen

Gledhill (2000<sup>2</sup>: 124ff.) nennt einige Beispiele für Kollokationen, von denen hier zwei wiedergegeben werden:

Beispiel 2.6.8-1: *vera miraklo* ein wahres Wunder  
 (*vera* und *miraklo* kommen durchaus  
 auch in Verbindung mit anderen  
 Wörtern vor)

Beispiel 2.6.8-2: *krokodilaj larmoj* Krokodilstränen  
 (Vorkommen der Wörter analog  
 Beispiel 2.6.8-1)

Ein aktuelles Beispiel liefert auch der aktuelle Hörbeleg (Beispiel 2.6.8-3), eine Kollokation, die der Autor der vorliegenden Arbeit öfter in der Esperanto-Sendung des chinesischen Auslandsrundfunks hörte:

Beispiel 2.6.8-3 *chio iris glate* alles ging glatt  
 (Vorkommen der Wörter analog  
 Beispiel 2.6.8-1)

Gledhill (2000<sup>2</sup>: 131) verweist darauf, dass „Esperantisten sich vor Versuchen hüten, nationale Unterschiede zwischen Sprechern zu identifizieren [...], hauptsächlich weil von charakteristischen Sprechmustern angenommen wird, dass sie möglicherweise das universelle Ideal unterminieren ...“. Er zeigt anhand einiger beispielhafter, offensichtlich unter dem Einfluss von Ethnosprachen im Esperanto entstandener Kollokationen und einer Korpus-Untersuchung, dass es derartige Unterschiede gibt (Gledhill 2000<sup>2</sup>: 130ff.).

## 2.7 Worthäufigkeit

### 2.7.1 Offizielle Basis-Wortwurzelammlung (BRO)

Auf der Basis von Untersuchungen zur Worthäufigkeit im Esperanto erarbeitete die Esperanto-Akademie eine „offizielle Basis-Wortwurzelammlung“ („Baza Radikaro Oficiala“; BRO, Akademio de Esperanto [Hrsg.] 2007<sup>2</sup>: 23ff.). Diese beruht auf einer Metaanalyse von acht in der einschlägigen Literatur verfügbaren Worthäufigkeitslisten verschiedener Autoren<sup>24</sup>, Listen, die z. T. nach sehr unterschiedlichen Kriterien ermittelt worden waren, sowie einer neunten Liste<sup>25</sup>, die von der Esperanto-Akademie selbst erstellt wurde. Ergebnis dieser Analysen sind neun Häufigkeitsgruppen BRO1 bis BRO9, wobei BRO1 alle diejenigen Wörter umfasst, die in allen neun Worthäufigkeitslisten vorkommen; BRO2 jene, die in acht Worthäufigkeitslisten vorkommen usw.

### 2.7.2 Andere Worthäufigkeitslisten

Im Frequenzwörterbuch Dietze (1989) sind die 1093 häufigsten Wortwurzeln der Esperanto-Literatursprache gelistet. Grundlage ist eine rechnergestützte Analyse von politischen, kulturpolitischen und populärwissenschaftlichen Texten folgender Esperanto-Zeitschriften aus dem Zeitraum zwischen 1969 und 1987 (im Folgenden wird dieser Korpus als Dietze-Korpus bezeichnet):

- (a) „La pacaktivulo“ (Belgien)
- (b) „Monato“ (Belgien)
- (c) „Kontakto“ (Finnland)
- (d) „Paco“ (Deutsche Demokratische Republik)
- (e) „Paco“ (Tschechoslowakei)
- (f) „Hungara vivo“ (Ungarn)
- (g) „Bulgara esperantisto“ (Bulgarien)
- (h) „La espero el Koreujo“ (Republik Korea)
- (i) „El Popola Chinio“ (Volksrepublik China)

<sup>24</sup> Blaas (1951); Bovet, Ith (1956); Ministère de l'éducation nationale (1954) (zwei Worthäufigkeitslisten); Rublev (1927); Setälä (1949) (drei Worthäufigkeitslisten)

<sup>25</sup> Einzelheiten zu dieser neunten Liste: Akademio de Esperanto (Hrsg.) (2007<sup>2</sup>: 24)

Tabelle 6 gibt statistische Merkmale des Dietze-Korpus.

<b>Merkmal</b>	<b>Größe (Wert)</b>
Textlänge	31699
Wortschatzumfang	2058
Mittlere Häufigkeit des Wortschatzes (= Textlänge/Wortschatzumfang)	15,4
Extensität (= Wortschatzumfang/Textlänge)	0,065
Anzahl der Wortwurzeln mit Häufigkeit > 2	1093
Anzahl der Wortwurzeln mit Häufigkeit 1 oder 2	965
Anteil der 50 häufigsten Wortwurzeln am Wortschatzumfang	2,4%
Konzentration des Wortschatzes (Anteil der 50 häufigsten Wortwurzeln an der Textlänge)	50,3%
Die vier häufigsten Wörter und ihre Häufigkeiten	„la“ (= der, die, das): 3540 „de“ (= von): 1797 „kaj“ (= und): 1059 „en“ (= in, an): 744
Gravität (= Anzahl der Wortwurzeln mit Häufigkeit > 2/Textlänge)	0,034
Exklusivität (= Anzahl der Wortwurzeln mit Häufigkeit 1 oder 2/Textlänge)	0,030
Variabilität (= Anzahl der Wortwurzeln mit Häufigkeit 1 oder 2/Wortschatzumfang)	46,9%
Rhythmus der Wurzelwiederholungen (= [Textlänge - Wortschatzumfang]/Anzahl der Wortwurzeln mit Häufigkeit > 2)	27,12

*Tabelle 6: Dietze-Korpus: statistische Merkmale (erarbeitet unter Rückgriff auf die Angaben von Dietze 1989: 4ff.)*

Es ist festzustellen, dass auch hier das Zipsche Gesetz gilt: Im Dietze-Korpus ist das Wort mit dem Rang 1 etwa zweimal so häufig wie das Wort mit dem Rang 2, etwa dreimal so häufig wie das Wort mit dem Rang 3 und etwa viermal so häufig wie das Wort mit dem Rang 4. Eine Worthäufigkeitsliste findet sich ebenfalls in Gledhill (2000<sup>2</sup>: 144ff.). Darin sind allerdings nur die ersten 100 Ränge aufgeführt. Sie ist daher für eine Untersuchung des Zusammenhangs zwischen der Zugehörigkeit eines Wortes zu einer der BRO-Häufigkeitsgruppen und seiner Häufigkeit nicht geeignet (s. Kapitel 3.7).

## **2.8 Zusammenhang zwischen Bekanntheit und Häufigkeit von Wörtern und Phraseologismen**

Bisherige esperantologische Untersuchungen hatten hauptsächlich den Zweck, Erkenntnisse zum Erlernen und Verstehen von Esperanto bzw. Esperanto-Texten zu gewinnen, und fokussierten dabei kaum auf eine Ermittlung des Zusammenhangs zwischen Bekanntheit von Esperanto-Wörtern in Abhängigkeit von ihrer Häufigkeit. Selbst experimentelle Arbeiten erzielten daher oft nur wenig ergiebige bis vage Erkenntnisse wie z. B. solche, dass ein englischsprachiger Schüler während sechs Monaten so viel Esperanto lernt, wie er Französisch während vier Jahren lernt (vgl. Janton 1993<sup>2</sup>). Es existieren nur sehr wenige aktuelle Arbeiten, die sich explizit auf den Esperanto-Wortschatz beziehen und einen empirisch quantifizierenden Zugang ermöglichen. So ist man hier auf alte Untersuchungen wie z. B. die von Walther (Hrsg.) (1970) angewiesen, die zumindest so interessante Korrelationen wie die folgenden nannte: Um 99% eines Esperanto-Textes zu verstehen, muss man 2000 Wörter kennen; analog dazu im Englischen 7000 Wörter; um 99,75% eines Esperanto-Textes zu verstehen, muss man 2800 Wörter kennen; analog dazu im Englischen 10000 Wörter.

Köhler, Rapp (2007) haben eine aktuelle Studie zum Zusammenhang zwischen Wort-Vertrautheit und Worthäufigkeit des Englischen vorgelegt. Das Maß für die Wort-Vertrautheit ergibt sich darin aus einer entsprechenden Einschätzung durch Testpersonen. Die Ergebnisse dieser Studie auf andere Sprachen zu verallgemeinern, ist jedoch aufgrund der bisher spärlichen Vergleichszahlen für andere Sprachen noch sehr eingeschränkt.

Juska-Bacher (2006) präsentiert eine hochinteressante statistische Arbeit zum Zusammenhang zwischen Bekanntheit und Häufigkeit von Phraseologismen des Deutschen. Sie resümiert besonders über den desolaten Stand der empirischen Forschung in der Phraseologie der letzten 15 Jahre und sieht die Ursache dafür in der bisher zu geringen Nutzung statistischer Verfahren. Sie plädiert überzeugend für eine systematische Verwendung statistischer Methoden in empirischen Ansätzen der modernen Linguistik.

Bezüglich des Zusammenhangs zwischen Häufigkeit und Geläufigkeit von Phraseologismen des Deutschen ergaben sich widersprüchliche Ergebnisse:

- (a) „Die hier vorhandene Korrelation zwischen Frequenz und Geläufigkeit (viele Belegstellen – hohe Geläufigkeit) bei Phraseologismen entspricht nicht den Ergebnissen aus Ďurčos Untersuchungen zur Frequenz von Sprichwörtern im Korpus und zur Geläufigkeit bei Muttersprachlern“ (Hallsteinsdóttir, Šajánková, Quasthoff 2006).
- (b) „... hat sich das erwartete Ergebnis darin bestätigt, dass der Bekanntheitsgrad, die lexikographische Erfassung und die textuelle Vorkommenshäufigkeit keine korrelierenden Werte sind“ (Ďurčo 2001: 102).

Offensichtlich ist eine unterschiedliche Festlegung von Kriterien der Merkmale Häufigkeit und Bekanntheit die Ursache für die unterschiedlichen Ergebnisse: „Bei der Festlegung der Kriterien für die Gruppeneinteilung stößt man auf die Problematik, dass die Bestimmung, wann ein Phraseologismus als 'frequent' oder 'geläufig' anzusehen ist, einen gewissen Grad an Beliebigkeit aufweist.“ und „Ďurčos Gruppierung enthält jedoch nur eine Gruppe [...], die für Fremdsprachenlerner relevant ist“ (Hallsteinsdóttir, Šajánková, Quasthoff 2006).

### **3 Weiterführende quantitative Studien zum Esperanto: Phonologie, Morphologie, Syntax, Lexik, lexikalische Semantik, Worthäufigkeit**

#### **3.1 Phonemhäufigkeit**

##### **3.1.1 Ziel der Untersuchung zur Phonemhäufigkeit**

Ziel der Untersuchung zur Phonemhäufigkeit ist, für die Plansprache Esperanto die in Strauss, Altmann, Best (2007) aufgestellte Hypothese zu testen, nach der die geordnete Reihenfolge der Phonemhäufigkeiten einer regulären Wahrscheinlichkeitsfunktion oder regulären monoton abnehmenden Funktion folgt. Dazu sollen an die empirisch ermittelte Phonem-Häufigkeitsverteilung Modellfunktionen angepasst werden, die sich bereits bei Phonem-Häufigkeitsverteilungen anderer Sprachen bewährt haben (vgl. Kapitel 2.2).

##### **3.1.2 Anwendung verschiedener Modelle zur Beschreibung der Phonemhäufigkeit**

Die in Kapitel 2.2 genannten Modelle<sup>26</sup> wurden auf die empirisch erhaltene relative Phonemhäufigkeit des Esperanto angewendet. Tabelle 7 zeigt die Zusammenfassung.

---

<sup>26</sup> Damit in Einklang mit den anderen Modellen die Zahl der Parameter nicht größer als drei ist, wurde in der vorliegenden Arbeit im Modell von Gleichung 8 nur der erste Summand berücksichtigt (also zwei Parameter). Da sich eine gute Anpassung zeigte (s. Tabelle 7), wurde auf die Heranziehung weiterer Summanden (und damit weiterer Parameter) verzichtet.

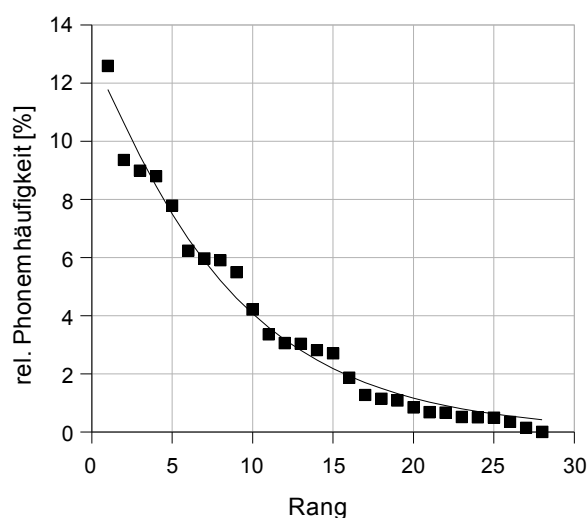
Rang x	Pho- nem	empirisch ermittelte abs. Phonem- häufigk. $f_{x,abs}$	empirisch ermittelte rel. Phonem- häufigk. $f_x$	Modell							
				Tuldava	Zipf	Yule	Naranan/B.	Altmann	Sigurd	Good	Altmann/P.
1	a	62471	12,59	13,36	14,76	11,78	11,66	12,59	12,59	14,03	12,01
2	e	46444	9,36	10,56	9,55	10,64	11,80	11,00	11,09	10,45	10,65
3	i	44608	8,99	8,92	7,40	9,50	9,68	9,63	9,78	8,67	9,44
4	o	43665	8,80	7,76	6,18	8,46	7,95	8,44	8,61	7,48	8,36
5	n	38654	7,79	6,86	5,37	7,51	6,66	7,41	7,59	6,59	7,41
6	l	30963	6,24	6,12	4,79	6,65	5,69	6,52	6,69	5,87	6,57
7	r	29623	5,97	5,50	4,35	5,89	4,95	5,74	5,89	5,28	5,82
8	s	29325	5,91	4,96	4,00	5,21	4,37	5,06	5,19	4,77	5,16
9	t	27291	5,50	4,49	3,71	4,61	3,90	4,47	4,58	4,32	4,57
10	k	20939	4,22	4,06	3,47	4,07	3,51	3,95	4,03	3,92	4,05
11	u	16672	3,36	3,68	3,27	3,60	3,19	3,50	3,55	3,57	3,59
12	m	15184	3,06	3,32	3,10	3,18	2,92	3,10	3,13	3,24	3,18
13	d	15084	3,04	3,00	2,95	2,81	2,69	2,75	2,76	2,94	2,82
14	p	13993	2,82	2,70	2,81	2,48	2,49	2,45	2,43	2,67	2,5
15	j	13447	2,71	2,42	2,69	2,19	2,32	2,18	2,14	2,41	2,22
16	v	9279	1,87	2,16	2,58	1,93	2,17	1,94	1,89	2,17	1,96
17	g	6351	1,28	1,92	2,49	1,70	2,03	1,73	1,66	1,95	1,74
18	b	5706	1,15	1,69	2,40	1,50	1,91	1,55	1,47	1,74	1,54
19	f	5409	1,09	1,47	2,32	1,33	1,81	1,38	1,29	1,54	1,37
20	c	4218	0,85	1,26	2,25	1,17	1,71	1,24	1,14	1,36	1,21
21	ĝ	3424	0,69	1,07	2,18	1,03	1,62	1,11	1,00	1,18	1,07
22	ĉ	3325	0,67	0,88	2,12	0,91	1,54	0,99	0,88	1,01	0,95
23	z	2580	0,52	0,70	2,06	0,80	1,47	0,89	0,78	0,84	0,84
24	ŭ	2531	0,51	0,53	2,00	0,71	1,40	0,80	0,69	0,69	0,75
25	h	2481	0,50	0,36	1,95	0,62	1,34	0,72	0,61	0,54	0,66
26	ŝ	1737	0,35	0,20	1,91	0,55	1,29	0,65	0,53	0,40	0,59
27	ĵ	744	0,15	0,05	1,86	0,49	1,23	0,59	0,47	0,26	0,52
28	ĥ	50	0,01	0,00	1,82	0,43	1,19	0,53	0,41	0,13	0,46
Bestimmtheitsmaß:				D=0,98	D=0,82	D=0,98	D=0,92	D=0,98	D=0,98	D=0,96	D=0,98
Abweichungsquadratsumme:				$r^2=7,79$	$r^2=58,83$	$r^2=5,85$	$r^2=26,45$	$r^2=7,59$	$r^2=6,82$	$r^2=12,04$	$r^2=5,96$
Parameter:				a=13,3562 b=-4,0370	A=14,7604 b=0,6284	a=13,3814 b=-0,0361 d=0,8805	$a_1=1,1568$ $a_2=1,6273$ C=59,3324	a=87,2247 b=75,9729	q=0,8812		a=13,5557 b=0,1207

*Tabelle 7: Anwendung verschiedener Modelle zur Beschreibung der relativen Phonemhäufigkeit im Esperanto*

Mit dem Zipf- und Naranan/Balasubrahmanyam-Modell ergibt sich eine vergleichsweise schlechte Anpassung (Bestimmtheitsmaß: 0,82 bzw. 0,92). Ein mit dem letztgenannten Wert vergleichbarer Wert hatte sich mit diesen Modellen nach Strauss, Altmann, Best (2007) auch bezüglich des Hawaiischen ergeben. Mit den anderen Modellen ergeben sich bezüglich des Esperanto Anpassungen mit Bestimmtheitsmaßen von 0,96 und 0,98. Die beste Anpassung ergibt sich mit dem Yule-Modell (Bestimmtheitsmaß: 0,98; Abweichungsquadratsumme: 5,85), was auch in Abbildung 6 gezeigt wird. Auch bezüglich des Hawaiischen hatte sich die beste Anpassung mit dem Yule-Modell ergeben (Bestimmtheitsmaß: 0,94; nach Strauss, Altmann, Best 2007). Das Ergebnis ist ebenfalls in Einklang mit dem Ergebnis der 95 Sprachen umfassenden Untersuchung von Tambovtsev, Martindale (2007), der zufolge



das Yule-Modell die beste Anpassung an die Phonemhäufigkeitsverteilung liefert.



*Abbildung 6: Relative Phonemhäufigkeit des Esperanto gegen den Rang sowie Anpassung mit dem Yule-Modell*

*Erläuterungen zu den Anpassungen der Modellfunktionen:*

- (a) Die Auswertung erfolgte mit dem Tabellenkalkulationsprogramm OpenOffice Calc 2.3.
- (b) Zur Anpassung des Good-Modells an die empirisch erhaltene relative Phonemhäufigkeit war keine Variation von Parametern erforderlich.
- (c) Zur Anpassung der übrigen Modelle an die empirisch erhaltene relative Phonemhäufigkeit war die Variation von einem bis drei Parametern erforderlich. Dies erfolgte mit Hilfe des (iterativen) Gauß-Newton-Verfahrens<sup>27</sup> (Björck 1996). In Anhang A ist beispielhaft das Tabellenkalkulationsblatt zur Anpassung des Yule-Modells mit dem Gauß-Newton-Verfahren gezeigt (die Anpassung der anderen Modelle erfolgte analog). Es sei angemerkt, dass die im Gauß-Newton-Verfahren benötigten Ableitungen durch Differenzenquotienten genähert wurden.

<sup>27</sup> Eine gute Beschreibung des Gauß-Newton-Verfahrens wird in der Internet-Enzyklopädie [http://en.wikipedia.org/wiki/Gauss%E2%80%93Newton\\_algorithm](http://en.wikipedia.org/wiki/Gauss%E2%80%93Newton_algorithm) (10.07.2008) gegeben.

### 3.1.3 Wertung von Affrikaten

Zur Klärung der Frage, ob die Affrikaten *c*, *ĉ* und *ĝ* mono- oder biphonematisch gewertet werden sollten (vgl. Kapitel 2.2), wurde das Yule-Modell auf die Phonemhäufigkeitsverteilung von Haszpra (2001) angewendet, Letztere allerdings dahingehend modifiziert, dass die genannten Affrikaten biphonematisch gewertet wurden. Es zeigt sich ein Bestimmtheitsmaß von 0,98 (wie bei monophonematischer Wertung) und eine Abweichungsquadratsumme von 7,32 (bei monophonematischer Betrachtung: 5,85). Offensichtlich ist die Anpassung bei biphonematischer Wertung etwas schlechter als bei monophonematischer Wertung. Einzelheiten sind in Anhang B.

### 3.1.4 Zusammenfassung: Untersuchung zur Phonemhäufigkeit

In der vorliegenden Arbeit wurden das Inventar der Phoneme des Esperanto und ihre Häufigkeitsverteilung dargestellt. An Letztere wurden verschiedene Modellfunktionen angepasst, die sich bereits bei Phonemhäufigkeitsverteilungen anderer Sprachen bewährt haben: Tuldava (1988); Zipf (1929, 1935, 1949); Yule (1924); Naranan, Balasubrahmanyan (1992a, b, 2000); Altmann (1993); Sigurd (1968); Good (1969); Altmann, Popescu<sup>28</sup>. Damit wurde für die Plansprache Esperanto die in Strauss, Altmann, Best (2007) aufgestellte Hypothese überprüft, nach der die geordnete Reihenfolge der Phonemhäufigkeiten einer regulären Wahrscheinlichkeitsfunktion oder regulären monoton abnehmenden Funktion folgt. Folgendes ist festzuhalten:

- (a) Esperanto verfügt über 28 Phoneme (23 Konsonanten; fünf Vokale); sie werden phonemisch verschriftet.
- (b) Die absteigende Rangfolge der relativen Phonemhäufigkeit sieht wie folgt aus (nach Haszpra 2001; Korpus-Umfang: 496.196 Phoneme; aus Haszpra 1998): 1. a (12,59%); 2. e (9,36%); 3. o (8,99%); 4. i (8,80%); 5. n (7,79%); 6. l (6,24%); 7. r (5,97%); 8. s (5,91%); 9. t (5,50%); 10. k (4,22%); 11. u (3,36%); 12. m (3,06%); 13. d (3,04%); 14. p (2,82%); 15. j (2,71%); 16. v (1,87%); 17. g (1,28%); 18. b (1,15%); 19. f (1,09%);

---

<sup>28</sup> Siehe Fußnote 2.

20. c (0,85%); 21. ĝ (0,69%); 22. ĉ (0,67%); 23. z (0,52%); 24. ŭ (0,51%); 25. h (0,50%); 26. ŝ (0,35%); 27. ĵ (0,15%); 28. ĥ (0,01%). Andere Untersuchungen (Jung 1926; Stancliff 1933; Sadler 1959; Harry 1967; Dominte 2001) ergaben ähnliche Ergebnisse.

- (c) Nach Dominte (2001) kann Esperanto mit seinem durchschnittlichen Vokal-zu-Konsonant-Verhältnis von 43 zu 57 als relativ vokalische Sprache betrachtet werden (mit einem Platz zwischen den vorwiegend vokalischen Sprachen Finnisch, Türkisch, Italienisch, Neugriechisch, Serbisch, Kroatisch, Rumänisch, Litauisch, Latein, Slowakisch, Französisch und den eher konsonantischen Sprachen Ungarisch, Schwedisch, Tschechisch, Deutsch und amerikanisches Englisch).
- (d) Mit dem Zipf- und Naranan/Balasubrahmanyam-Modell ergibt sich eine verhältnismäßig schlechte Anpassung (Bestimmtheitsmaß: 0,82 bzw. 0,92). Ein mit dem letztgenannten Wert vergleichbarer Wert hatte sich mit diesen Modellen auch bezüglich des Hawaiischen ergeben (nach Strauss, Altmann, Best 2007). Mit den anderen Modellen ergeben sich bezüglich des Esperanto Anpassungen mit Bestimmtheitsmaßen von 0,96 und 0,98. Die beste Anpassung ergibt sich mit dem Yule-Modell (Bestimmtheitsmaß: 0,98; Abweichungsquadratsumme: 5,85): relative Phonemhäufigkeit (in %)  $y = 13,3814 x^{0,0361} - 0,8805x$  (x: Rang). Auch bezüglich des Hawaiischen hatte sich die beste Anpassung mit dem Yule-Modell ergeben (Bestimmtheitsmaß: 0,94). Das Ergebnis ist ebenfalls in Einklang mit dem Ergebnis einer aktuellen Untersuchung von 95 Sprachen (Tambovtsev, Martindale 2007), der zufolge das Yule-Modell die beste Anpassung an die Phonemhäufigkeitsverteilung liefert.
- (e) Unter Zugrundelegung des Yule-Modells und der Phonemhäufigkeiten von Haszpra (2001) ergibt sich bei biphonematischer Wertung der drei Affrikaten c, ĉ und ĝ eine etwas schlechtere Anpassung als bei monophonematischer Wertung.

Die o. g. Hypothese wurde bestätigt. Dies war zu erwarten, weil es sich bei Esperanto um eine aposteriorische<sup>29</sup> Plansprache handelt und es somit

---

<sup>29</sup> Siehe Fußnote 19.

naheliegender ist, dass Gesetzmäßigkeiten, die für Ethnosprachen gelten, auch für Esperanto gelten.

## **3.2 Sprachtypologische Einordnung**

### **3.2.1 Ziel der Untersuchung zur sprachtypologischen Einordnung**

In Sikosek (2003<sup>2</sup>: 204) wird resümiert, dass nach Meinung der meisten Autoren und in passendster Verallgemeinerung Esperanto agglutinierend ist und einige Merkmale einer isolierenden Sprache aufweist (vgl. Kapitel 2.3.11). Ziel der Untersuchung zur sprachtypologischen Einordnung ist, dieses Resümee anhand eigener (quantitativer) Untersuchungen zu bestätigen oder zu widerlegen.

### **3.2.2 Isolierende Merkmale**

Vietnamesisch wird als „typische isolierende Sprache“ in Lyons (1995<sup>8</sup>: 191) genannt. Darin wird weiterhin ein Isoliertheitsgrad als Verhältnis von Morphemzahl zu Wörterzahl mit Beispielen für Englisch (1,68), Sanskrit (2,59) und Eskimo (3,72) angegeben (der Idealtyp einer Sprache mit isolierendem Sprachbau hat einen Isoliertheitsgrad von 1).

In Wendt (1977: 198-199) wird noch weiter unterschieden zwischen:

(a) wurzelisolierend:

Wörter als nicht durch Affixe erweiterbare Phonemgruppen; Beispiel:  
Chinesisch

(b) wortisolierend:

isoliert im Satz bleibende und durch Affixe erweiterbare Wörter; Funktion der Affixe: Wortbildung und Andeutung grammatischer Funktionen;  
Beispiel: Indonesisch

Isolierende Merkmale des Esperanto finden sich u. a. in seinen freien Morphemen (s. Kapitel 2.3.1). Viele von ihnen sind aber durchaus durch Affixe erweiterbar, z. B. /tuj/a/ (= sofortig). Daher handelt es sich vorwiegend um wortisolierende Merkmale. Exemplarisch wurden als Esperanto-Textprobe die Wörter und Morpheme des Märchentextes in Zamenhof (1991<sup>10</sup>: 107-108)

gezählt; Ergebnis:  $M = 210$  Morpheme und  $W = 109$  Wörter<sup>30</sup>. Damit ergibt sich ein Isoliertheitsgrad von 1,93. Auf der Grundlage dessen kann Esperanto als isolierender als Sanskrit sowie Eskimo und als weniger isolierend als Englisch eingestuft werden.

Der Kehrwert des Isoliertheitsgrades ist als Index des Synthetismus  $W/M$  einer der zehn in Greenberg (1960) genannten Indizes zur quantitativen Beschreibung morphologischer Merkmale. Auch Altmann, Lehfeld (1973), Kempgen, Lehfeldt (2004) und Hoffmann (2005) greifen auf diese Indizes zurück. Nach einem in Altmann, Lehfeldt (1973: 46-47) dargestellten und im Vergleich dazu in Kempgen, Lehfeldt (2004: 1243) leicht modifizierten klassifikatorischen Schlüssel spricht ein  $W/M \leq 0,90$  (wie im Fall des Esperanto mit  $W/M = 0,52$ ) gegen eine Einordnung als eine Sprache der in den beiden letztgenannten Quellen mit A bezeichneten Klasse (Repräsentant: Vietnamesisch). In Kempgen, Lehfeldt (2004: 1243) wird sie als „hochgradig analytisch, maximal isolierend“ beschrieben.

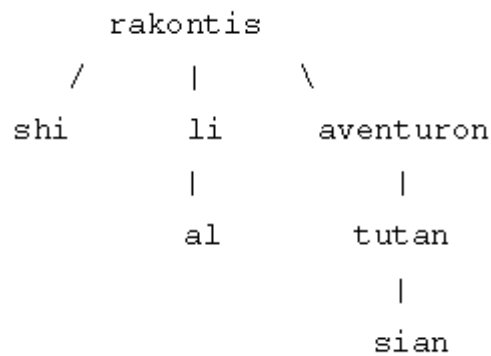
Als weiteres Entscheidungskriterium kann nach den genannten Quellen der Index der Isolation  $O/N$  herangezogen werden: Laut Kempgen, Lehfeldt (2004: 1236) ist dabei  $N$  die Anzahl der syntaktischen Konnexionen<sup>31</sup> und  $O$  die Anzahl der lediglich durch die Stellung signalisierten syntaktischen Konnexionen<sup>32</sup>. Der folgende Beispielsatz aus der untersuchten Esperanto-Textprobe veranschaulicht, wie in dieser Arbeit der Isolationsindex  $O/N$  für Esperanto ermittelt wurde: „Shi rakontis al li sian tutan aventuron“ = „Sie erzählte ihm [wörtlich: 'zu er'] ihr ganzes Abenteuer.“  
Abbildung 7 veranschaulicht die syntaktischen Konnexionen.

---

30 Die Anzahl der Stichproben (nämlich eine!) und deren Wortzahl (etwa 100) liegt somit in der Größenordnung dessen, was (je untersuchte Sprache) als Untersuchungsgrundlage von Greenberg (1960) herangezogen wurde. Bezüglich der statistischen Aussagekraft siehe Kempgen, Lehfeldt (2004: 1237).

31 In Greenberg (1960) wird eine syntaktische Konnexion als „nexus“ bezeichnet, also jeder Fall der Verwendung einer Richtlinie, um Beziehungen zwischen Wörtern anzuzeigen.

32 In Greenberg (1960) wurde die Abwesenheit eines Flexionsmorphems in einem Wort als Anzeige für Stellung („order“) als diejenige Methode angenommen, mit der zu diesem Wort eine Beziehung hergestellt wird.



*Abbildung 7: Syntaktische  
Konnexionen in einem Beispielsatz*

Die Striche in Abbildung 7 (also /, | und \) stellen die syntaktischen Konnexionen dar; hier  $N = 6$ . In diesem Beispielsatz gibt es nur eine syntaktische Konnexion, auf die ausschließlich durch Stellung hingewiesen wird: die syntaktische Konnexion zwischen „al“ und „li“. Das Pronomen „al“ (= zu) könnte auch vor „shi“ (= sie) stehen (ohne dass gegen die Grammatik verstoßen würde). Das würde die Bedeutung dahingehend ändern, dass nicht „sie ihm“, sondern „er ihr“ etwas erzählt (anders als der Akkusativ wird im Esperanto der Dativ nicht durch ein Flexionsmorphem markiert, sondern durch /al/). Daher ist in dem Beispielsatz  $O = 1$ . Analog wurden die übrigen Sätze der Textprobe behandelt. Es ergab sich für die Esperanto-Textprobe  $O = 46$  und  $N = 93$  somit ein Index der Isolation von  $O/N = 0,49$ .

Bei einem Index des Synthetismus von  $W/M \leq 0,90$  spricht (nach dem oben erwähnten klassifikatorischen Schlüssel) ein Index der Isolation von  $O/N < 0,27^{33}$  für eine Einordnung als eine Sprache der in Altmann, Lehfelddt (1973: 47) und Kempgen, Lehfelddt (2004: 1243) mit B bezeichneten Klasse (Repräsentanten: Altpersisch, Altenglisch, Sanskrit, Rigveda, Eskimo). Demnach kann Esperanto auch nicht dieser Klasse zugeordnet werden. Sie wird in Kempgen, Lehfelddt (2004: 1243) mit den Worten „gering isolierend, relativ freie Wortfolge“ beschrieben.

Die bisherigen Untersuchungen lassen also den Schluss zu, dass Esperanto bezüglich isolierenden Sprachbaus zwischen den Extremen maximal isolierend und gering isolierend liegt.

<sup>33</sup> Nach Altmann, Lehfelddt (1973: 47) ist das Kriterium  $O/N < 0,30$ .

### 3.2.3 Synthetische Merkmale

Merkmale agglutinierenden Sprachbaus sind typisch für Esperanto: Die meisten grammatischen Funktionen sind durch agglutinierte Affixe markiert, die Wortbildung erfolgt ebenfalls durch Agglutination. Dies wurde besonders durch die Ausführungen der Kapitel 2.3.1 bis 2.3.5 und 2.3.9 deutlich gemacht. Esperanto wird in Wendt (1977: 356) als „analytisch agglutinierend“ charakterisiert (wobei darin agglutinierend, isolierend, polysynthetisch und flektierend als die vier Haupttypen genannt sind). In Greenberg (1960) wird der Index der Agglutination  $A/J$  eingeführt:  $A$  ist die Anzahl der agglutinierenden Konstruktionen,  $J$  die Anzahl der Morph-Junktoren. Beispielsweise besteht das Esperanto-Wort */arb/ar/o/* (= Wald) aus den zwei agglutinierenden Konstruktionen */arb/ + /ar/* und */ar/ + /o/*. Dabei ist */arb/* das Morph, das das „Baum“-Morphem realisiert, */ar/* das Morph, das das „Sammlung von“-Morphem realisiert, und */o/* das Morph, das das Substantiv-Morphem realisiert. In diesem Beispiel gibt es zwei Morph-Junktoren, nämlich zwischen */arb/* und */ar/* und zwischen */ar/* und */o/*. In der in Kapitel 3.2.2 erwähnten Esperanto-Textprobe ist  $A = 103$  und  $J = 103$ , und daher der Index der Agglutination  $A/J = 1,00^{34}$ , d. h. der höchstmögliche Wert, was Esperanto klar als agglutinierende Sprache kennzeichnet. Zum Vergleich: Das in Lyons (1995<sup>8</sup>: 191) als Beispiel für eine agglutinierende Sprache genannte Türkisch hat laut Altmann, Lehfeldt (1973: 40) einen Index der Agglutination von  $A/J = 0,60$  (Schriftsprache) und  $A/J = 0,67$  (mündliche Sprache).

Flexion zeigen im Esperanto Substantive, Adjektive, Verben, Pronomina sowie substantivierte und adjektivierte Numeralien (zu Letzteren gehören im Esperanto Ordinalzahlen), und zwar mit Kongruenz bezüglich Numerus und Kasus, z. B. */bon/a/j/n/ /patr/o/j/n/* (= gute Väter [Akkusativ]). Allerdings sind die flexionstypischen Merkmale Fusion und Polysemasie nicht bzw. in nur vernachlässigbarem Umfang vorhanden. Weiterhin spricht die Tatsache, dass (im Gegensatz zu beispielsweise den Wörtern im Lateinischen) die weitaus meisten Wörter des Esperanto eindeutig oder konsequent in Morphe segmentierbar<sup>35</sup> sind, dagegen, Esperanto als typisch flektierende Sprache,

<sup>34</sup> Man beachte Fußnote 17.

<sup>35</sup> Der Mangel einer solchen Segmentierbarkeit charakterisiert laut Lyons (1995<sup>8</sup>: 194) Lateinisch als flektierende Sprache.

sondern es als agglutinierende Sprache einzustufen. In Greenberg (1960) wird der Index der Flexion I/W eingeführt und in Altmann, Lehfeldt (1973: 39) in I/M modifiziert: I ist die Anzahl der Flexionsmorpheme, W die der Wörter, M die Anzahl der Morpheme. So besteht das Esperanto-Wort /shi/n/ (= sie [Akkusativ]) aus dem Morph /shi/, das das „Pronomen 'sie“-Morphem realisiert, und dem Morph /n/, das das Akkusativ-Morphem realisiert. In diesem Beispiel gibt es zwei Morpheme, von denen eines ein Flexionsmorphem ist. In der in Kapitel 3.2.2 erwähnten Esperanto-Textprobe ist I = 28 und M = 210, daher ist der Index der Flexion I/M = 0,13. Zum Vergleich: Von den in Altmann, Lehfeldt (1973: 40) gelisteten 20 Sprachen hat nur Vietnamesisch mit I/M = 0,00 einen kleineren Index der Flexion als Esperanto.

Es soll nun durch fortgesetzte Anwendung des bereits in Kapitel 3.2.2 benutzten klassifikatorischen Schlüssel von (Kempgen, Lehfeldt 2004: 1243) herausgefunden werden, inwieweit damit die bisher in diesem Punkt bezüglich Esperanto genannten Merkmale (also viel agglutinierender und wenig flektierender Sprachbau) bestätigt oder widerlegt werden.

Da die bisher in Kapitel 3.2.2 herangezogenen Entscheidungskriterien der quantitativen morphologischen Typologie (Index des Synthetismus W/M und Index der Isolation O/N) noch zu keiner Einordnung geführt haben, kann als weiteres Entscheidungskriterium die Präfixation P/M herangezogen werden: Dabei ist P die Anzahl der Präfixe. In der in Kapitel 3.2.2 erwähnten Esperanto-Textprobe ist P = 15 und M = 210, daher ist die Präfixation P/M = 0,07.

Bei einem Index des Synthetismus von  $W/M \leq 0,90$  und einem Index der Isolation  $O/N \geq 0,27$  spricht (nach dem oben erwähnten klassifikatorischen Schlüssel) eine Präfixation von  $P/M > 0,10^{36}$  für eine Einordnung als eine Sprache der mit C bezeichneten Klasse (Repräsentant: Suaheli). Demnach kann Esperanto auch nicht dieser Klasse zugeordnet werden. Sie wird in Kempgen, Lehfeldt (2004: 1243) als „stark präfigierend, stark agglutinierend“ beschrieben.

Als weiteres Entscheidungskriterium wird (nach dem klassifikatorischen Schlüssel) die Suffixation S/M herangezogen: Dabei ist S die Anzahl der

---

36 Nach Altmann, Lehfeldt (1973: 47) ist das Kriterium  $P/M > 0,20$ .



Suffixe. In der Esperanto-Textprobe ist  $S = 19^{37}$ , und daher die Suffixation  $S/M = 0,07$ .

Bei einem Index des Synthetismus von  $W/M \leq 0,90$ , einem Index der Isolation von  $O/N \geq 0,27$  und einer Präfixation von  $P/M \leq 0,10$  spricht (nach dem klassifikatorischen Schlüssel von Kempgen, Lehfeldt 2004: 1243) eine Suffixation von  $S/M < 0,45$  für eine Einordnung als eine Sprache der mit D bezeichneten Klasse (Repräsentanten: mündliches Türkisch, Bengali, Neuenglisch, Neupersisch, Neugriechisch). Demnach kann Esperanto (nach dem erwähnten klassifikatorischen Schlüssel abschließend) dieser Klasse zugeordnet werden. Sie wird in Kempgen, Lehfeldt (2004: 1243) als „schwach präfigierend, schwach suffigierend“ beschrieben (das sagt aber nur wenig über z. B. agglutierende Merkmale aus: Eine Sprache, die zwar schwach präfigierend und schwach suffigierend ist, könnte beispielsweise wegen ihrer großen Zahl von Infixen durchaus stark agglutinierend sein). Daher ist ein Vergleich mit den „Klasse D“-Sprachen bezüglich ausgewählter Indizes, die sich auf isolierende, agglutierende und flektierende Merkmale beziehen, möglicherweise ergiebiger (Tabelle 8).

<b>Sprache</b>	<b>Index des Synthetismus W/M</b>	<b>Index der Isolation O/N</b>	<b>Index der Agglutination A/J</b>	<b>Index der Flexion I/M</b>
Türkisch (mündl.)	0,57	0,69	0,67	0,38
Neuenglisch	0,60	0,75	0,30	0,32
Neupersisch	0,66	0,52	0,34	0,26
Bengali	0,53	0,57	0,46	0,28
Neugriechisch	0,55	0,53	0,40	0,37
Esperanto	0,52	0,49	1,00	0,13

*Tabelle 8: Ausgewählte morphologische Indizes von „Klasse D“-Sprachen und Esperanto (die Werte der Indizes [mit Ausnahme der für Esperanto genannten] entstammen Altmann, Lehfeldt 1973: 40)*

Aus Tabelle 8 kann man Folgendes schließen: Esperanto ist agglutierender als (mündliches) Türkisch, kaum flektierend und weist isolierende Merkmale in

<sup>37</sup> Morpheme mit rein grammatischer Funktion, also die in Kapitel 2.3.3 unter (a) sowie (c) bis (n) aufgezählten, wurden nicht mit einbezogen. Doch selbst wenn man sie mit einbezieht (und dann für die Esperanto-Textprobe  $S = 67$  bekommt), ist die Suffixation  $S/M = 0,32 < 0,45$ .

einer Ausprägung auf, die geringer als die des Neuenglischen und in etwa vergleichbar mit der des Neugriechischen, Neupersischen und Bengali ist.

### **3.2.4 Inkorporierende und polysynthetische Merkmale**

Auch im Esperanto gibt es Merkmale polysynthetischen Sprachbaus; Beispiel (aufbauend auf dem Beispiel von Kapitel 2.3.4): /li/ /mal/san/igh/is/ (= er erkrankte; /igh/ realisiert das Morphem des Werdens oder des In-einen-Zustand-Kommens). Dieser Satz hat einen Index des Synthetismus von  $W/M = 0,40$ . Die alternative (isolierendere) Ausdrucksweise /li/ /igh/is/ /mal/san/a/ hat einen Index des Synthetismus von  $W/M = 0,50$ . Hier ein Beispiel für inkorporierenden Sprachbau: /li/ /en/mal/san/ul/ej/igh/is/ (= er kam ins Krankenhaus). Das Morphem /en/ kann auch frei sein und hat als lokale Präposition auch eine lexikalische Bedeutung (vgl. Bußmann 2002<sup>3</sup>: 402). Daher kann der Sprachbau dieses Beispielsatzes als inkorporierend angesehen werden (seine anderen Morpheme kommen entweder nicht frei vor oder haben keine lexikalische Bedeutung). Der Index des Synthetismus ist hier  $W/M = 0,25$ . Die alternative (isolierendere) Ausdrucksweise /li/ /en/igh/is/ /en/ /mal/san/ul/ej/o/n/ hat einen Index des Synthetismus von  $W/M = 0,36$ . Bis zu einem gewissen Grad besteht also Freiheit in der Verwendung einerseits eher isolierender, andererseits aber eher polysynthetischer oder inkorporierender sprachlicher Mittel. Inkorporierender und auch polysynthetischer Sprachbau werden daher zwar als mögliche, jedoch zumindest nicht vorherrschende sprachtypologische Merkmale des Esperanto angesehen.

### **3.2.5 Zusammenfassung: Untersuchung zur sprachtypologischen Einordnung**

Esperanto weist in erster Linie synthetischen Sprachbau auf. Dabei dominiert Agglutination im Vergleich zu inkorporierenden, polysynthetischen und kaum vorhandenen typisch flektierenden Merkmalen; Merkmale isolierenden Sprachbaus sind festzustellen. Diese Aussagen wurden durch eine quantitative morphologische Untersuchung einer kurzen Esperanto-Textprobe (Zamenhof 1991<sup>10</sup>: 107-108) untermauert: Es ergab sich unter Verwendung des klassifikatorischen Schlüssels von Altmann, Lehfelddt (1973: 47) und Kempgen, Lehfelddt (2004: 1243) eine Einordnung als Sprache der Klasse D,

dessen Repräsentanten mündliches Türkisch, Bengali, Neuenglisch, Neupersisch und Neugriechisch sind. – Alles in allem wird das Resümee von Sikosek (2003<sup>2</sup>: 204) bestätigt, dem zufolge nach Meinung der meisten Autoren und in passendster Verallgemeinerung Esperanto agglutinierend ist und einige Merkmale einer isolierenden Sprache aufweist.

### 3.3 Detaillierte Analyse zur Häufigkeitsverteilung von Morphemstrukturen des Esperanto

In Kapitel 2.3.10 wurde auf die Modellierbarkeit der Häufigkeitsverteilung der Morphemstrukturen des Esperanto mittels des Zipfschen Gesetzes hingewiesen. Eine detaillierte Analyse (Anhang C) führt auf Gleichung 9 (Bestimmtheitsmaß: 0,98); Abbildung 8 zeigt die graphische Auftragung der Häufigkeit der 20 häufigsten Morphemstrukturen des Esperanto (vgl. Kapitel 2.3.10, Tabelle 4) gegen den Rang und Modellierung nach Gleichung 9.

$$y = Ax^{-b} \quad (9)$$

$y$ : Häufigkeit (der Morphemstruktur)

$A$ : Parameter; hier:  $A = 12969,4608$

$b$ : Parameter  $b = 1,5392$

$x$ : Rang

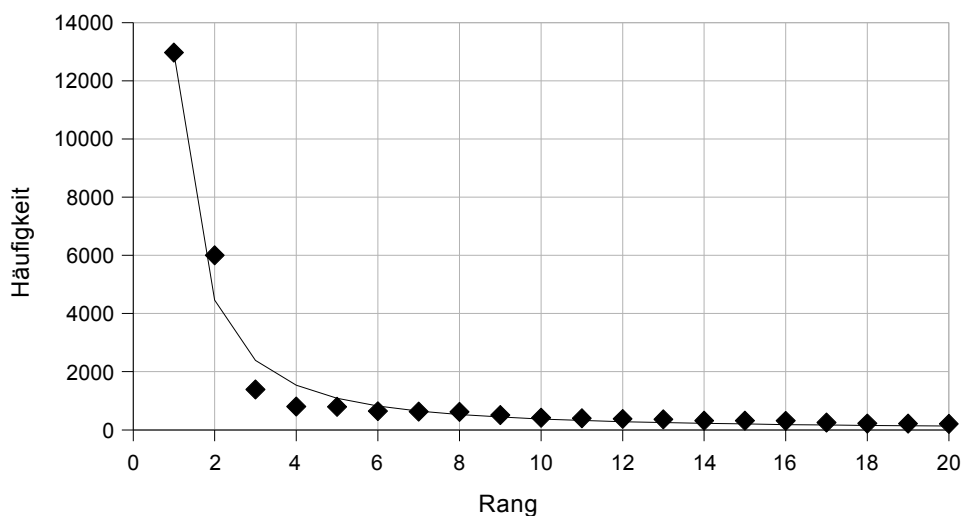


Abbildung 8: Häufigkeit der 20 häufigsten Morphemstrukturen des Esperanto (vgl. Kapitel 2.3.10, Tabelle 4) gegen den Rang sowie Modellierung nach Gleichung 9

### 3.4 Rechts- und Linkseinbettung von Phrasen und Relativsätzen

#### 3.4.1 Ziel der Untersuchung zur Rechts- und Linkseinbettung von Phrasen und Relativsätzen

Einbettungen verschiedenartiger Phrasen und Relativsätze sollen dargestellt und erläutert werden (weitgehend mit Erwähnung von Ergebnissen quantitativer Untersuchungen). Zur Darstellung wird die in Kapitel 2.4.2 erwähnte X-Bar-Theorie verwendet, weil damit Links- und Rechtseinbettungen sofort erkennbar sind.

#### 3.4.2 Flexionsphrase und Einbettung von Subjekt-Nominalphrasen

Der ganze Satz kann als eine Flexionsphrase IP (aus dem Englischen „inflectional phrase“) in der X-Bar-Syntax aufgefasst werden (Abbildung 9).

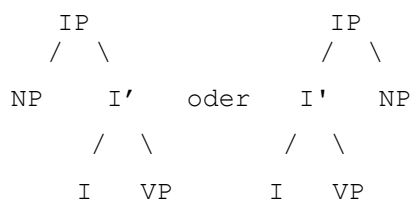


Abbildung 9: Flexionsphrase

Dabei können die das Subjekt darstellende NP als Spezifizierer und die das Prädikat darstellende VP als Komplement des Kopfes I betrachtet werden (vgl. die Analogie zu der in Abbildung 3 dargestellten Gesamtgrundstruktur). Der Kopf I wird in der Regel nicht lexikalisch ausgedrückt. Bei Sätzen mit finitem Verb (oder Hilfsverb) ist es üblich, das Verb-Flexionsmorphem (bzw. Hilfsverb) als Kopf I zu sehen (wie in Abbildung 2; darin ist *is* das Morph des Präteritums-Morphems des Verbs mit der Wortwurzel *don*<sup>38</sup>).

Die Linkseinbettung der Subjekt-Nominalphrase in der Flexionsphrase ist nach einer Untersuchung von Jansen (2007: 80) mit 97% die klar vorherrschende Einbettungsart bei den 280 Aussage-Hauptsätzen der grundlegenden Esperanto-Werke Zamenhof (1887) und Zamenhof (1991<sup>10</sup>). Auf die gleiche Quote kommt Gledhill (2000<sup>2</sup>: 87) (der von ihm untersuchte Korpus enthält

<sup>38</sup> Auf die (durchaus nicht unwichtige) Frage nach der Stellung des Kopfes I wird im Rahmen der vorliegenden Arbeit nicht eingegangen.

1045 relevante Sätze und besteht aus 156 Texten aus einem breiten Spektrum von Themen; Einzelheiten sind in Gledhill 2000<sup>2</sup>: 144 zu finden). Ein Beispiel für eine in der Flexionsphrase rechts<sup>39</sup> eingebettete Subjekt-Nominalphrase zeigt Abbildung 10.

*dangheron prezentus la shanghado* (Zamenhof 1991<sup>10</sup> : 46)  
eine Gefahr [Objekt] würde darstellen das Ändern

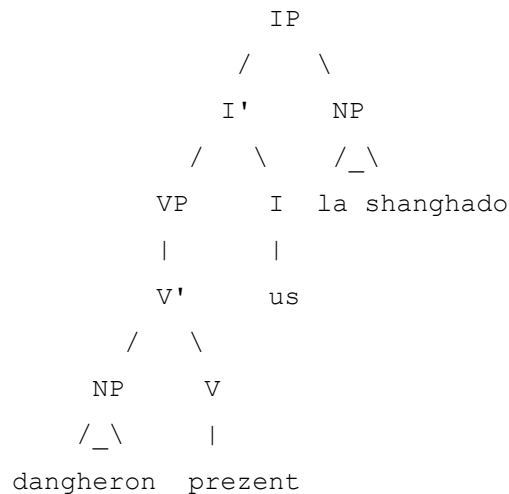


Abbildung 10: Flexionsphrase mit rechts eingebetteter Subjekt-Nominalphrase

Wells (1978: 41) schreibt bezüglich der fünf verschiedenen Möglichkeiten, Subjekt, Verb und Objekt anders als in dieser Reihenfolge anzuordnen, dass sie einen „gänzlich grammatischen Satz mit derselben Grundbedeutung, nur eventuell plus definierten Stilnuancen“ liefern.

Eine Besonderheit stellen Sätze mit einem Kopulativverb dar, z. B. *esti* (= sein), *ighi* (= werden). Damit das Subjekt als solches erkennbar ist, muss (mangels morphologischer Markierung) die Subjekt-Nominalphrase links in der Flexionsphrase eingebettet sein: „*glavo ighis plugilo* [= ein Schwert wurde ein Pflug] ... *plugilo ighis glavo* [= ein Pflug wurde ein Schwert]“ (Wells 1978: 43). Dies kann aus einem Beispielsatz wie „*leono estas besto* [= ein Löwe ist ein Tier]“ (Zamenhof 1991<sup>10</sup>: 83) mittelbar geschlossen werden.

<sup>39</sup> Alternativ kann die Rechtseinbettung als Folge einer Bewegungstransformation (Bewege- $\alpha$ ) der Objekt-Nominalphrase ins Vorfeld aus einer zugrundeliegenden Tiefenstruktur erklärt werden. Die Bewegungstransformation ist ein Transformationstyp, bei dem eine Phrase im Strukturbaum bewegt wird, indem eine Kopie der bewegten Phrase an der Ausgangsposition getilgt wird (Bußmann 2002<sup>3</sup>: 123).

### 3.4.3 Einbettung von Objekt-Nominalphrasen

Bei einer Verbalphrase stellt die Objekt-Nominalphrase das Komplement des Kopfes dar (also hier eines [transitiven] Verbs). Die Objekt-Nominalphrase kann rechts (wie in dem Beispiel von Abbildung 2) oder links in der Verbalphrase eingebettet sein (Abbildung 11).

Die Untersuchung Jansen (2007: 80) ergab eine vorherrschende Rechtseinbettung (83%) der Objekt-Nominalphrase in der Verbalphrase. Die größte Tendenz zur Linkseinbettung gibt es, wenn zumindest einer der Köpfe der Subjekt- oder Objekt-Nominalphrase ein Pronomen ist. Abbildung 11 zeigt ein hierzu Beispiel.

*mi ghin manghis* (Zamenhof 1991<sup>10</sup>: 106)

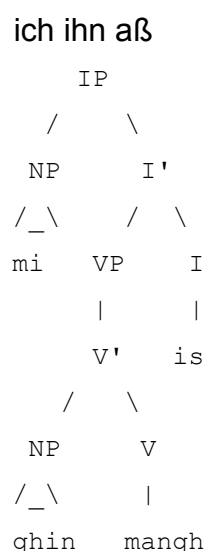


Abbildung 11: Links eingebettete Objekt-Nominalphrase in einer Verbalphrase

### 3.4.4 Einbettung von Adjektivphrasen

In Abbildung 2 gibt es eine Nominalphrase (*dolchan pomon*) mit links eingebetteter Adjektivphrase (*dolchan*). Laut der Untersuchung von Jansen (2007: 112) ist Linkseinbettung von Adjektivphrasen vorherrschend (88%) gegenüber Rechtseinbettung. Wells (1978: 49) verweist darauf, dass Esperanto mit seiner vorherrschenden Linkseinbettung von Adjektivphrasen zu einer (allerdings großen) Minderheit der Sprachen gehört (nach Greenberg 1974 gibt es diesbezüglich etwas mehr Sprachen mit Rechtseinbettung der Adjektivphrase). Bezüglich Esperanto schließt Jansen (2007: 112) auf folgende Motive für die (seltener) Rechtseinbettung:

Emphase des Adjektivs; besondere Umstände (z. B. Euphonie); Nominalphrase ist Korrelativwort (Abbildung 12); Adjektivphrase ist komplex (Abbildung 13).

*ni atendis audi ion novan* (Zamenhof 1991<sup>10</sup> : 131)

wir erwarteten [zu] hören etwas Neues

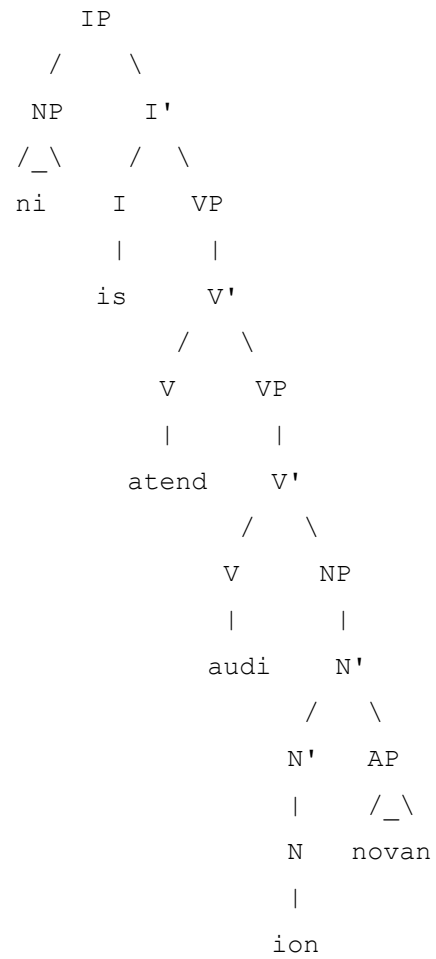


Abbildung 12: Rechts eingebettete Adjektivphrase bei Korrelativwort



*glaso da vino estas glas plena je vino* (Zamenhof 1991<sup>10</sup>: 131)

ein Glas [von] Wein ist ein Glas voll mit Wein

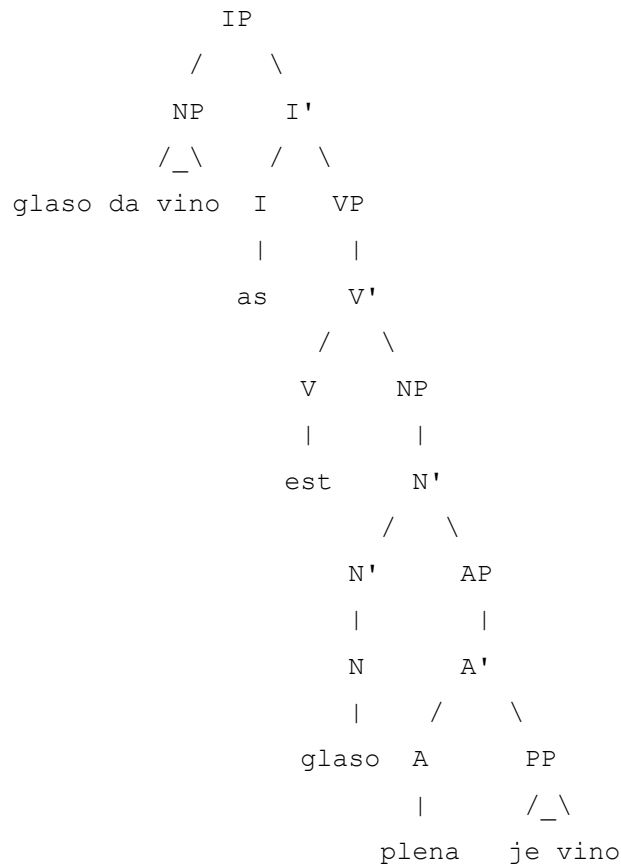


Abbildung 13: Rechts eingebettete komplexe Adjektivphrase

### 3.4.5 Einbettung von Präpositionalphrasen

Laut Greenberg (1963) sind Sprachen, in deren Syntax das Verb zuerst kommt, ausnahmslos präpositional und Sprachen, in deren Syntax das Verb zuletzt kommt, mit nur sehr wenigen Ausnahmen postpositional. Sprachen, bei denen das Verb überwiegend zwischen Subjekt und Objekt (oder zwischen Objekt und Subjekt) steht, können prä- oder postpositional sein (im Englischen und Deutschen gibt es wenige Beispiele für Postpositionen: *the world over*; *die Straße entlang*). Esperanto ist ausnahmslos präpositional. In Abbildung 2 wird beispielhaft auch eine Präpositionalphrase (*al mi*) gegeben. Sie kann als ein links eingebettetes Adjunkt in der Verbalphrase (*al mi donis dolchan pomon*) betrachtet werden. Was Präpositionalphrasen betrifft, beschränkt sich die in Jansen (2007: 117) veröffentlichte quantitative Untersuchung der grundlegenden Esperanto-Werke Zamenhof (1887) und Zamenhof (1991<sup>10</sup>) auf die Stellung des indirekten Objekts in Bezug auf die Objekt-Nominalphrase (bei



*li estas amiko de mia filo* (Sakaguchi 1998: 196)

er ist [ein] Freund von meinem Sohn

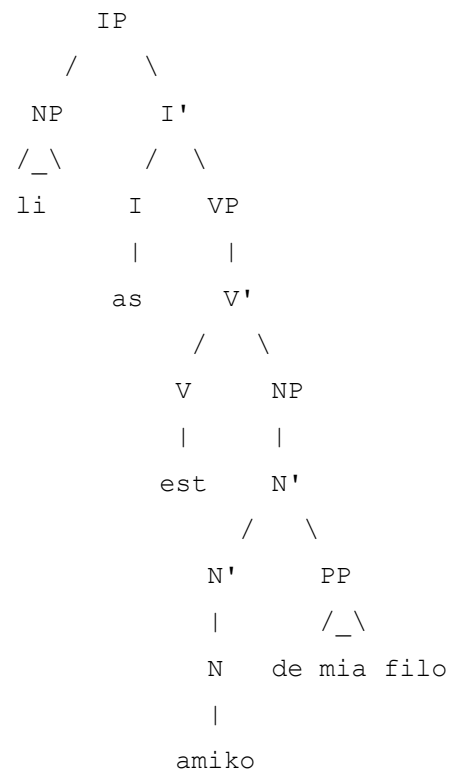


Abbildung 15: Rechts vom Bezugssubstantiv eingebettete Präpositionalphrase

*li estas de mia filo amiko* (Sakaguchi 1998: 196)

er ist von meinem Sohn [ein] Freund

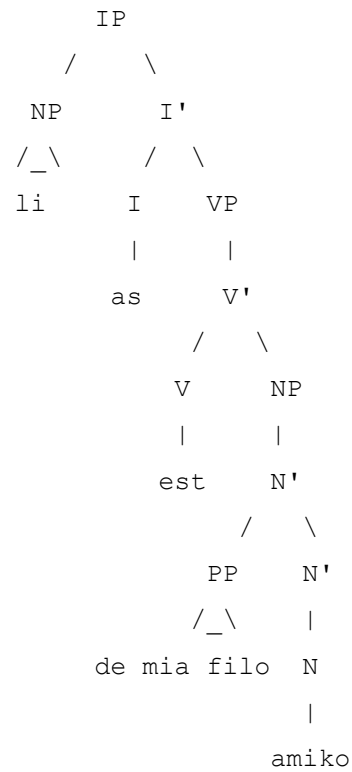


Abbildung 16: Links vom Bezugssubstantiv eingebettete Präpositionalphrase

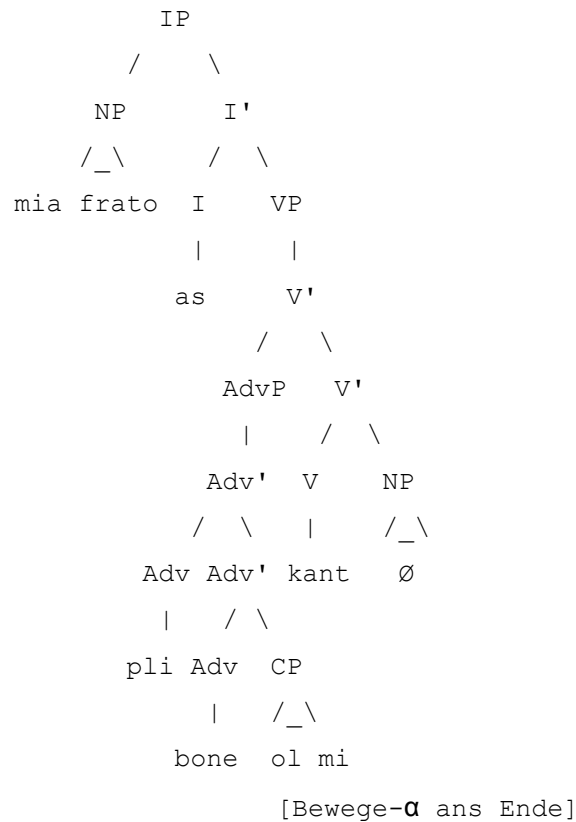
### 3.4.6 Einbettung von Adverbphrasen

Abbildung 17 zeigt einen Esperanto-Satz mit einer Adverbphrase als rechts eingebettetes Adjunkt in einer Verbalphrase.



*mia frato pli bone kantas ol mi* (Zamenhof 1991<sup>10</sup> : 54; 66; 72; 77)

mein Bruder besser [wörtl.: mehr gut] singt als ich



*Abbildung 18: Adverbphrase als links eingebettetes Adjunkt in einer Verbalphrase*

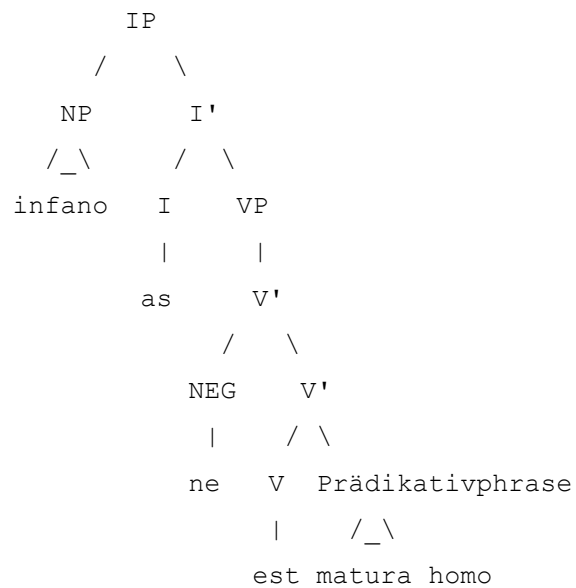
Eine Untersuchung von 122 relevanten Sätzen in den grundlegenden Esperanto-Werken Zamenhof (1887) und Zamenhof (1991<sup>10</sup>) mit Adverbphrasen ergab eine gleichgroße Zahl von Links- und Rechtseinbettungen in Verbalphrasen (Jansen 2007: 114).

### 3.4.7 Einbettung von Negationen

Abbildung 19 zeigt einen Satz mit der Negation *ne* (= nicht; nein) als links eingebettetes Adjunkt in einer Verbalphrase.

*Infano ne estas matura homo* (Zamenhof 1991<sup>10</sup>: 84)

[ein] Kind nicht ist [ein] reifer Mensch



*Abbildung 19: Negation ne als links eingebettetes Adjunkt in einer Verbalphrase*

Eine Untersuchung von 82 relevanten Vorkommen von *ne* in den grundlegenden Esperanto-Werken Zamenhof (1887) und Zamenhof (1991<sup>10</sup>) ergab zu 100% Linkseinbettung bei Fehlen eines Kontrastes und zu 62% Rechtsseinbettung bei Vorhandensein eines Kontrastes (Jansen 2007: 102). Zu Letzterer zeigt Abbildung 20 ein Beispiel.

*tio chi estis jam ne simpla pluvo, sed pluvego* (Zamenhof 1991<sup>10</sup>: 129)

das war schon nicht [ein] einfacher Regen, sondern Starkregen

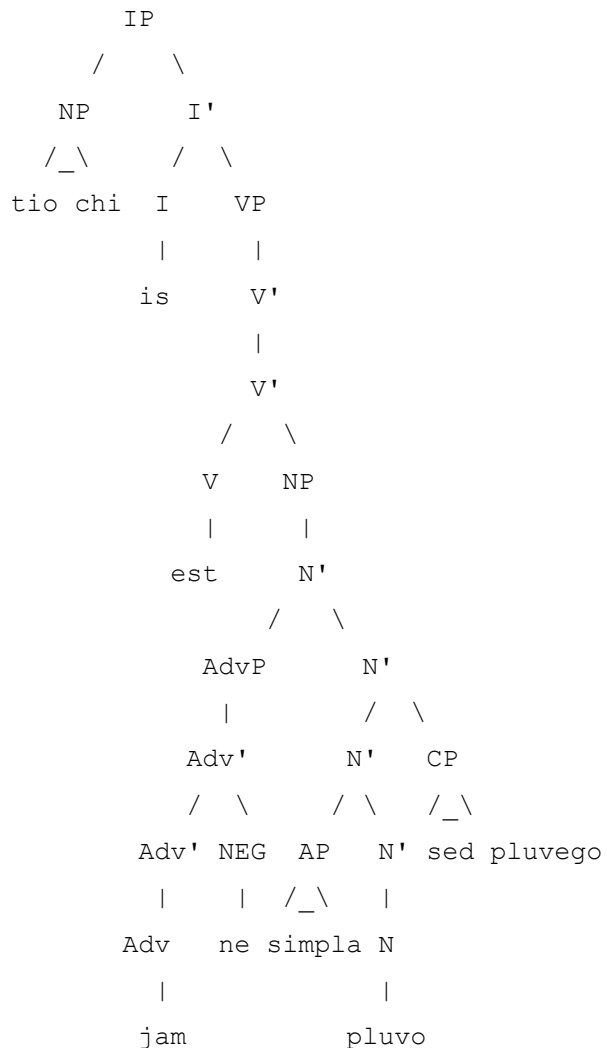


Abbildung 20: Negation *ne* als rechts eingebettetes Adjunkt in einer Verbalphrase (bei Kontrast)

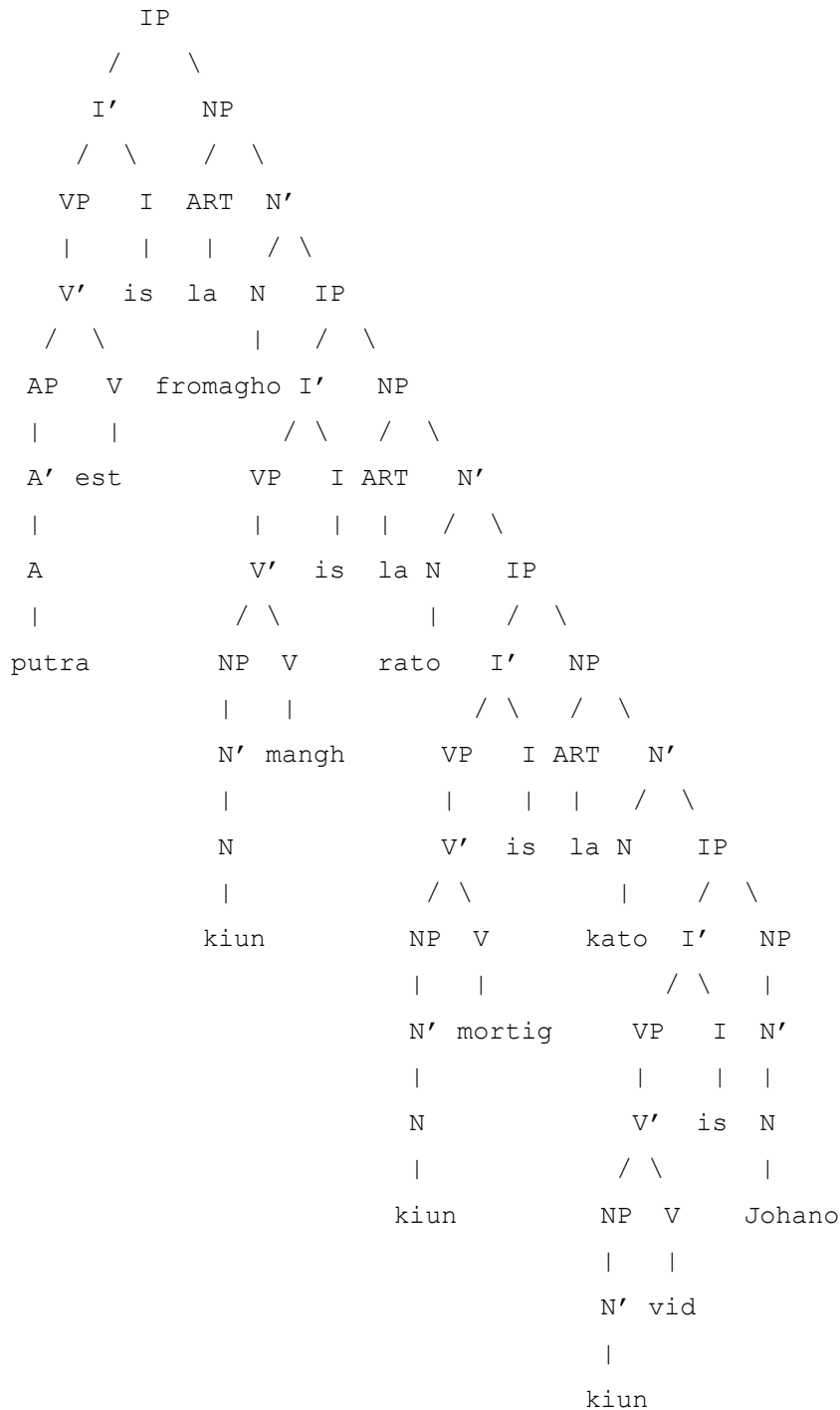
### 3.4.8 Einbettung von Relativsätzen

Sakaguchi (1998: 196) erläutert (mit Bezugnahme auf die Ausführungen von Wells 1978: 49ff.), dass Relativsätze im Esperanto gewöhnlich nach dem betreffenden Substantiv, also rechts, eingebettet werden. (Das Gegenbeispiel einer Linkseinbettung konnte vom Autor der vorliegenden Arbeit in den grundlegenden Esperanto-Werken Zamenhof 1887 und Zamenhof 1991<sup>10</sup> nicht gefunden werden.) Abbildung 21 zeigt ein Beispiel der Rechtseinbettung von Relativsätzen im Esperanto.



*putra estis la fromagho, kiun manghis la rato, kiun mortigis la kato, kiun vidis  
Johano* (Wells 1978: 51 und Sakaguchi 1998: 197)

verdorben war der Käse, den fraß die Ratte, die tötete der Kater, den sah  
Johano



*Abbildung 21: Rechtseinbettung von Relativsätzen im Esperanto*

Zum Vergleich zeigt Abbildung 22 die Übersetzung des Satzes aus Abbildung 21 ins Japanische (in dieser Sprache gibt es übrigens kein Relativpronomen).

*Johano ga mita neko ga korosita nezumi ga tabeta tiizu wa kusatte-ita*  
 (Wells 1978: 51ff. zitiert diesen Satz aus Kuno 1973)

Johano sah Kater [der] tötete Ratte [die] fraß Käse [der] verdorben war

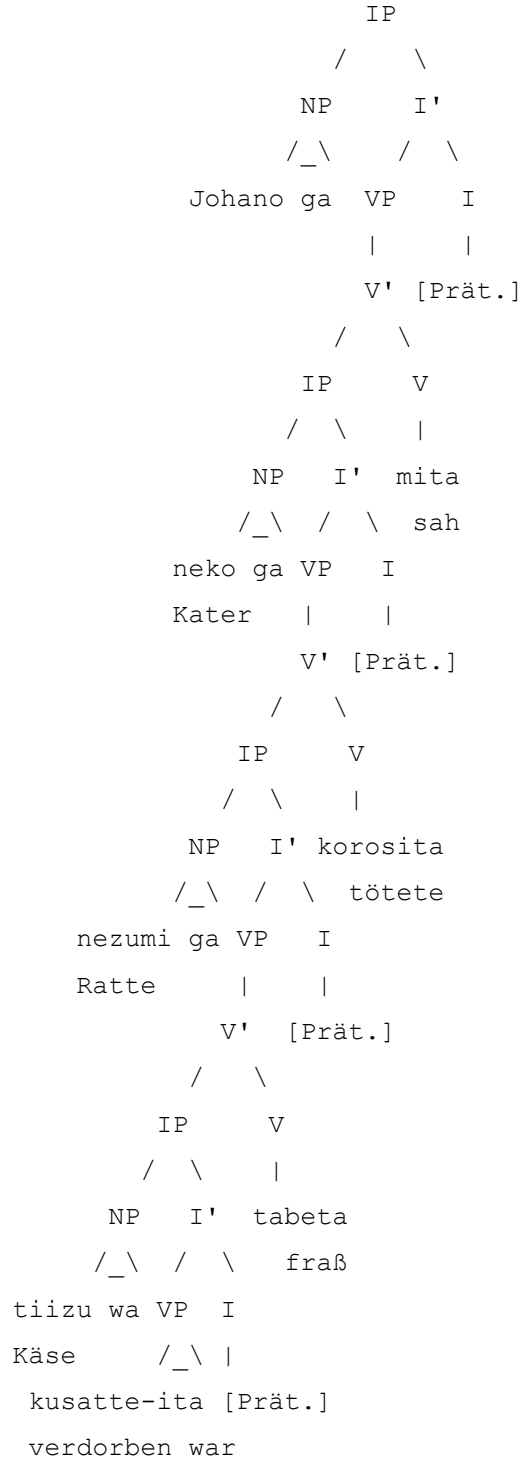


Abbildung 22: Linkseinbettung von Relativsätzen im Japanischen

### 3.4.9 Zusammenfassung: Untersuchung zur Rechts- und Linkseinbettung von Phrasen und Relativsätzen

In der vorliegenden Arbeit wurde die Rechts- und Linkseinbettung von verschiedenartigen Phrasen und Relativsätzen im Esperanto dargestellt. Dazu wurden einige Beispielsätze in X-Bar-Theorie analysiert. Beachtenswert ist, dass Esperanto als Plansprache zwar eine präskriptive Grundlage in Zamenhof (1991<sup>10</sup>) hat, es darin jedoch keine unmittelbaren expliziten Regeln zur Einbettung von Phrasen und Relativsätzen gibt. Etwaige diesbezügliche Regeln oder Empfehlungen können nur mittelbar aus Beispielsätzen hergeleitet werden. Tabelle 9 fasst die Ergebnisse zusammen.

Fazit:

- (a) Subjekt-Nominalphrasen sind überwiegend links in der Flexionsphrase, Objekt-Nominalphrasen überwiegend rechts in der Verbalphrase eingebettet. Daher kann die Reihenfolge Subjekt-Verb-Objekt als Basis-Wortstellung im Esperanto identifiziert werden (vgl. Jansen 2007: 81, Gledhill 2000<sup>2</sup>: 87 und Wells 1978: 41). Bei einem Kopulativverb ist die Linkseinbettung der Nominalphrase praktisch obligatorisch (Wells 1978: 47).
- (b) Der Artikel ist links in der Nominalphrase eingebettet (identifizierte Ausnahme: Rechtseinbettung nach Personennamen vor Ordnungszahl/Adjektiv [Wells 1978: 47]).
- (c) Adjektivphrasen sind überwiegend links in Nominalphrasen eingebettet (Rechtseinbettung vor allem bei Emphase, Euphonie, Korrelativwort, Komplexität [Jansen 2007: 112]).
- (d) Präpositionalphrasen sind überwiegend links in Verbalphrasen eingebettet.
- (e) Keine Vorzugsrichtung bezüglich Einbettung konnte für Adverbphrasen in Verbalphrasen gefunden werden.
- (f) Negationen ohne Kontrast sind (ausnahmslos) links eingebettet; mit Kontrast dagegen überwiegend rechts.
- (g) Relativsätze sind rechts eingebettet.

Einbettung von ...	Linkseinbettungen	Rechtseinbettungen	Bemerkung
Subjekt-Nominalphrase in Flexionsphrase	97% (bei Kopulativverb: 100%; [Wells 1978: 43])	3% (bei Kopulativverb: 0; [Wells 1978: 43])	Jansen (2007: 81) identifiziert Subjekt - Verb - Objekt als Basis-Wortstellung im Esperanto)
Objekt-Nominalphrase (in Verbalphrase)	17% (größte Tendenz bei Pronomen als Subjekt u./o. Objekt)	83%	
Artikel in Nominalphrase	100%	0 (laut Wells 1978: 47 zwischen Personennamen und Ordnungszahl oder Adjektiv möglich)	
Adjektivphrase (in Nominalphrase)	88%	12% (vor allem bei Emphase, Euphonie, Korrelativwort, Komplexität [Jansen 2007: 112])	
Präpositionalphrase (in Verbalphrase)	88%	12%	Diese Angaben beziehen sich nur auf solche Präpositionalphrasen, die ein indirektes Objekt darstellen
Adverbphrase (in Verbalphrase)	50%	50%	
Negation ohne Kontrast	100%	0	
Negation mit Kontrast	38%	62%	
Relativsatz	0	100%	
Zugrundegelegter Korpus: Zamenhof (1887) und Zamenhof (1991 <sup>10</sup> ); die Zahlen (außer bei „Relativsatz“) beruhen auf den in Jansen (2007) gemachten Angaben			

*Tabelle 9: Zusammenfassung der Ergebnisse zur Einbettung von Phrasen und Relativsätzen*

### 3.5 Zahl der grundlegenden und offiziellen Lexeme

In Kapitel 2.5 wurde die Modellierbarkeit der Zahl  $n$  der grundlegenden und offiziellen Wörter gegen  $i$  (Rang in der Offizialisierungsfolge) mit einer Zipf-Funktion gezeigt.

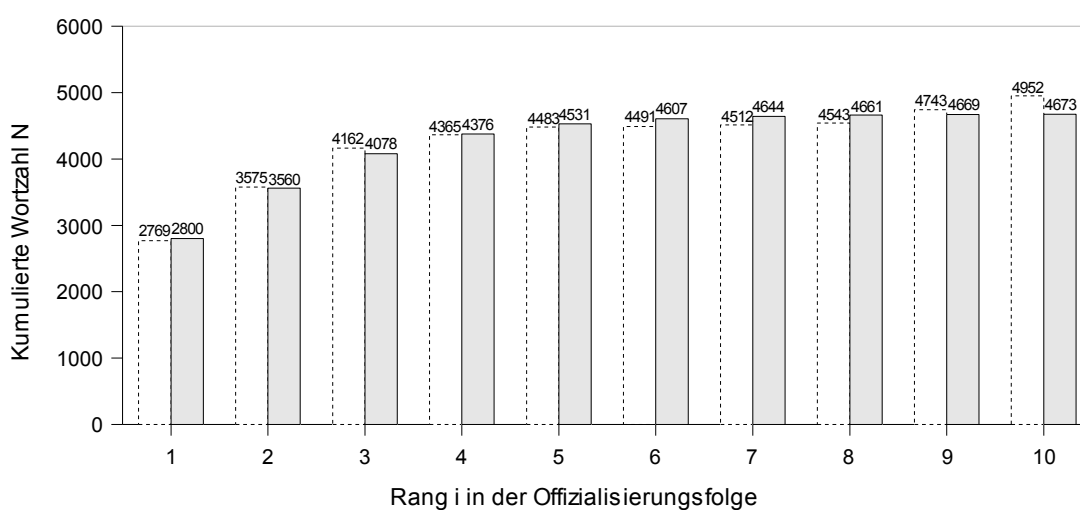
In Anlehnung an die in Strauss, Altmann (2008) angegebene Gleichung (3) (Piotrowski-Gesetz) wurde die kumulierte Wortzahl  $N$  nach Gleichung 10

modelliert, wobei hier anstelle der Zeit der Rang in der Offizialisierungsfolge verwendet wurde.

$$N(i) = \frac{C}{1 + a \exp(-bi)} \quad (10)$$

$a, b, C$ : Modell-Parameter  
 $i$ : Rang in der Offizialisierungsfolge  
 (vgl. Kapitel 2.5)  
 $N$ : kumulierte Wortzahl

Die beste Anpassung (mit einem Bestimmtheitsmaß von 0,96) ergibt sich für  $a = 1,4314$ ;  $b = 0,7580$  und  $C = 4676,5221$  (siehe Abbildung 23).



**Abbildung 23:** Kumulierte Wortzahl  $N$  gegen Rang  $i$  in der Offizialisierungsfolge; gestrichelte farblose Säulen:  $N$  (aus Tabelle 5 ermittelt, dabei nur „Fundamento“ [ $i = 1$ ] und „Beifügungen“ berücksichtigt); graue Säulen:  $N$  (modelliert nach Gleichung 10)

Die kumulierte Wortzahl der grundlegenden und offiziellen Wörter des Esperanto in Abhängigkeit vom Rang in der Offizialisierungsfolge kann also auch in Anlehnung an das Piotrowski-Gesetz beschrieben werden.

### 3.6 Polysemie

Wie bereits in Kapitel 2.6.2 erwähnt, wurden für einige Sprachen gesetzmäßige (quantitative) Zusammenhänge zwischen der Länge von Lexemen und Polysemie gefunden.

Diesbezügliche Untersuchungen für Esperanto konnten vom Autor der vorliegenden Arbeit nicht ausfindig gemacht werden. Daher wurde eine Stichprobe von 100 Lexemen des Esperanto bezüglich Länge und Polysemie

untersucht. Die Stichprobe wurde wie folgt erhoben: Aus Waringhien (2002), einem renommierten einsprachigen Esperanto-Wörterbuch, wurden zu den Polysemien 1, 2, 3, 4, 5, 6, 7 und 8 je zehn bis 17 Wörter zufällig ausgewählt (durch Sichtung des Autors dieser Arbeit beim Durchblättern und ohne maschinelle Unterstützung; pseudozufällige Stichprobe). Bei Polysemien >8 konnten auf diese Weise jedoch insgesamt nur fünf Wörter (im Rahmen einer angebrachten Zeitvorgabe) auffindig gemacht werden. Sie wurden ebenfalls in die Stichprobe aufgenommen, damit ein möglichst großer Umfang der Polysemie erhalten wird. Abbildung 24 zeigt das Ergebnis der Auswertung; Einzelheiten sind in Anhang D, Tabellen 19 bis 22 und Abbildung 44.

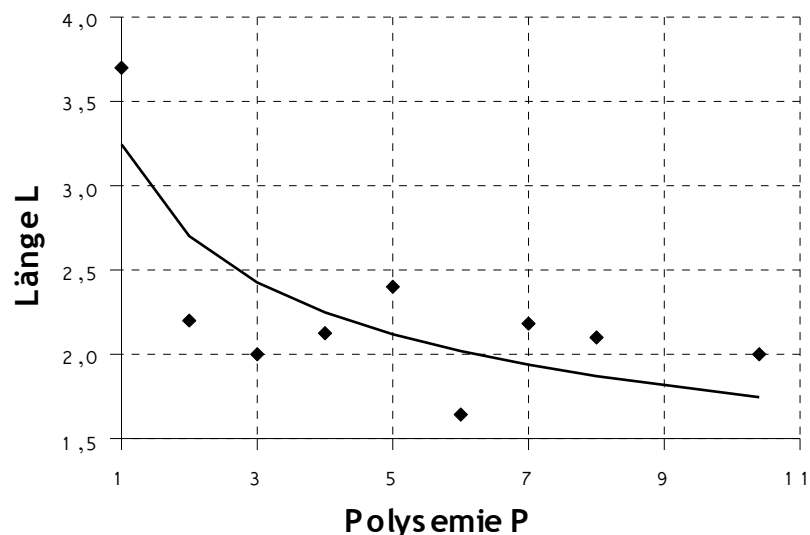


Abbildung 24: Länge von Lexemen als Funktion der Polysemie im Esperanto

Sowohl für Maori (nach Köhler 1999) als auch für Esperanto kann der Zusammenhang zwischen  $L$  und  $P$  durch die Gleichung 11 beschrieben werden.

$$L = C P^b \quad (11)$$

Ein Vergleich der Parameter

für Maori:  $C = 3,3505$ ;  $b = -0,2010$ ; Bestimmtheitsmaß: 0,92

für Esperanto:  $C = 3,2540$ ;  $b = -0,2649$ ; Bestimmtheitsmaß: 0,61

bestätigt die Gesetzmäßigkeit auch für Esperanto.

### 3.7 Zusammenhang zwischen offizieller Basis-Wortwurzelammlung (BRO) und Häufigkeiten nach dem Dietze-Korpus

Um festzustellen, welcher Zusammenhang zwischen der Zugehörigkeit eines Wortes zu einer der BRO-Häufigkeitsgruppen und seiner Häufigkeit im Dietze-Korpus besteht, wurde wie folgt verfahren:

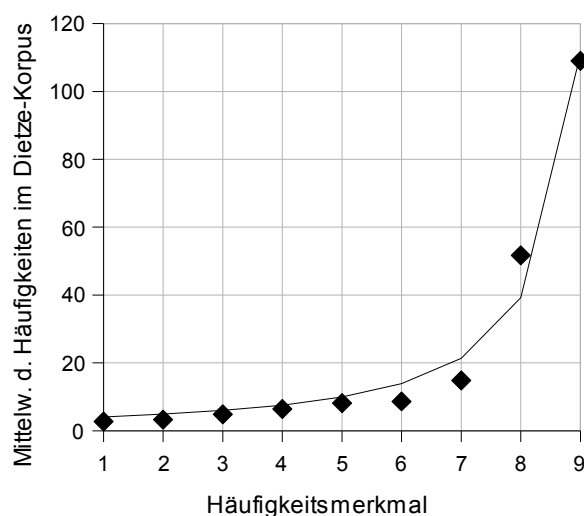
Für jede der neun BRO-Häufigkeitsgruppen wurden die Häufigkeiten ihrer jeweiligen Wörter im Dietze-Korpus festgestellt (anhand von Dietze 1989). BRO-Wörtern, die nicht in Dietze (1989) aufgeführt sind, haben demzufolge eine Häufigkeit von 0, 1 oder 2 im Dietze-Korpus. Ihnen wurde in der vorliegenden Arbeit die Häufigkeit 1 zugewiesen. Anschließend wurde für jede der neun BRO-Häufigkeitsgruppen (mit Häufigkeitsmerkmal *a*) ihr Mittelwert *d* der Häufigkeiten im Dietze-Korpus gebildet (s. Tabelle 10 und Abbildung 25).

Häufigkeitsmerkmal	BRO-Häufigkeitsgruppe	Anzahl der Wörter <sup>A</sup>	Häufigkeit im Dietze-Korpus	
			Mittelwert	Standardabweichung
1	9	755	2,71	5,50
2	8	471	3,25	6,85
3	7	294	4,85	12,19
4	6	241	6,41	14,84
5	5	155	8,17	13,88
6	4	152	8,65	11,04
7	3	121	14,83	23,59
8	2	102	51,70	79,59
9	1	144	108,99	344,61
Summe		2435		

<sup>A</sup> Etwaige Abweichungen zu den in Akademio de Esperanto (Hrsg.) (2007<sup>2</sup>: 31ff.) angegebenen Zahlen lassen sich wie folgt erklären:

- (a) Unberücksichtigt aus BRO1 blieben nicht-lexikalische Morpheme und Morpheme, die nicht in Dietze (1989) aufgeführt sind.
- (b) Wörter, die mehr als eine Form haben, und Wörter, die mehr als einer Wortkategorie angehören, wurden in dieser Tabelle einzeln mitberücksichtigt.

*Tabelle 10: Häufigkeiten der BRO-Wörter im Dietze-Korpus*



*Abbildung 25: Mittelwert  $d$  der Häufigkeiten im Dietze-Korpus gegen Häufigkeitsmerkmal  $a$*

Da das Häufigkeitsmerkmal  $a$  (genau genommen  $10-a$ ) eine Art Rang darstellt, liegt es nahe zu vermuten, dass auch hier das Zipfsche Gesetz gilt. Tatsächlich kann dahingehend die  $d$ - $a$ -Auftragung durch Gleichung 12 genähert beschrieben werden (die durchgezogene Linie in Abbildung 25 stellt dies dar<sup>40</sup>; Bestimmtheitsmaß 0,98):

$$d = 111,0696(10 - a)^{-1,5011} \quad (12)$$

$d$ : Mittelwert der Häufigkeiten der Wörter mit dem Häufigkeitsmerkmal  $a$

Für  $a$  gilt Gleichung 13:

$$a = 10 - (d / 111,0696)^{-1/1,5011} \quad (13)$$

Das Häufigkeitsmerkmal  $a$  eines Wortes, das durch seine Zugehörigkeit zu einer der neun Häufigkeitsgruppen der offiziellen Basis-Wortwurzelammlung (BRO1, ..., BRO9; Akademio de Esperanto [Hrsg.] 2007<sup>2</sup>: 23ff.) definiert ist, korreliert also mit einer Häufigkeitsverteilung, die mit dem Zipfschen Gesetz modellierbar ist. Ein derartiges Ergebnis war nicht unbedingt erwartet worden: Der Dietze-Korpus wurde (nach Dietze 1989) anhand von politischen, kulturpolitischen und populär-wissenschaftlichen Texten zusammengestellt; bei

<sup>40</sup> Die Anpassung erfolgte analog zu der beispielsweise in Anhang A beschriebenen Anpassung (also mit dem Gauß-Newton-Verfahren).



der offiziellen Basis-Wortwurzelammlung gibt es keine thematische Beschränkung der dafür herangezogenen Texte.

## 4 Weiterführende quantitative Studien zum Esperanto: Wortbekanntheit

### 4.1 Annahmen zur Wortbekanntheit

Es soll der Zusammenhang zwischen Bekanntheit und Häufigkeit von Esperanto-Wörtern ermittelt werden. Folgende Annahmen werden gemacht:

- (a) Je häufiger ein Esperanto-Wort vorkommt, desto bekannter ist es.
- (b) Verschiedene Esperanto-Wörter mit gleicher Häufigkeit können signifikant unterschiedlich bekannt sein.
- (c) Andere Faktoren neben der Häufigkeit beeinflussen die Bekanntheit eines Esperanto-Wortes.

### 4.2 Vergleichbare Aufgabenstellung in einem anderen Fachgebiet

Die Annahmen von Kapitel 4.1 haben eine Analogie in der Werkstoffprüftechnik<sup>41</sup> (siehe Tabelle 11).

Charakteristikum	Sprachwissen- schaft	Werkstoffprüftechnik
Herausforderung	(Esperanto-)Wort (das es zu kennen gilt)	Riss (den es zu finden gilt)
Merkmal	Häufigkeit(smerk- mal) des (Esperanto-)Wortes	Größe(nmerkmal) des Risses
Zu ermittelnde Größe	Bekanntheit des (Esperanto-)Wortes	Findungswahrscheinlichkeit des Risses

*Tabelle 11: Analogie zwischen Sprachwissenschaft und Werkstoffprüftechnik*

Zur Bestimmung der Findungswahrscheinlichkeit von Rissen in Abhängigkeit von dem quantitativen Merkmal „Größe des Risses“ wird in der Werkstoffprüfung ein mathematisches Verfahren in spezieller Weise verwendet, das „logistische Regression“ heißt (Berens 1989).

<sup>41</sup> Die Assoziierung zur Werkstoffprüftechnik mag zunächst seltsam erscheinen. Der Autor der vorliegenden Arbeit fand den Zugang zur logistischen Regression durch eine analoge Aufgabenstellung aus der Werkstoffprüftechnik, in der er langjährige Berufserfahrung hat.

Ein solches Verfahren soll nun analog in der vorliegenden Arbeit verwendet werden, um den Bekanntheitsgrad von Esperanto-Wörtern in Abhängigkeit von ihrem Häufigkeitsmerkmal zu ermitteln. Weiterhin soll das Verfahren auch dazu dienen, den Zusammenhang zwischen dem selbst eingeschätzten Esperanto-Anwendungsvermögen und der Esperanto-Lernzeit zu ermitteln. Da bei letztgenannter Anwendung der mathematische Ansatz gleich ist, wird das Prinzip im Folgenden nur anhand der erstgenannten Anwendung erläutert.

### 4.3 Quantitative Beschreibung der Häufigkeit

Für die vorgesehene Anwendung in der Esperantologie ist es erforderlich, die Begriffe „Häufigkeit“ und „Bekanntheit“ in Bezug auf Esperanto-Wörter quantitativ zu beschreiben. Wie in Kapitel 2.7 beschrieben, gibt es bezüglich Esperanto eine fundierte Analyse von Worthäufigkeiten (die BRO, also die „offizielle Basis-Wortwurzelsammlung“ nach Akademio de Esperanto 2007<sup>2</sup>: 23ff.): In der ersten Häufigkeitsgruppe (BRO1) sind die häufigsten Wörter, in der neunten (BRO9) die am wenigsten häufigen (aus den rund 2500 häufigsten Wörtern). Diese bereits vorhandene (und relative) Charakterisierung des Merkmals „Häufigkeit“ wird in der vorliegenden Arbeit zur Definition des Häufigkeitsmerkmals zu Grunde gelegt:

Das Häufigkeitsmerkmal (abgekürzt  $a$ ) von Esperanto-Wörtern aus der Häufigkeitsgruppe BRO9 wird als  $a = 1$  (quantitativ) definiert. Die Häufigkeitsmerkmale der Esperanto-Wörter der weiteren Häufigkeitsgruppen werden folgendermaßen quantitativ definiert: BRO8:  $a = 2$ ; BRO7:  $a = 3$ ; BRO6:  $a = 4$ ; BRO5:  $a = 5$ ; BRO4:  $a = 6$ ; BRO3:  $a = 7$ ; BRO2:  $a = 8$  und BRO1:  $a = 9$  (vgl. auch Tabelle 10).

Diese Art der Charakterisierung der Häufigkeit heranzuziehen, wird damit begründet, dass sie diejenige Charakterisierung ist, die von der Esperanto-Akademie (also den Untersuchungsgegenstand) auf der Grundlage mehrerer Einzeluntersuchungen erarbeitet wurde<sup>42</sup>. Darüber hinaus ist der Verzicht auf die unmittelbare Heranziehung der Häufigkeit (in einem bestimmten Korpus)

---

<sup>42</sup> In Akademio de Esperanto (Hrsg.) (2007<sup>2</sup>: 24) ist diesbezüglich angemerkt, dass die „Überlegenheit von BRO darin besteht, dass sie aus dem Vergleich verschiedener Listen resultiert, die nach sehr unterschiedlichen Kriterien erstellt wurden: um in den ersten Gruppen zu erscheinen, muss eine Wurzel als 'häufig' nach sehr unterschiedlichen Kriterien deklariert worden sein; aber dennoch ... sogar eine häufige Wurzel kann man übersehen“.

kein Nachteil. Grund: Sie kann, wie z. B. beim Dietze-Korpus, dem Häufigkeitsmerkmal *a* zugeordnet werden (siehe Kapitel 3.7, vor allem Tabelle 10).

## 4.4 Quantitative Beschreibung der Bekanntheit

### 4.4.1 Angestrebte Aussageform zur Bekanntheit

Erfahrungsgemäß weiß der Esperanto-Sprecher, dass er ein aufgenommenes Wort entweder kennt oder nicht kennt (Analogie in der Werkstoffprüftechnik: Mit der angewendeten Werkstoffprüftechnik findet der einzelne Prüfer entweder einen Riss oder nicht). Es steht also an, aus dieser auf die Einzelperson bezogene dichotome Beschreibung der Bekanntheit eines Esperanto-Wortes eine statistisch abgesicherte und zwischen 0 und 100% gestufte (gradueller) Aussage zur Bekanntheit zu gewinnen. Ziel ist also eine Aussage in der Art: „Die Esperanto-Wörter der BRO-Häufigkeitsgruppe 7 haben einen Bekanntheitsgrad von 90% bei einem Vertrauensgrad<sup>43</sup> von 95%.“ Mit dem Vertrauensgrad soll der statistischen Sicherheit<sup>44</sup> Rechnung getragen werden.

### 4.4.2 Eine mögliche Vorgehensweise

Eine mögliche Vorgehensweise zur Definition und Ermittlung des Bekanntheitsgrades wird im Folgenden an einem einfachen (fiktiven) Beispiel veranschaulicht:

- (1) Man erstellt einen Fragebogen. Dieser listet mehrere Esperanto-Wörter aus jeder Häufigkeitsgruppe, z. B. 12 Esperanto-Wörter mit dem Häufigkeitsmerkmal 3 nach Kapitel 4.3.
- (2) Man bittet z. B. 6 Befragte, bei jedem der Esperanto-Wörter anzugeben, ob sie es kennen oder nicht.

43 Anstelle des Begriffs „Vertrauensgrad“ werden anderweitig auch die Begriffe „Vertrauensniveau“ und „Konfidenzniveau“ verwendet. In der vorliegenden Arbeit (und z. B. auch in Meier 1999) wird der Begriff „Vertrauensgrad“ verwendet. Deutsches Institut für Normung (Hrsg.) (1982) definiert das „Vertrauensniveau“ folgendermaßen: „Mindestwert  $1-\alpha$  der Wahrscheinlichkeit, der für die Berechnung eines Vertrauensbereichs [...] vorgegeben ist.“ Deutsches Institut für Normung (Hrsg.) (1982) definiert weiterhin den „Vertrauensbereich“ folgendermaßen: „Aus Stichprobenergebnissen berechneter Schätzbereich, der den wahren Wert  $\theta$  des zu schätzenden Parameters auf dem vorgegebenen Vertrauensniveau  $1-\alpha$  einschließt.“ Näheres: Kapitel 4.4.2.

44 Gottwald (2000: 51) bemerkt eine im Laboralltag anzutreffende Verwendung der Begriffe „Vertrauensniveau“ oder „statistische Sicherheit“ für dieselbe Sache.

- (3) Ein Beispiel zu einem möglichen Ergebnis und zur Auswertung: Von den maximal möglichen 72 „Treffern“ (= 6 x 12) gibt es 69mal die Aussage „bekannt“ und dreimal die Aussage „nicht bekannt“. Der Anteil  $p_1$  der gekannten Wörter des betreffenden Häufigkeitsmerkmals wird auf  $69/72 = 96\%$  geschätzt.
- (4) Zur Berücksichtigung der statistischen Sicherheit legt man einen Vertrauensgrad von z. B. 95% zu Grunde. Der Bekanntheitsgrad  $p'_{29;V}$  beim gewählten Vertrauensgrad  $V$  ergibt sich mit folgender Gleichung 14 (in Anlehnung an Kühlmeyer 2001: 100 und Deutsches Institut für Normung [Hrsg.] 1985: 2); warum hier „29“ in der Bezeichnung auftaucht, wird weiter unten deutlich:

$$p'_{29;V} = \frac{x}{x + (n - x + 1) F(100\% - V; f_1; f_2)} \quad (14)$$

- $n$ : Stichprobenumfang; z. B. 72  
 $x$ : Zahl der Gekannt-Aussagen; z. B. 69  
 $V$ : Vertrauensgrad; z. B. 95%  
 $F$ : Quantil der F-Verteilung<sup>45</sup>; im Beispiel 2,01  
 $p'_{29;V}$ : Bekanntheitsgrad bei Vertrauensgrad  $V$ ; im Beispiel: 90%  
 $f_1$ : Freiheitsgrad 1:  $f_1 = 2(n - x + 1)$   
 $f_2$ : Freiheitsgrad 2:  $f_2 = 2x$

Mit dem Vertrauensgrad wird berücksichtigt, dass ja nur eine Stichprobe aus der Grundgesamtheit verwendet wurde. Der Bekanntheitsgrad ist bei sonst gleichen Parametern umso kleiner, desto größer der Vertrauensgrad gewählt wird (in dem obigen Beispiel ergibt sich der Bekanntheitsgrad ohne Vertrauensgrad zu 96%, und mit einem Vertrauensgrad von 95% ergibt sich der Bekanntheitsgrad zu 90%). Man kann sich das Ergebnis wie folgt veranschaulichen: Ein Bekanntheitsgrad von 90% bei einem Vertrauensgrad von 95% bedeutet: Bei mindestens 95 aus 100 Befragungen ergibt sich ein Bekanntheitsgrad von mindestens 90%.

<sup>45</sup> Der Wert dieser Funktion kann in Standard-Tabellenkalkulationsprogrammen mit  $\text{FINV}(1-V; f_1; f_2)$  o. Ä. aufgerufen oder Tabellen entnommen werden.

In der Werkstoffprüftechnik ist die auf Gleichung 14 beruhende Methode<sup>46</sup> zum Ermitteln der zum Bekanntheitsgrad analogen Kenngröße Findungswahrscheinlichkeit unter der Bezeichnung 29/29-Methode<sup>47</sup> bekannt. Grund: Häufig wird von einer Werkstoffprüftechnik gefordert, dass sie Risse einer festgelegten Größe mit einer Findungswahrscheinlichkeit von 90% bei einem Vertrauensgrad von 95% nachweist. Eine derartig geforderte Tauglichkeit der Werkstoffprüftechnik kann folgendermaßen nachgewiesen werden: Man zeigt in einem Experiment, dass es bei 29 Inspektionen an einem Riss (oder auch an mehreren Rissen) der festgelegten Größe 29 Findungen gibt. Mit  $n = 29$  und  $x = 29$  ergibt sich nach Gleichung 14 eine Findungswahrscheinlichkeit von 90% bei einem Vertrauensgrad von 95%. Im Fall einer Nicht-Findung sind für den angestrebten Tauglichkeitsnachweis übrigens 45 Findungen erforderlich ( $n = 46$  und  $x = 45$ ), bei zwei Nicht-Findungen sind 61 Findungen ( $n = 61$  und  $x = 59$ ) erforderlich<sup>48</sup> usw.

Die 29/29-Methode wird in der Prüftechnik jedoch kontrovers diskutiert<sup>49</sup>. Daher wird wie dort auch in der vorliegenden Arbeit eine alternative Vorgehensweise erwogen und angewendet. Es sei in diesem Zusammenhang bereits Folgendes vorweggenommen: In Kapitel 4.10.10 wird es einen Vergleich der 29/29-Methode mit der in Kapitel 4.4.3 beschriebenen alternativen Vorgehensweise geben. Weiterhin wird die Tauglichkeit der alternativen Vorgehensweise statistisch getestet (in Kapitel 4.12).

---

46 Diese Methode beruht auf einem Binomialansatz, der in Air Force Research Laboratory (Hrsg.) (2002) gut erläutert wird. Gleichung 14 ist nur eine andere Darstellung des Binomialansatzes. Übrigens wird ein solcher Binomialansatz auch in Altmann, Lehfeldt (1973: 98) für Wahrscheinlichkeitsaussagen bezüglich Phonemdistributionen herangezogen.

47 Annis (2001): „'29 of 29' method“

48 Diese und weitere Kombinationen sind auch in Air Force Research Laboratory (Hrsg.) (2002) und Natrella (1966) dargestellt.

49 Laut Annis (2001) beruht die 29/29-Methode auf unhaltbaren statistischen Untermauerungen. Laut U. S. Department of Defense (Hrsg.) (1999) (Abschnitt G.3.1) bestehen analytische Schwierigkeiten, laut Air Force Research Laboratory (Hrsg.) (2002) ist die 29/29-Methode u. U. zu konservativ (d. h. es wird u. U. nur ein zu schlechter Wert statistisch abgesichert).

### 4.4.3 Logistische Regression nach Berens

Die von Berens vorgestellte Methode (Berens 1989) einer logistischen Regression<sup>50</sup>, um Findungswahrscheinlichkeiten zu ermitteln, hat folgenden Vorteil gegenüber der in Kapitel 4.4.2 erwähnten 29/29-Methode: In die Ermittlung der Findungswahrscheinlichkeit von Rissen einer bestimmten Größe gehen auch Findungen und Nicht-Findungen von Rissen anderer Größen ein. Weiterhin wird eine fein abgestufte Findungswahrscheinlichkeit für jede Rissgröße gewonnen (die so genannte POD-Kurve; POD: Englisch: „probability of detection“, also Findungswahrscheinlichkeit). In Analogie zu der hier vorliegenden Aufgabenstellung heißt das: In die Ermittlung des Bekanntheitsgrades von Esperanto-Wörtern eines bestimmten Häufigkeitsmerkmals gehen auch Bekanntheit und Nicht-Bekanntheit von Esperanto-Wörtern anderer Häufigkeitsmerkmale ein.

Im Folgenden wird die Berens-Methode in Bezug auf die vorliegende Aufgabenstellung der Sprachwissenschaft erläutert.

Zunächst wird die Annahme von Kapitel 4.1 aufgegriffen: Verschiedene Esperanto-Wörter mit gleicher Häufigkeit können signifikant unterschiedlich bekannt sein. Man geht daher von Folgendem aus: Jedes Esperanto-Wort aus einer Gruppe mit gegebenem Häufigkeitsmerkmal  $a$  hat seine eigene Bekanntheitswahrscheinlichkeit  $p'$ , und die Verteilungsfunktion von  $p'$  sei  $f_a(p')$  (die kleine glockenförmige Kurve in Abbildung 26 in Anlehnung an Berens 1989).

---

<sup>50</sup> Eine allgemeine Einführung in die logistische Regression geben z. B. Tiede (1995) und Urban (1993).

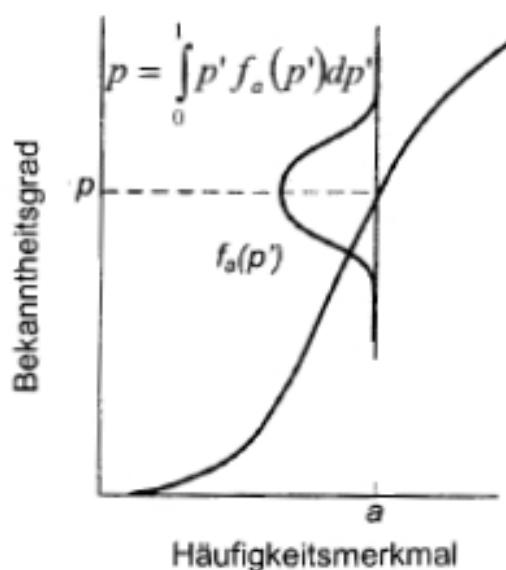


Abbildung 26: Annahmen im Berens-Modell; in der Vorlage (Berens 1989) steht „defect size“ statt „Häufigkeitsmerkmal“ und „probability of detection“ statt „Bekanntheitsgrad“

Das heißt: Die bedingte Wahrscheinlichkeit eines zufällig aus der Gruppe mit dem Häufigkeitsmerkmal  $a$  ausgewählten Esperanto-Wortes, bekannt zu sein, ist  $p'f_a(p')dp'$  (die Bedingtheit besteht darin, dass das Esperanto-Wort die Bekanntheitswahrscheinlichkeit  $p'$  hat). Die unbedingte Wahrscheinlichkeit eines zufällig aus der Gruppe mit dem Häufigkeitsmerkmal  $a$  ausgewählten Esperanto-Wortes, bekannt zu sein, ist das Integral über alle Bekanntheitswahrscheinlichkeiten  $p'$  (siehe die in Abbildung 26 oben gegebene Gleichung) und wird in der vorliegenden Arbeit als Bekanntheitsgrad  $p$  definiert. Der Bekanntheitsgrad  $p$  ist daher der Mittelwert (oder Erwartungswert) der Bekanntheitswahrscheinlichkeiten  $p'$  für Esperanto-Wörter mit einem gegebenen Häufigkeitsmerkmal  $a$ .

Ziel: Aus den Ergebnissen einer Stichproben-Befragung mit Ja/Nein-Aussage zur Bekanntheit der abgefragten Esperanto-Wörter wird der Bekanntheitsgrad  $p$  als Funktion des Häufigkeitsmerkmals unter Berücksichtigung eines festgelegten Vertrauensgrades ermittelt.

Ein mathematisches Modell für den funktionalen Zusammenhang zwischen  $p$  und  $a$  (ohne Berücksichtigung des Vertrauensgrades) ist die in Gleichung 15 gegebene Funktion<sup>51</sup>.

<sup>51</sup> In Berens (1989) wird dieses mathematische Modell gewählt. Grund: Unter den zuvor in



$$p_0(a; \mu; \sigma) = \frac{\exp\left(\frac{\pi}{\sqrt{3}}[\ln(a) - \mu]/\sigma\right)}{1 + \exp\left(\frac{\pi}{\sqrt{3}}[\ln(a) - \mu]/\sigma\right)} \quad (15)$$

$a$ : Häufigkeitsmerkmal<sup>52</sup>

$\mu$ : Mittelwert-Parameter

$\sigma$ : Streuungs-Parameter

$p_0$ : Bekanntheitsgrad ohne Vertrauensgrad im Modell der logistischen Regression

Gleichung 15 stellt eine Näherung der kumulativen logarithmischen Normalverteilung mit den Parametern  $\mu$  und  $\sigma$  dar.

Es ist interessant, dass man Gleichung 15 in eine solche Form bringen kann, die der Gleichung (2) in Köhler, Rapp (2007) recht ähnlich ist:

Division durch den Zähler ergibt:

$$p_0(a; \mu; \sigma) = \frac{1}{1 + \exp\left(\frac{-\pi}{\sqrt{3}}[\ln(a) - \mu]/\sigma\right)} \quad (16)$$

Weiteres Umformen ergibt:

$$p_0(a; \mu; \sigma) = \frac{1}{1 + \exp\left(\frac{\pi}{\sqrt{3}\sigma}\mu\right) \exp\left(\frac{-\pi}{\sqrt{3}\sigma}\ln(a)\right)} \quad (17)$$

Anders geschrieben:

$$p_0(a; \mu; \sigma) = \frac{1}{1 + \exp\left(\frac{\pi}{\sqrt{3}\sigma}\mu\right) a^{\frac{-\pi}{\sqrt{3}\sigma}}} \quad (18)$$

oder:

$$p_0(a; \mu; \sigma) = \frac{1}{1 + A' a^{-B'}} \quad (19)$$

$$A' := \exp\left(\frac{\pi}{\sqrt{3}\sigma}\mu\right), B' := \frac{\pi}{\sqrt{3}\sigma}$$

---

Berens, Hovey (1981) untersuchten sieben mathematischen Modellen stellte es sich als das beste heraus. Es ist analytisch gut zu bearbeiten und ähnlich der Normalverteilung. Anstelle des in Berens (1989) verwendeten Faktors  $\pi/\sqrt{3}$  wird in anderen Ansätzen (z. B. Annis 2001) meist der Faktor 1 verwendet. Das der Gleichung 15 zugrundeliegende Modell wird auch in Cox (1970) genannt.

52 In Gleichung 15 geht der natürliche Logarithmus des Häufigkeitsmerkmals  $a$  ein. Ansätze, nach denen der Logarithmus der Worthäufigkeit (des Englischen) gewissen statistischen Verteilungsfunktionen folgt, gibt es in Herdan (1960), Carroll (1967) und Baayen (2001).

Mit Hilfe von Gleichung 13 lässt sich Gleichung 19 in folgende Form bringen (Gleichung 20):

$$p_0(a; \mu; \sigma) = \frac{1}{1 + A'(10 - (d/111,0696)^{-1/1,5011})^{-B'}} \quad (20)$$

Gleichung 19 weist eine gewisse Ähnlichkeit mit der folgenden Gleichung 21 auf:

$$y = \frac{1}{\frac{1}{A} x^{-B}} \quad (21)$$

Gleichung 21 ist lediglich eine andere Darstellung von Gleichung (2) in Köhler, Rapp (2007):

$$y = Ax^B \quad (22)$$

A: Konstante  
 B: Proportional-Operator  
 x: Worthäufigkeit  
 y: Wortvertrautheit

Diese Ähnlichkeit wird als Beleg dafür gesehen, dass das durch Gleichung 15 beschriebene Modell nicht prinzipiell verschieden von dem Modell in Köhler, Rapp (2007) ist.

Die Größe der Parameter  $\mu$  und  $\sigma$  der Gleichung 15 ist zunächst nicht bekannt. Sie wird auf der Grundlage der Stichproben-Befragungsergebnisse mittels einer Schätzung der maximalen Mutmaßlichkeit<sup>53</sup> (Englisch: „maximum likelihood estimation“, MLE) bestimmt.

Dazu wird jedem einzelnen Stichproben-Befragungsergebnis  $i$  ein Wert  $p_i$  zugeordnet, falls das betreffende Wort bekannt war; ansonsten wird  $1-p_i$  zugeordnet. Der Wert  $p_i$  wird nach Gleichung 15 mit solchen Größen der Parametern  $\mu$  und  $\sigma$  berechnet, die aufgrund einer ersten Näherung aus den Stichproben-Befragungsergebnissen als Startwerte  $\mu_1$  und  $\sigma_1$ <sup>54</sup> angesetzt werden.

<sup>53</sup> Laut Tiede (1995: 19) wird der Begriff Wahrscheinlichkeit „also in den Begriff Likelihood (Mutmaßlichkeit) geändert, weil dem Parameter keine Wahrscheinlichkeit zugeordnet werden kann, wenn x-Werte bereits realisiert sind“ (den x-Werten entsprechen die a-Werte der vorliegenden Arbeit).

<sup>54</sup> In Berens (1989) wird beschrieben, wie man solche Startwerte ermittelt.

Es ergibt sich für die Schätzung der maximalen Mutmaßlichkeit (MLE) folgende Mutmaßlichkeitsfunktion (Gleichung 23):

$$L = \prod_{i=1}^n p_i^{y_i} (1 - p_i)^{1 - y_i} \quad (23)$$

- $i$ : Nummer des einzelnen Stichproben-Befragungsergebnisses
- $n$ : Anzahl der Stichproben-Befragungsergebnisse
- $p$ : Bekanntheitsgrad (zunächst nur nach einer ersten Näherung)
- $y$ : Dichotomie-Merkmal (1: falls Wort bekannt; 0: sonst)
- $L$ : Mutmaßlichkeit (= „likelihood“)

In der Praxis wird (um zu kleine Zahlen zu vermeiden) auf beiden Seiten von Gleichung 23 der natürliche Logarithmus gebildet. Man erhält so Gleichung 24:

$$\ln L = \sum_{i=1}^n y_i \ln(p_i) + (1 - y_i) \ln(1 - p_i) \quad (24)$$

Es sind nun diejenigen Größen  $\mu_0$  und  $\sigma_0$  der Parameter  $\mu$  und  $\sigma$  zu finden, mit denen  $\ln L$  in Gleichung 24 maximal wird. In einem solchen Fall gibt Gleichung 15 die erhaltenen Stichproben-Befragungsergebnisse bestmöglich wieder und stellt die beste Anpassung dar. Dazu werden die Nullstellen der beiden Ableitungsfunktionen von  $\ln L$  nach  $\mu$  und  $\sigma$  numerisch bestimmt (z. B. nach dem modifizierten Newton-Verfahren mit zwei Variablen, Hempel 1996). Mit  $\mu = \mu_0$  und  $\sigma = \sigma_0$  erhält man mittels Gleichung 15 den Bekanntheitsgrad  $p$  als Funktion des Häufigkeitsmerkmals  $a$  – allerdings ohne Berücksichtigung eines Vertrauensgrades.

Um den Vertrauensgrad zu berücksichtigen, greift Berens auf Cheng, Iles (1983) und Cheng, Iles (1988) zurück. Überträgt man die Ergebnisse jener Arbeiten auf die vorliegende Aufgabenstellung, lässt sich der in Gleichung 25 gegebene Zusammenhang zwischen Bekanntheitsgrad  $p_V$  (für den gewählten Vertrauensgrad  $V$ ) und dem Häufigkeitsmerkmal  $a$  aufstellen:

$$p_V = \Phi \left[ \frac{\ln a - \mu_0}{\sigma_0} - \sqrt{\frac{(k_0 \frac{\ln a - \mu_0}{\sigma_0} + k_1)^2}{nk_0 \left(1 + \frac{\sigma_0^2}{k_0 k_2 - k_1^2}\right)}} \right] \quad (25^{55})$$

- $\Phi$ : kumulative Standard-Normalverteilung (also Mittelwert 0; Standardabweichung 1)<sup>56</sup>  
 $a$ : Häufigkeitsmerkmal  
 $V$ : Vertrauensgrad  
 $\mu_0, \sigma_0$ : Parameter, mit denen die Mutmaßlichkeitsfunktion  $\ln L$  (Gleichung 24) ein Maximum hat  
 $n$ : Anzahl der Stichproben-Befragungsergebnisse

Die weiteren Größen sind durch die nachfolgend aufgeführten Gleichungen 28 bis 34 definiert<sup>57</sup>:

$$k_0 = -\frac{\partial^2 \ln L}{\partial \mu^2}(\mu_0, \sigma_0) \frac{\sigma_0^2}{n} = \frac{\pi^2}{3n} \sum_{i=1}^n \frac{h_i}{(h_i + 1)^2} \quad (28)$$

$$k_1 = \frac{\partial^2 \ln L}{\partial \mu \partial \sigma}(\mu_0, \sigma_0) \frac{\sigma_0^2}{n} = -\frac{\pi}{\sqrt{3}n} \sum_{i=1}^n \frac{h_i \ln h_i}{(h_i + 1)^2} \quad (29)$$

$$k_2 = -\frac{\partial^2 \ln L}{\partial \sigma^2}(\mu_0, \sigma_0) \frac{\sigma_0^2}{n} = \frac{1}{n} \sum_{i=1}^n \frac{(\ln h_i)^2 h_i}{(h_i + 1)^2} \quad (30)$$

$$h = \exp\left(\frac{\pi}{\sqrt{3}}[\ln(a) - \mu_0]/\sigma_0\right) \quad (31)$$

55 In Cheng, Iles (1983) ist zusätzlich eine gleichwertige nach  $a$  aufgelöste Gleichung gegeben; daran ist Gleichung 26 angelehnt:

$$a(p_V; V) = \exp\left(\mu_0 + \sigma_0 q + \sigma_0 \frac{b}{N' - k_1 - qk_0}\right) \quad (26)$$

$$b = k_2 + 2k_1 q + k_0 q^2, \quad N' = \sqrt{(k_0 k_2 - k_1^2) \left(\frac{nb}{y} - 1\right)} \quad (27)$$

$q$ : Quantil der Standard-Normalverteilung (also Mittelwert 0 und Streuung 1) mit  $p_V$  als weiterem Parameter; z. B. ist für  $p_V = 90\%$  dieses Quantil  $q = 1,282$ .

56 Der Wert dieser Funktion kann in Standard-Tabellenkalkulationsprogrammen mit NOMVERT([...]; 0; 1, 1) o. Ä. aufgerufen werden; [...] ist das Argument (also das, was in Gleichung 25 in eckigen Klammern steht).

57 In Berens (1989) sind für  $k_1$  und  $k_2$  andere Endergebnisse angegeben als hier in den Gleichungen 29 und 30. Der Autor der vorliegenden Arbeit ist sich jedoch sicher, dass die hier für  $k_1$  und  $k_2$  angegebenen Endergebnisse richtig sind; Herleitung: vgl. Kapitel 4.4.4.

Der Parameter  $\gamma$  hängt von  $n$  und  $V$  ab und ist so definiert, dass damit Gleichung 32<sup>58</sup> erfüllt ist:

$$0,5 \int_{1-\sqrt{\frac{\gamma}{2n}}}^{1+\sqrt{\frac{\gamma}{2n}}} \int_{\sqrt{\frac{\gamma-2n(\sigma-1)^2}{n}}}^{\sqrt{\frac{\gamma-2n(\sigma+1)^2}{n}}} f(\mu)g(\sigma) d\mu d\sigma + 0,5 \int_{1-\sqrt{\frac{\gamma}{2n}}}^{1+\sqrt{\frac{\gamma}{2n}}} g(\sigma) d\sigma = V \quad (32)$$

wobei:

$$f(\mu) = \sqrt{\frac{n}{2\pi}} \exp(-n\mu^2/2) \quad (33)$$

$$g(\sigma) = 2([n-1]\sigma^2/2)^{(n-1)/2} \frac{\exp(-[n-1]\sigma^2/2)}{\sigma \Gamma([n-1]/2)} \quad (34)$$

Der Parameter  $\gamma$  wurde einmalig numerisch mittels der FORTRAN-Routinen D01DAF und D01AHF nach Gleichung 32 speziell für den Vertrauensgrad 95% ermittelt, und zwar für in Einer-Schritten gestufte Werte von  $n$ . Die ermittelten Werte sind in Tabelle 12 gelistet. Darin finden sich die in Cheng, Iles (1988) und Berens (1989) für nur ausgewählte Werte von  $n$  angegebenen  $\gamma$ -Werte wieder (was eine Art Kontrolle darstellt).

---

<sup>58</sup> Diese Gleichung wird in Cheng, Iles (1988) mit Worten beschrieben. In persönlichem Briefverkehr teilte Cheng dem Autor der vorliegenden Arbeit die darin dargestellte mathematische Formulierung mit.

n	$\gamma(n; 95\%)$	n	$\gamma(n; 95\%)$	n	$\gamma(n; 95\%)$
3	5.600	43	5.187	221-243	5.147
4	5.576	44	5.186	244-264	5.146
5	5.519	45	5.185	265-289	5.145
6	5.467	46	5.184	290-319	5.144
7	5.426	47	5.183	320-363	5.143
8	5.393	48	5.182	364-417	5.142
9	5.366	49	5.181	418-535	5.141
10	5.344	50	5.180	536	5.140
11	5.326	51-52	5.179	537-540	5.139
12	5.311	53	5.178	$\geq 541$	5.138
13	5.298	54	5.177		
14	5.287	55-56	5.176		
15	5.277	57	5.175		
16	5.269	58-59	5.174		
17	5.261	60-61	5.173		
18	5.254	62-63	5.172		
19	5.248	64	5.171		
20	5.243	65-67	5.170		
21	5.238	68-69	5.169		
22	5.233	70-72	5.168		
23	5.229	73	5.167		
24	5.225	74-77	5.166		
25	5.222	78-80	5.165		
26	5.219	81-83	5.164		
27	5.216	84-86	5.163		
28	5.213	87-91	5.162		
29	5.210	92-94	5.161		
30	5.208	95-98	5.160		
31	5.206	99-104	5.159		
32	5.204	105-109	5.158		
33	5.202	110-117	5.157		
34	5.200	118-122	5.156		
35	5.198	123-129	5.155		
36	5.196	130-137	5.154		
37	5.195	138-142	5.153		
38	5.194	143-158	5.152		
39	5.192	159-162	5.151		
40	5.191	163-184	5.150		
41	5.189	185-197	5.149		
42	5.188	198-220	5.148		

*Tabelle 12: Parameter  $\gamma$  in Abhängigkeit von  $n$  bei 95% Vertrauensgrad (bestimmt nach Gleichung 32; Hinweis: Dezimalpunkt statt Dezimalkomma verwendet)*

#### 4.4.4 Vertiefung der mathematischen Grundlagen

Im Folgenden wird die Herleitung der Gleichungen 28 bis 30 detailliert beschrieben. Die Parameter  $k_0$ ,  $k_1$  und  $k_2$  werden nach den Gleichungen 28 bis 30 aus den zweiten Ableitungen der Mutmaßlichkeitsfunktion (Gleichung 35) ermittelt.

$$\ln L = \sum_{i=1}^n y_i \ln(p_i) + (1-y_i) \ln(1-p_i) \quad (35)$$

Für die Bestimmung von  $k_0$  wird zunächst die erste Ableitung nach  $\mu$  berechnet:

$$\frac{\partial \ln L}{\partial \mu} = \sum_{i=1}^n \left\{ \frac{y_i}{p_i} \frac{\partial p_i}{\partial \mu} - \frac{(1-y_i)}{(1-p_i)} \frac{\partial p_i}{\partial \mu} \right\} = \sum_{i=1}^n \left\{ \left[ \frac{y_i}{p_i} - \frac{(1-y_i)}{(1-p_i)} \right] \frac{\partial p_i}{\partial \mu} \right\} \quad (36)$$

Es ist

$$\frac{\partial p}{\partial \mu} = \frac{\partial p}{\partial h} \frac{\partial h}{\partial \mu} \quad (37)$$

mit

$$\frac{\partial p}{\partial h} = \frac{-h}{(1+h)^2} + \frac{1}{h+1} = \frac{1}{(h+1)^2} \quad (38)$$

und

$$\frac{\partial h}{\partial \mu} = \frac{-\pi}{\sqrt{3}} \frac{h}{\sigma} \quad (39)$$

und somit

$$\frac{\partial p}{\partial \mu} = \frac{-\pi}{\sqrt{3}} \frac{h}{\sigma} \frac{h}{(h+1)^2} = \frac{-\pi}{\sqrt{3}} \frac{p}{\sigma h+1} \quad (40)$$

Damit ist

$$\frac{\partial \ln L}{\partial \mu} = \sum_{i=1}^n \left\{ \left[ \frac{y_i}{p_i} - \frac{(1-y_i)}{(1-p_i)} \right] \frac{\partial p_i}{\partial \mu} \right\} = \sum_{i=1}^n \left\{ \left[ \frac{y_i}{p_i} - \frac{(1-y_i)}{(1-p_i)} \right] \frac{-\pi}{\sqrt{3}} \frac{p_i}{\sigma h_i+1} \right\} \quad (41)$$

$$\frac{\partial \ln L}{\partial \mu} = \frac{-\pi}{\sqrt{3}} \sum_{i=1}^n \left\{ \left[ \frac{y_i}{h_i+1} - (1-y_i) p_i \right] \right\} \quad (42)$$

weil

$$\frac{p}{1-p} = h \quad \text{und} \quad \frac{h}{h+1} = p \quad (43)$$

Somit gilt

$$\frac{\partial \ln L}{\partial \mu} = \frac{-\pi}{\sqrt{3}} \sum_{i=1}^n \left\{ [y_i(1-p_i) - (1-y_i)p_i] \right\} = \frac{-\pi}{\sqrt{3}} \sum_{i=1}^n \{ y_i - y_i p_i - p_i + y_i p_i \} \quad (44)$$

$$\frac{\partial \ln L}{\partial \mu} = \frac{-\pi}{\sqrt{3}\sigma} \sum_{i=1}^n y_i + \frac{\pi}{\sqrt{3}\sigma} \sum_{i=1}^n p_i \quad (45)$$

$$\frac{\partial^2 \ln L}{\partial \mu^2} = \frac{-\pi^2}{3\sigma^2} \sum_{i=1}^n \frac{p_i}{h_i+1} = \frac{-\pi^2}{3\sigma^2} \sum_{i=1}^n \frac{h_i}{(h_i+1)^2} \quad (46)$$

Damit ergibt sich

$$k_0 = -\frac{\partial^2 \ln L}{\partial \mu^2}(\mu_0, \sigma_0) \frac{\sigma_0^2}{n} = \frac{\pi^2}{3n} \sum_{i=1}^n \frac{h_i}{(h_i+1)^2} \quad (47)$$

Für die Bestimmung von  $k_2$  wird zunächst die erste Ableitung nach  $\sigma$  berechnet:

$$\frac{\partial \ln L}{\partial \sigma} = \sum_{i=1}^n \left\{ \frac{y_i}{p_i} \frac{\partial p_i}{\partial \sigma} - \frac{(1-y_i)}{(1-p_i)} \frac{\partial p_i}{\partial \sigma} \right\} = \sum_{i=1}^n \left\{ \left[ \frac{y_i}{p_i} - \frac{(1-y_i)}{(1-p_i)} \right] \frac{\partial p_i}{\partial \sigma} \right\} \quad (48)$$

Es ist

$$\frac{\partial p}{\partial \sigma} = \frac{\partial p}{\partial h} \frac{\partial h}{\partial \sigma} \quad (49)$$

mit

$$\frac{\partial h}{\partial \sigma} = \frac{-\pi}{\sqrt{3}\sigma^2} (\ln a - \mu) h \quad (50)$$

und somit

$$\frac{\partial p}{\partial \sigma} = \frac{-\pi}{\sqrt{3}\sigma^2} (\ln a - \mu) \frac{h}{(h+1)^2} = \frac{-\pi}{\sqrt{3}\sigma^2} (\ln a - \mu) \frac{p}{h+1} \quad (51)$$

Damit ist

$$\frac{\partial \ln L}{\partial \sigma} = \sum_{i=1}^n \left\{ \left[ \frac{y_i}{p_i} - \frac{(1-y_i)}{(1-p_i)} \right] \frac{\partial p_i}{\partial \sigma} \right\} = \sum_{i=1}^n \left\{ \left[ \frac{y_i}{p_i} - \frac{(1-y_i)}{(1-p_i)} \right] \frac{-\pi}{\sqrt{3}\sigma^2} (\ln a_i - \mu) \frac{p_i}{h_i+1} \right\} \quad (52)$$

$$\frac{\partial \ln L}{\partial \sigma} = \frac{-\pi}{\sqrt{3}\sigma^2} \sum_{i=1}^n \left\{ \left[ \frac{y_i}{p_i} - \frac{(1-y_i)}{(1-p_i)} \right] (\ln a_i - \mu) \frac{p_i}{h_i+1} \right\} \quad (53)$$

$$\frac{\partial \ln L}{\partial \sigma} = \frac{-\pi}{\sqrt{3}\sigma^2} \sum_{i=1}^n \left\{ (\ln a_i - \mu) \left[ \frac{y_i}{h_i+1} - \frac{(1-y_i)}{(h_i+1)} \frac{p_i}{1-p_i} \right] \right\} \quad (54)$$

$$\frac{\partial \ln L}{\partial \sigma} = \frac{-\pi}{\sqrt{3}\sigma^2} \sum_{i=1}^n \left\{ (\ln a_i - \mu) \left[ \frac{y_i - (1-y_i)h_i}{h_i+1} \right] \right\} \quad (55)$$

$$\frac{\partial \ln L}{\partial \sigma} = \frac{-\pi}{\sqrt{3}\sigma^2} \sum_{i=1}^n \left\{ (\ln a_i - \mu) \left[ \frac{y_i(1+h_i) - h_i}{h_i+1} \right] \right\} \quad (56)$$

$$\frac{\partial \ln L}{\partial \sigma} = \frac{-\pi}{\sqrt{3}\sigma^2} \sum_{i=1}^n \left\{ (\ln a_i - \mu) (y_i - p_i) \right\} \quad (57)$$



$$\frac{\partial \ln L}{\partial \sigma} = \frac{-\pi}{\sqrt{3}\sigma^2} \sum_{i=1}^n \{[(\ln a_i - \mu) y_i - (\ln a_i - \mu) p_i]\} \quad (58)$$

Es ist

$$\frac{\partial^2 \ln L}{\partial \sigma^2} = \frac{2\pi}{\sqrt{3}\sigma^3} \sum_{i=1}^n \{[(\ln a_i - \mu) y_i - (\ln a_i - \mu) p_i]\} - \frac{\pi}{\sqrt{3}\sigma^2} \frac{\pi}{\sqrt{3}\sigma^2} \sum_{i=1}^n \{(\ln a_i - \mu)^2 \frac{h_i}{(h_i+1)^2}\} \quad (59)$$

$$\frac{\partial^2 \ln L}{\partial \sigma^2} = \frac{\pi}{\sqrt{3}\sigma^2} \sum_{i=1}^n \{(\ln a_i - \mu) [(y_i - p_i) \frac{2}{\sigma} - \frac{\pi}{\sqrt{3}\sigma^2} (\ln a_i - \mu) \frac{h_i}{(h_i+1)^2}]\} \quad (60)$$

$$\frac{\partial^2 \ln L}{\partial \sigma^2} = \frac{\pi}{\sqrt{3}\sigma^2} \sum_{i=1}^n (\ln a_i - \mu) (y_i - p_i) \frac{2}{\sigma} - \frac{\pi}{\sqrt{3}\sigma^2} \frac{\pi}{\sqrt{3}\sigma^2} \sum_{i=1}^n (\ln a_i - \mu)^2 \frac{h_i}{(h_i+1)^2} \quad (61)$$

Wegen Gleichung 57 kann dies auch so geschrieben werden:

$$\frac{\partial^2 \ln L}{\partial \sigma^2} = \frac{\pi}{\sqrt{3}\sigma^2} \frac{\pi}{\sqrt{3}\sigma^2} \frac{2}{\sigma} \frac{\partial \ln L}{\partial \sigma} \Big|_{\mu=\mu_0; \sigma=\sigma_0} - \frac{\pi}{\sqrt{3}\sigma^2} \frac{\pi}{\sqrt{3}\sigma^2} \sum_{i=1}^n (\ln a_i - \mu)^2 \frac{h_i}{(h_i+1)^2} \quad (62)$$

Der erste Summand ist null, weil die darin vorkommende partielle erste Ableitung beim Maximum ( $\mu = \mu_0$ ;  $\sigma = \mu_0$ ) null ist. Weiterhin ist

$$\ln a - \mu = \frac{\sqrt{3}\sigma}{\pi} \ln h \quad \text{wegen } h = \exp\left(\frac{\pi}{\sqrt{3}} \frac{(\ln a - \mu)}{\sigma}\right) \quad (63)$$

Somit ist

$$\frac{\partial^2 \ln L}{\partial \sigma^2} = \frac{-\pi^2}{3\sigma^4} \sum_{i=1}^n \left\{ \frac{3\sigma^2}{\pi^2} \frac{h_i (\ln h_i)^2}{(h_i+1)^2} \right\} = -\frac{1}{\sigma^2} \sum_{i=1}^n \left\{ \frac{h_i (\ln h_i)^2}{(h_i+1)^2} \right\} \quad (64)$$

Damit ergibt sich

$$k_2 = -\frac{\partial^2 \ln L}{\partial \sigma^2} (\mu_0, \sigma_0) \frac{\sigma_0^2}{n} = \frac{1}{n} \sum_{i=1}^n \frac{(\ln h_i)^2 h_i}{(h_i+1)^2} \quad (65)$$

Für die Bestimmung von  $k_1$  wird zunächst Folgendes berechnet (ausgehend von Gleichung 45):

$$\frac{\partial^2 \ln L}{\partial \sigma \partial \mu} = \frac{\partial}{\partial \sigma} \left[ \frac{\pi}{\sqrt{3}\sigma} \sum_{i=1}^n \{p_i - y_i\} \right] \quad (66)$$

$$\frac{\partial^2 \ln L}{\partial \sigma \partial \mu} = \frac{-\pi}{\sqrt{3}\sigma^2} \sum_{i=1}^n \{p_i - y_i\} - \frac{\pi}{\sqrt{3}\sigma} \frac{\pi}{\sqrt{3}\sigma^2} \sum_{i=1}^n \left\{ (\ln a_i - \mu) \frac{h_i}{(h_i+1)^2} \right\} \quad (67)$$

$$\frac{\partial^2 \ln L}{\partial \sigma \partial \mu} = \frac{-\pi}{\sqrt{3}\sigma^2} \left[ \sum_{i=1}^n \{p_i - y_i\} + \frac{\pi}{\sqrt{3}\sigma} \sum_{i=1}^n \left\{ (\ln a_i - \mu) \frac{h_i}{(h_i+1)^2} \right\} \right] \quad (68)$$

$$\frac{\partial^2 \ln L}{\partial \sigma \partial \mu} = \frac{-\pi}{\sqrt{3}\sigma^2} \left[ \sum_{i=1}^n \left\{ \frac{\sqrt{3}\sigma}{\pi} \frac{\partial \ln L}{\partial \mu} \Big|_{\mu=\mu_0; \sigma=\sigma_0} \right\} + \frac{\pi}{\sqrt{3}\sigma} \sum_{i=1}^n \left\{ (\ln a_i - \mu) \frac{h_i}{(h_i+1)^2} \right\} \right] \quad (69)$$

Der erste Summand in der eckigen Klammer ist null, weil die darin vorkommende partielle erste Ableitung beim Maximum ( $\mu = \mu_0$ ;  $\sigma = \mu_0$ ) null ist. Weiterhin ist

$$\ln a - \mu = \frac{\sqrt{3}\sigma}{\pi} \ln h \quad \text{wegen } h = \exp\left(\frac{\pi}{\sqrt{3}} \frac{(\ln a - \mu)}{\sigma}\right) \quad (70)$$

Somit ist

$$\frac{\partial^2 \ln L}{\partial \sigma \partial \mu} = \frac{-\pi}{\sqrt{3}\sigma^2} \sum_{i=1}^n \left\{ \frac{h_i \ln h_i}{(h_i + 1)^2} \right\} \quad (71)$$

Damit ergibt sich

$$k_1 = \frac{\partial^2 \ln L}{\partial \mu \partial \sigma}(\mu_0, \sigma_0) \frac{\sigma_0^2}{n} = -\frac{\pi}{\sqrt{3}n} \sum_{i=1}^n \frac{h_i \ln h_i}{(h_i + 1)^2} \quad (72)$$

#### 4.4.5 Finden der maximalen Mutmaßlichkeit

Um die maximale Mutmaßlichkeit zu finden, ist das Maximum der Mutmaßlichkeitsfunktion (Gleichung 24) zu bestimmen. Dazu werden mit dem modifizierten Newton-Verfahren die Nullstellen der folgenden Ableitungsfunktionen ermittelt:

$$\frac{\partial \ln L}{\partial \mu} = \frac{-\pi}{\sqrt{3}\sigma} \sum_{i=1}^n y_i + \frac{\pi}{\sqrt{3}\sigma} \sum_{i=1}^n p_i \quad (73)$$

(dies entspricht Gleichung 45)

und

$$\frac{\partial \ln L}{\partial \sigma} = \frac{-\pi}{\sqrt{3}\sigma^2} \sum_{i=1}^n \{[(\ln a_i - \mu) y_i - (\ln a_i - \mu) p_i]\} \quad (74)$$

(dies entspricht Gleichung 58)

Für die Nullstellensuche wird Gleichung 73 wie folgt vereinfacht:

$$\sum_{i=1}^n y_i - \sum_{i=1}^n p_i =: f \quad (75)$$

Und Gleichung 74 wird

$$\frac{\partial \ln L}{\partial \sigma} = \frac{-\pi}{\sqrt{3}\sigma^2} \left[ \sum_{i=1}^n \{y_i \ln a_i - p_i \ln a_i\} + \mu \sum_{i=1}^n \{p_i - y_i\} \right] \quad (76)$$

Der zweite Summand in der eckigen Klammer ist null (vgl. Gleichung 69). Damit ergibt sich:

$$\frac{\partial \ln L}{\partial \sigma} = \frac{-\pi}{\sqrt{3} \sigma^2} \sum_{i=1}^n \{y_i \ln a_i - p_i \ln a_i\} =: g \quad (77)$$

Allgemein lässt sich (unter gewissen Voraussetzungen) die verbesserte genäherte Nullstelle  $x_{k+1}$  einer Funktion  $q(x)$ , deren Nullstelle  $x_k$  genähert bekannt ist, wie folgt mit dem modifizierten Newton-Verfahren bestimmen, einem iterativen Näherungsverfahren (Hempel 1996: 36):

$$x_{k+1} = x_k - m_k \frac{q(x)}{q'(x)} \quad (78)$$

$$m_k = \frac{1}{1 - \frac{q(x_k) q''(x_k)}{q'(x_k)^2}} \quad (79)$$

(die hochgestellten Striche bedeuten Ableitungen nach  $x$ )

Die Übertragung der Gleichungen 78 und 79 auf Gleichung 75 führt auf folgende Gleichung zur Bestimmung einer verbesserten genäherten Nullstelle:

$$\left( \frac{\mu_{k+1}}{\sigma_{k+1}} \right) = \left( \frac{\mu_k}{\sigma_k} \right) - m_k \frac{f(\mu_k, \sigma_k) \cdot f'(\mu_k, \sigma_k)}{|f'(\mu_k, \sigma_k)|^2} \quad (80)$$

$$m_k = \frac{1}{1 - \frac{f \cdot [(f'' \cdot f') \cdot f']}{|f'|^4}} \quad (81)$$

Für die Funktion  $g$  gibt es analog zu den Gleichungen 80 und 81 zwei (hier nicht explizit aufgeführte) Gleichungen mit  $g$  statt  $f$ . Es sind also folgende Ableitungen nötig:

$$\frac{\partial f}{\partial \mu} = \frac{\pi}{\sqrt{3} \sigma} \sum_{i=1}^n \frac{p_i}{h_i + 1} =: fb \quad (82)$$

$$\frac{\partial f}{\partial \sigma} = \frac{1}{\sigma} \sum_{i=1}^n \frac{p_i \ln h_i}{h_i + 1} =: fs \quad (83)$$

$$\frac{\partial^2 f}{\partial \mu^2} = \frac{-\pi^2}{3 \sigma^2} \sum_{i=1}^n \frac{h_i (1 - h_i)}{(h_i + 1)^3} =: fbb \quad (84)$$

$$\frac{\partial^2 f}{\partial \sigma^2} = \frac{-1}{\sigma^2} \sum_{i=1}^n \frac{h_i \ln h_i}{(h_i + 1)^2} \left[ 2 + \ln h_i - \frac{2 h_i \ln h_i}{h_i + 1} \right] =: fss \quad (85)$$

$$\frac{\partial^2 f}{\partial \mu \partial \sigma} = \frac{-\pi}{\sqrt{3} \sigma^2} \sum_{i=1}^n \frac{h_i}{(h_i + 1)^2} \left[ 1 + \ln h_i - \frac{2 h_i \ln h_i}{h_i + 1} \right] =: fbs \quad (86)$$

$$\frac{\partial^2 f}{\partial \sigma \partial \mu} = \frac{\partial^2 f}{\partial \mu \partial \sigma} =: fbs \quad (87)$$

$$\frac{\partial \underline{g}}{\partial \mu} = \frac{\pi}{\sqrt{3}\sigma} \sum_{i=1}^n \frac{h_i \ln a_i}{(h_i+1)^2} =: gb \quad (88)$$

$$\frac{\partial \underline{g}}{\partial \sigma} = \frac{1}{\sigma} \sum_{i=1}^n \frac{p_i \ln a_i \ln h_i}{h_i+1} =: gs \quad (89)$$

$$\frac{\partial^2 \underline{g}}{\partial \mu^2} = \frac{-\pi^2}{3\sigma^2} \sum_{i=1}^n \frac{p_i \ln a_i (1-h_i)}{(h_i+1)^2} =: gbb \quad (90)$$

$$\frac{\partial^2 \underline{g}}{\partial \sigma^2} = \frac{-1}{\sigma^2} \sum_{i=1}^n \frac{h_i \ln a_i \ln h_i}{(h_i+1)^2} \left[ 2 + \ln h_i - \frac{2 h_i \ln h_i}{h_i+1} \right] =: gss \quad (91)$$

$$\frac{\partial^2 \underline{g}}{\partial \mu \partial \sigma} = \frac{-\pi}{\sqrt{3}\sigma^2} \sum_{i=1}^n \frac{h_i \ln a_i}{(h_i+1)^2} \left[ 1 + \ln h_i - \frac{2 h_i \ln h_i}{h_i+1} \right] =: gbs \quad (92)$$

$$\frac{\partial^2 \underline{g}}{\partial \sigma \partial \mu} = \frac{\partial^2 \underline{g}}{\partial \mu \partial \sigma} =: gsb \quad (93)$$

Die in den Gleichungen 82 bis 93 gewählten Kurzbezeichnungen *fb*, *fs*, *gb* usw. entsprechen den Variablennamen im Makro ModNewton (Anhang G). Damit ergeben sich die in den Gleichungen 80 und 81 benötigten Größen wie folgt:

$$f' = \begin{pmatrix} fb \\ fs \end{pmatrix} \quad (94)$$

$$f'' = \begin{pmatrix} fbb & fbs \\ fbs & fss \end{pmatrix} \quad (95)$$

(und analog für *g'* und *g''*).

## 4.5 Entwurf einer Befragung und Ausführung

Für die Berechnung des Bekanntheitsgrades  $p_v$  in Abhängigkeit vom Häufigkeitsmerkmal *a* nach Gleichung 25 mussten Daten beschafft werden. Dazu wurden fünf Fragebogen A bis E in Esperanto entworfen. Darin wird jeder Befragte nach Namen, Netzpostadresse, Muttersprache und Beginn des Esperanto-Lernens gefragt sowie, ob er sich selbst so einschätzt, dass er Esperanto ausreichend gut verwenden kann („ja“) oder nicht („nein“). Weiterhin enthält jeder Fragebogen eine Liste von 100 Esperanto-Wörtern<sup>59</sup>. Zu jedem dieser Wörter wird der Befragte gebeten anzugeben, ob er es kennt („ja“) oder nicht („nein“). Der Befragte kann außerdem einen Kommentar im Fragebogen eingeben. Abbildung 27 zeigt beispielhaft den Fragebogen A.

<sup>59</sup> Berens (1989), Air Force Research Laboratory (Hrsg.) (2002) und Georgiou (2006: 7) empfehlen einen Stichprobenumfang von mindestens 60; Grundlage: im Wesentlichen Erfahrungen (laut Air Force Research Laboratory [Hrsg.] 2002).

Die 100 abgefragten Wörter eines jeden Fragebogens wurden als pseudozufällige Stichproben aus den Häufigkeitsgruppen BRO1 bis BRO9 (Kapitel 4.3) gewählt. Eine weitgehend verhältnismäßige Gewichtung nach der Mächtigkeit der einzelnen Häufigkeitsgruppen wurde berücksichtigt: Beispielsweise stammen 102 Wörter aus BRO2 und 755 Wörter aus BRO9; daher sind in den 100 Wörtern des Fragebogens 5 Wörter aus BRO2 und 30 Wörter aus BRO9. Weiterhin wurde innerhalb einer jeden Häufigkeitsgruppe entsprechend der Zugehörigkeit gemäß den Wortkategorien Substantiv, Adjektiv, Verb und Sonstiges gewichtet: Beispielsweise sind von den 30 Wörtern, die die Häufigkeitsgruppe BRO9 repräsentieren, 21 Substantive, 2 Adjektive, 6 Verben und ein „Sonstiges“-Wort (in BRO9 sind 518 Substantive (darunter drei doppelte Formen<sup>60</sup>), 54 Adjektive, 164 Verben und 19 „Sonstige“-Wörter).

Es wurde, wenn möglich, vermieden, dass ein und dasselbe Wort in mehr als einem Fragebogen vorkam. Außerdem wurden nur gebunden auftretende nicht-lexikalische Morpheme vermieden. Die Fragebogen wurden im Internet veröffentlicht.

---

<sup>60</sup> In Akademio de Esperanto (Hrsg.) (2007<sup>2</sup>: 43) ist als Anzahl 516 angegeben (mit Hinweis auf drei doppelte Formen). Für die Ermittlung des Zusammenhangs zwischen offizieller Basis-Wortwurzelsammlung (BRO) und Häufigkeiten nach dem Dietze-Korpus (Kapitel 3.7) und auch für die Gewichtung der für die Fragebogen ausgewählten Wortzahlen werden die doppelten Formen jedoch in der vorliegenden Arbeit mitberücksichtigt, sodass hier die Anzahl von 518 zugrundegelegt ist (vgl. auch Tabelle 10, Fußnote A).

## Enketo

### Por eltrovi la konatecgradon de Esperanto-vortoj

Andreas Kück, Rechtenfleth, Germanujo

#### Enkonduko

Per tiu ĉi enketo la aŭtoro celas eltrovi la konatecgradon de Esperanto-vortoj laŭ speciala statistika metodo. Tiucele via helpo estas bezonata. Bonvole plenigu jenan enketilon kaj ĝin sendu al la aŭtoro, premante la butonon "Sendu" en la fino de tiu ĉi interretaĵo. Viaj nomo kaj retpoŝtadreso estos nek pludonataj nek publikigataj. La aŭtoro anticipe dankas.

#### Enketilo

Mia nomo estas ..... Mia retpoŝtadreso estas .....  
 Mia gepatra lingvo estas ..... Mi komencis lerni Esperanton en la jaro .....

Kiu el la du sekvaj eblaĵoj trafas laŭ vi pri via uzo de Esperanto?

- Mi scipovas uzi Esperanton sufiĉe bone.  
 Mi ne scipovas uzi Esperanton sufiĉe bone.

Jen cent Esperanto-vortoj. Bonvole marku apud ĉiu el ili, ĉu vi scias ĝian signifon (aŭ "jes" aŭ "ne").

Bonvole ne rigardu en vortaron kaj ne uzu la helpon de io alia aŭ de iu alia, sed decidu tuj laŭ via unua impresio. (Vi ne bezonas doni tradukojn aŭ difinojn.)

alia	<input type="radio"/> jes <input type="radio"/> ne	formulo	<input type="radio"/> jes <input type="radio"/> ne	leporo	<input type="radio"/> jes <input type="radio"/> ne	reno	<input type="radio"/> jes <input type="radio"/> ne	ies	<input type="radio"/> jes <input type="radio"/> ne
periodo	<input type="radio"/> jes <input type="radio"/> ne	generacio	<input type="radio"/> jes <input type="radio"/> ne	sango	<input type="radio"/> jes <input type="radio"/> ne	densa	<input type="radio"/> jes <input type="radio"/> ne	kaso	<input type="radio"/> jes <input type="radio"/> ne
komponi	<input type="radio"/> jes <input type="radio"/> ne	dikti	<input type="radio"/> jes <input type="radio"/> ne	saĝa	<input type="radio"/> jes <input type="radio"/> ne	krepusko	<input type="radio"/> jes <input type="radio"/> ne	angulo	<input type="radio"/> jes <input type="radio"/> ne
forĝi	<input type="radio"/> jes <input type="radio"/> ne	foiro	<input type="radio"/> jes <input type="radio"/> ne	balono	<input type="radio"/> jes <input type="radio"/> ne	jungo	<input type="radio"/> jes <input type="radio"/> ne	ledo	<input type="radio"/> jes <input type="radio"/> ne
komitato	<input type="radio"/> jes <input type="radio"/> ne	bulbo	<input type="radio"/> jes <input type="radio"/> ne	pinto	<input type="radio"/> jes <input type="radio"/> ne	araneo	<input type="radio"/> jes <input type="radio"/> ne	acida	<input type="radio"/> jes <input type="radio"/> ne
remi	<input type="radio"/> jes <input type="radio"/> ne	ekvatoro	<input type="radio"/> jes <input type="radio"/> ne	celo	<input type="radio"/> jes <input type="radio"/> ne	petoli	<input type="radio"/> jes <input type="radio"/> ne	Julio	<input type="radio"/> jes <input type="radio"/> ne
neniu	<input type="radio"/> jes <input type="radio"/> ne	ok	<input type="radio"/> jes <input type="radio"/> ne	demandi	<input type="radio"/> jes <input type="radio"/> ne	segi	<input type="radio"/> jes <input type="radio"/> ne	malvo	<input type="radio"/> jes <input type="radio"/> ne
vendredo	<input type="radio"/> jes <input type="radio"/> ne	agiti	<input type="radio"/> jes <input type="radio"/> ne	kerno	<input type="radio"/> jes <input type="radio"/> ne	ŝuti	<input type="radio"/> jes <input type="radio"/> ne	lasi	<input type="radio"/> jes <input type="radio"/> ne
diveni	<input type="radio"/> jes <input type="radio"/> ne	pozicio	<input type="radio"/> jes <input type="radio"/> ne	ĉia	<input type="radio"/> jes <input type="radio"/> ne	kuri	<input type="radio"/> jes <input type="radio"/> ne	parko	<input type="radio"/> jes <input type="radio"/> ne
singularo	<input type="radio"/> jes <input type="radio"/> ne	ĉies	<input type="radio"/> jes <input type="radio"/> ne	resto	<input type="radio"/> jes <input type="radio"/> ne	publiko	<input type="radio"/> jes <input type="radio"/> ne	anonci	<input type="radio"/> jes <input type="radio"/> ne
avantaĝo	<input type="radio"/> jes <input type="radio"/> ne	ĉelo	<input type="radio"/> jes <input type="radio"/> ne	informi	<input type="radio"/> jes <input type="radio"/> ne	preskaŭ	<input type="radio"/> jes <input type="radio"/> ne	turismo	<input type="radio"/> jes <input type="radio"/> ne
pizo	<input type="radio"/> jes <input type="radio"/> ne	ankaŭ	<input type="radio"/> jes <input type="radio"/> ne	daktilo	<input type="radio"/> jes <input type="radio"/> ne	identa	<input type="radio"/> jes <input type="radio"/> ne	masto	<input type="radio"/> jes <input type="radio"/> ne
kategorio	<input type="radio"/> jes <input type="radio"/> ne	ŝpini	<input type="radio"/> jes <input type="radio"/> ne	komisii	<input type="radio"/> jes <input type="radio"/> ne	kaduka	<input type="radio"/> jes <input type="radio"/> ne	premio	<input type="radio"/> jes <input type="radio"/> ne
nul	<input type="radio"/> jes <input type="radio"/> ne	tago	<input type="radio"/> jes <input type="radio"/> ne	kreto	<input type="radio"/> jes <input type="radio"/> ne	minuto	<input type="radio"/> jes <input type="radio"/> ne	direktoro	<input type="radio"/> jes <input type="radio"/> ne
domo	<input type="radio"/> jes <input type="radio"/> ne	imagi	<input type="radio"/> jes <input type="radio"/> ne	malgraŭ	<input type="radio"/> jes <input type="radio"/> ne	klini	<input type="radio"/> jes <input type="radio"/> ne	cerbo	<input type="radio"/> jes <input type="radio"/> ne
ju	<input type="radio"/> jes <input type="radio"/> ne	zoologio	<input type="radio"/> jes <input type="radio"/> ne	ĵaŭdo	<input type="radio"/> jes <input type="radio"/> ne	amaso	<input type="radio"/> jes <input type="radio"/> ne	vendi	<input type="radio"/> jes <input type="radio"/> ne
kompromiti	<input type="radio"/> jes <input type="radio"/> ne	okcidento	<input type="radio"/> jes <input type="radio"/> ne	teksto	<input type="radio"/> jes <input type="radio"/> ne	urbo	<input type="radio"/> jes <input type="radio"/> ne	vulturo	<input type="radio"/> jes <input type="radio"/> ne
ezoko	<input type="radio"/> jes <input type="radio"/> ne	kajero	<input type="radio"/> jes <input type="radio"/> ne	apenaŭ	<input type="radio"/> jes <input type="radio"/> ne	brasiko	<input type="radio"/> jes <input type="radio"/> ne	dependi	<input type="radio"/> jes <input type="radio"/> ne
vera	<input type="radio"/> jes <input type="radio"/> ne	aparta	<input type="radio"/> jes <input type="radio"/> ne	sago	<input type="radio"/> jes <input type="radio"/> ne	atingi	<input type="radio"/> jes <input type="radio"/> ne	agento	<input type="radio"/> jes <input type="radio"/> ne
fantazio	<input type="radio"/> jes <input type="radio"/> ne	perlo	<input type="radio"/> jes <input type="radio"/> ne	kial	<input type="radio"/> jes <input type="radio"/> ne	kolektiva	<input type="radio"/> jes <input type="radio"/> ne	ami	<input type="radio"/> jes <input type="radio"/> ne

Mi havas jenan komenton:

Abbildung 27: Fragebogen A (Quellcode: Anhang E)

Hier die Übersetzung der Textpassagen der einführenden und abschließenden Textpassagen:

## „Befragung

### Um den Bekanntheitsgrad von Esperanto-Wörtern herauszufinden

Andreas Kück, Rechtenfleth, Deutschland

#### Einleitung

Mit dieser Befragung zieht der Autor darauf ab, den Bekanntheitsgrad von Esperanto-Wörtern nach einer speziellen statistischen Methode herauszufinden. Dazu wird Ihre Hilfe benötigt. Bitte füllen Sie folgenden Fragebogen aus und senden Sie ihn an den Autor, indem Sie die Schaltfläche „Sendu“ {Senden} am Ende dieser Netzseite drücken. Ihr Name und Ihre Netzpostadresse werden weder weitergegeben noch veröffentlicht. Der Autor dankt im Voraus.

#### Fragebogen

Ich heiÙe \_\_\_\_\_ Meine Netzpostadresse ist \_\_\_\_\_

Meine Muttersprache ist \_\_\_\_\_ Ich begann, Esperanto im Jahr ... zu lernen.

Welche der beiden folgenden Möglichkeiten trifft Ihrer Meinung nach für Ihre Anwendung des Esperanto zu?

Ich kann Esperanto ausreichend gut anwenden.

Ich kann Esperanto nicht ausreichend gut anwenden.

Hier sind hundert Esperanto-Wörter. Bitte kennzeichnen Sie neben jeden davon, ob Sie seine Bedeutung kennen (entweder „ja“ oder „nein“). Bitte sehen Sie nicht in ein Wörterbuch und verwenden Sie nicht die Hilfe von etwas oder jemand anderem, sondern entscheiden Sie gleich nach Ihrem ersten Eindruck. (Sie brauchen keine Übersetzungen oder Definitionen zu geben.)

{Es folgen die 100 Wörter.}

Ich habe folgenden Kommentar:

Senden“

Von den weiteren Fragebogen B bis E werden im Folgenden nur die abgefragten Wörter gezeigt (Abbildungen 28 bis 31), weil die einführenden und abschließenden Textpassagen gleich denen von Fragebogen A (Abbildung 27) sind.

antaŭ	<input type="radio"/> jes <input type="radio"/> ne	rafano	<input type="radio"/> jes <input type="radio"/> ne	ĝoji	<input type="radio"/> jes <input type="radio"/> ne	arko	<input type="radio"/> jes <input type="radio"/> ne	pugno	<input type="radio"/> jes <input type="radio"/> ne
tiom	<input type="radio"/> jes <input type="radio"/> ne	strigo	<input type="radio"/> jes <input type="radio"/> ne	kompreni	<input type="radio"/> jes <input type="radio"/> ne	ial	<input type="radio"/> jes <input type="radio"/> ne	nepo	<input type="radio"/> jes <input type="radio"/> ne
kio	<input type="radio"/> jes <input type="radio"/> ne	gvardio	<input type="radio"/> jes <input type="radio"/> ne	nenies	<input type="radio"/> jes <input type="radio"/> ne	honesta	<input type="radio"/> jes <input type="radio"/> ne	pinglo	<input type="radio"/> jes <input type="radio"/> ne
da	<input type="radio"/> jes <input type="radio"/> ne	mielo	<input type="radio"/> jes <input type="radio"/> ne	avida	<input type="radio"/> jes <input type="radio"/> ne	konduti	<input type="radio"/> jes <input type="radio"/> ne	azeno	<input type="radio"/> jes <input type="radio"/> ne
mil	<input type="radio"/> jes <input type="radio"/> ne	lado	<input type="radio"/> jes <input type="radio"/> ne	pretendi	<input type="radio"/> jes <input type="radio"/> ne	ŝiri	<input type="radio"/> jes <input type="radio"/> ne	sako	<input type="radio"/> jes <input type="radio"/> ne
tamen	<input type="radio"/> jes <input type="radio"/> ne	kolumno	<input type="radio"/> jes <input type="radio"/> ne	skui	<input type="radio"/> jes <input type="radio"/> ne	balai	<input type="radio"/> jes <input type="radio"/> ne	merkredo	<input type="radio"/> jes <input type="radio"/> ne
jen	<input type="radio"/> jes <input type="radio"/> ne	dando	<input type="radio"/> jes <input type="radio"/> ne	boji	<input type="radio"/> jes <input type="radio"/> ne	frunto	<input type="radio"/> jes <input type="radio"/> ne	serioza	<input type="radio"/> jes <input type="radio"/> ne
ĉiom	<input type="radio"/> jes <input type="radio"/> ne	angilo	<input type="radio"/> jes <input type="radio"/> ne	ĵuri	<input type="radio"/> jes <input type="radio"/> ne	kuvo	<input type="radio"/> jes <input type="radio"/> ne	sukcesi	<input type="radio"/> jes <input type="radio"/> ne
solena	<input type="radio"/> jes <input type="radio"/> ne	cirklo	<input type="radio"/> jes <input type="radio"/> ne	maĉi	<input type="radio"/> jes <input type="radio"/> ne	pilko	<input type="radio"/> jes <input type="radio"/> ne	kompati	<input type="radio"/> jes <input type="radio"/> ne
graveda	<input type="radio"/> jes <input type="radio"/> ne	hepato	<input type="radio"/> jes <input type="radio"/> ne	fulgo	<input type="radio"/> jes <input type="radio"/> ne	nevo	<input type="radio"/> jes <input type="radio"/> ne	vagono	<input type="radio"/> jes <input type="radio"/> ne
liveri	<input type="radio"/> jes <input type="radio"/> ne	bendo	<input type="radio"/> jes <input type="radio"/> ne	krado	<input type="radio"/> jes <input type="radio"/> ne	vango	<input type="radio"/> jes <input type="radio"/> ne	nubo	<input type="radio"/> jes <input type="radio"/> ne
sarki	<input type="radio"/> jes <input type="radio"/> ne	etapo	<input type="radio"/> jes <input type="radio"/> ne	rango	<input type="radio"/> jes <input type="radio"/> ne	adiaŭ	<input type="radio"/> jes <input type="radio"/> ne	doloro	<input type="radio"/> jes <input type="radio"/> ne
indulgi	<input type="radio"/> jes <input type="radio"/> ne	funelo	<input type="radio"/> jes <input type="radio"/> ne	kastelo	<input type="radio"/> jes <input type="radio"/> ne	klera	<input type="radio"/> jes <input type="radio"/> ne	nul	<input type="radio"/> jes <input type="radio"/> ne
mencii	<input type="radio"/> jes <input type="radio"/> ne	gramo	<input type="radio"/> jes <input type="radio"/> ne	kubuto	<input type="radio"/> jes <input type="radio"/> ne	nutri	<input type="radio"/> jes <input type="radio"/> ne	komuna	<input type="radio"/> jes <input type="radio"/> ne
muĝi	<input type="radio"/> jes <input type="radio"/> ne	konkordo	<input type="radio"/> jes <input type="radio"/> ne	sojlo	<input type="radio"/> jes <input type="radio"/> ne	koncerni	<input type="radio"/> jes <input type="radio"/> ne	frapi	<input type="radio"/> jes <input type="radio"/> ne
vagi	<input type="radio"/> jes <input type="radio"/> ne	limako	<input type="radio"/> jes <input type="radio"/> ne	vulpo	<input type="radio"/> jes <input type="radio"/> ne	frandi	<input type="radio"/> jes <input type="radio"/> ne	valori	<input type="radio"/> jes <input type="radio"/> ne
trupo	<input type="radio"/> jes <input type="radio"/> ne	karmo	<input type="radio"/> jes <input type="radio"/> ne	poezio	<input type="radio"/> jes <input type="radio"/> ne	aventuro	<input type="radio"/> jes <input type="radio"/> ne	okulo	<input type="radio"/> jes <input type="radio"/> ne
ribo	<input type="radio"/> jes <input type="radio"/> ne	sinjoro	<input type="radio"/> jes <input type="radio"/> ne	revuo	<input type="radio"/> jes <input type="radio"/> ne	pajlo	<input type="radio"/> jes <input type="radio"/> ne	afabla	<input type="radio"/> jes <input type="radio"/> ne
sekcio	<input type="radio"/> jes <input type="radio"/> ne	mondo	<input type="radio"/> jes <input type="radio"/> ne	komo	<input type="radio"/> jes <input type="radio"/> ne	ŝultro	<input type="radio"/> jes <input type="radio"/> ne	manki	<input type="radio"/> jes <input type="radio"/> ne
profeto	<input type="radio"/> jes <input type="radio"/> ne	multa	<input type="radio"/> jes <input type="radio"/> ne	kupro	<input type="radio"/> jes <input type="radio"/> ne	vato	<input type="radio"/> jes <input type="radio"/> ne	kosti	<input type="radio"/> jes <input type="radio"/> ne

Abbildung 28: Fragebogen B

anstataŭ	<input type="radio"/> jes <input type="radio"/> ne	inundi	<input type="radio"/> jes <input type="radio"/> ne	papilio	<input type="radio"/> jes <input type="radio"/> ne	malgraŭ	<input type="radio"/> jes <input type="radio"/> ne	minus	<input type="radio"/> jes <input type="radio"/> ne
ĉiam	<input type="radio"/> jes <input type="radio"/> ne	palpi	<input type="radio"/> jes <input type="radio"/> ne	sino	<input type="radio"/> jes <input type="radio"/> ne	milda	<input type="radio"/> jes <input type="radio"/> ne	sobra	<input type="radio"/> jes <input type="radio"/> ne
kiam	<input type="radio"/> jes <input type="radio"/> ne	tinti	<input type="radio"/> jes <input type="radio"/> ne	rimeno	<input type="radio"/> jes <input type="radio"/> ne	ornami	<input type="radio"/> jes <input type="radio"/> ne	fleksi	<input type="radio"/> jes <input type="radio"/> ne
kelo	<input type="radio"/> jes <input type="radio"/> ne	venĝi	<input type="radio"/> jes <input type="radio"/> ne	pendolo	<input type="radio"/> jes <input type="radio"/> ne	masoni	<input type="radio"/> jes <input type="radio"/> ne	perfidi	<input type="radio"/> jes <input type="radio"/> ne
kun	<input type="radio"/> jes <input type="radio"/> ne	aserti	<input type="radio"/> jes <input type="radio"/> ne	pruno	<input type="radio"/> jes <input type="radio"/> ne	esplori	<input type="radio"/> jes <input type="radio"/> ne	regali	<input type="radio"/> jes <input type="radio"/> ne
cent	<input type="radio"/> jes <input type="radio"/> ne	grumbli	<input type="radio"/> jes <input type="radio"/> ne	tubero	<input type="radio"/> jes <input type="radio"/> ne	violo	<input type="radio"/> jes <input type="radio"/> ne	konjekti	<input type="radio"/> jes <input type="radio"/> ne
aŭ	<input type="radio"/> jes <input type="radio"/> ne	grego	<input type="radio"/> jes <input type="radio"/> ne	tragedio	<input type="radio"/> jes <input type="radio"/> ne	rizo	<input type="radio"/> jes <input type="radio"/> ne	dungi	<input type="radio"/> jes <input type="radio"/> ne
ekster	<input type="radio"/> jes <input type="radio"/> ne	kolombo	<input type="radio"/> jes <input type="radio"/> ne	bati	<input type="radio"/> jes <input type="radio"/> ne	ŝvito	<input type="radio"/> jes <input type="radio"/> ne	abelo	<input type="radio"/> jes <input type="radio"/> ne
koro	<input type="radio"/> jes <input type="radio"/> ne	figo	<input type="radio"/> jes <input type="radio"/> ne	onklo	<input type="radio"/> jes <input type="radio"/> ne	ĉifono	<input type="radio"/> jes <input type="radio"/> ne	ekvilibro	<input type="radio"/> jes <input type="radio"/> ne
ĉambro	<input type="radio"/> jes <input type="radio"/> ne	lino	<input type="radio"/> jes <input type="radio"/> ne	neĝo	<input type="radio"/> jes <input type="radio"/> ne	nebulo	<input type="radio"/> jes <input type="radio"/> ne	drapo	<input type="radio"/> jes <input type="radio"/> ne
lasta	<input type="radio"/> jes <input type="radio"/> ne	beto	<input type="radio"/> jes <input type="radio"/> ne	fero	<input type="radio"/> jes <input type="radio"/> ne	sezono	<input type="radio"/> jes <input type="radio"/> ne	hajlo	<input type="radio"/> jes <input type="radio"/> ne
helpi	<input type="radio"/> jes <input type="radio"/> ne	alarmo	<input type="radio"/> jes <input type="radio"/> ne	nul	<input type="radio"/> jes <input type="radio"/> ne	tolo	<input type="radio"/> jes <input type="radio"/> ne	trezoro	<input type="radio"/> jes <input type="radio"/> ne
fari	<input type="radio"/> jes <input type="radio"/> ne	vertico	<input type="radio"/> jes <input type="radio"/> ne	sana	<input type="radio"/> jes <input type="radio"/> ne	ĉiel	<input type="radio"/> jes <input type="radio"/> ne	trikoto	<input type="radio"/> jes <input type="radio"/> ne
akvo	<input type="radio"/> jes <input type="radio"/> ne	palto	<input type="radio"/> jes <input type="radio"/> ne	tuŝi	<input type="radio"/> jes <input type="radio"/> ne	stranga	<input type="radio"/> jes <input type="radio"/> ne	pato	<input type="radio"/> jes <input type="radio"/> ne
tablo	<input type="radio"/> jes <input type="radio"/> ne	plago	<input type="radio"/> jes <input type="radio"/> ne	krajono	<input type="radio"/> jes <input type="radio"/> ne	flegi	<input type="radio"/> jes <input type="radio"/> ne	kliento	<input type="radio"/> jes <input type="radio"/> ne
pura	<input type="radio"/> jes <input type="radio"/> ne	splito	<input type="radio"/> jes <input type="radio"/> ne	ofico	<input type="radio"/> jes <input type="radio"/> ne	posedi	<input type="radio"/> jes <input type="radio"/> ne	kolono	<input type="radio"/> jes <input type="radio"/> ne
permesi	<input type="radio"/> jes <input type="radio"/> ne	romano	<input type="radio"/> jes <input type="radio"/> ne	stelo	<input type="radio"/> jes <input type="radio"/> ne	tremi	<input type="radio"/> jes <input type="radio"/> ne	proceso	<input type="radio"/> jes <input type="radio"/> ne
po	<input type="radio"/> jes <input type="radio"/> ne	obstaklo	<input type="radio"/> jes <input type="radio"/> ne	bruna	<input type="radio"/> jes <input type="radio"/> ne	ŝuo	<input type="radio"/> jes <input type="radio"/> ne	kruĉo	<input type="radio"/> jes <input type="radio"/> ne
pia	<input type="radio"/> jes <input type="radio"/> ne	tradicio	<input type="radio"/> jes <input type="radio"/> ne	pendi	<input type="radio"/> jes <input type="radio"/> ne	poto	<input type="radio"/> jes <input type="radio"/> ne	ripo	<input type="radio"/> jes <input type="radio"/> ne
amara	<input type="radio"/> jes <input type="radio"/> ne	spinaco	<input type="radio"/> jes <input type="radio"/> ne	ŝlosi	<input type="radio"/> jes <input type="radio"/> ne	rado	<input type="radio"/> jes <input type="radio"/> ne	legomo	<input type="radio"/> jes <input type="radio"/> ne

Abbildung 29: Fragebogen C



kontraŭ	<input type="radio"/> jes <input type="radio"/> ne	menso	<input type="radio"/> jes <input type="radio"/> ne	preni	<input type="radio"/> jes <input type="radio"/> ne	enigmo	<input type="radio"/> jes <input type="radio"/> ne	brusto	<input type="radio"/> jes <input type="radio"/> ne
naŭ	<input type="radio"/> jes <input type="radio"/> ne	stadiono	<input type="radio"/> jes <input type="radio"/> ne	ridi	<input type="radio"/> jes <input type="radio"/> ne	bildo	<input type="radio"/> jes <input type="radio"/> ne	ordinara	<input type="radio"/> jes <input type="radio"/> ne
ie	<input type="radio"/> jes <input type="radio"/> ne	stano	<input type="radio"/> jes <input type="radio"/> ne	plus	<input type="radio"/> jes <input type="radio"/> ne	monto	<input type="radio"/> jes <input type="radio"/> ne	promeni	<input type="radio"/> jes <input type="radio"/> ne
dum	<input type="radio"/> jes <input type="radio"/> ne	talio	<input type="radio"/> jes <input type="radio"/> ne	ĵaluzo	<input type="radio"/> jes <input type="radio"/> ne	dolĉa	<input type="radio"/> jes <input type="radio"/> ne	puŝi	<input type="radio"/> jes <input type="radio"/> ne
el	<input type="radio"/> jes <input type="radio"/> ne	vipo	<input type="radio"/> jes <input type="radio"/> ne	ŝovi	<input type="radio"/> jes <input type="radio"/> ne	dormi	<input type="radio"/> jes <input type="radio"/> ne	stacio	<input type="radio"/> jes <input type="radio"/> ne
kaj	<input type="radio"/> jes <input type="radio"/> ne	paliso	<input type="radio"/> jes <input type="radio"/> ne	velki	<input type="radio"/> jes <input type="radio"/> ne	rigardi	<input type="radio"/> jes <input type="radio"/> ne	besto	<input type="radio"/> jes <input type="radio"/> ne
inter	<input type="radio"/> jes <input type="radio"/> ne	sledo	<input type="radio"/> jes <input type="radio"/> ne	beni	<input type="radio"/> jes <input type="radio"/> ne	des	<input type="radio"/> jes <input type="radio"/> ne	avo	<input type="radio"/> jes <input type="radio"/> ne
ties	<input type="radio"/> jes <input type="radio"/> ne	pasero	<input type="radio"/> jes <input type="radio"/> ne	oscedi	<input type="radio"/> jes <input type="radio"/> ne	lama	<input type="radio"/> jes <input type="radio"/> ne	nul	<input type="radio"/> jes <input type="radio"/> ne
vanta	<input type="radio"/> jes <input type="radio"/> ne	volumeno	<input type="radio"/> jes <input type="radio"/> ne	ŝmiri	<input type="radio"/> jes <input type="radio"/> ne	peki	<input type="radio"/> jes <input type="radio"/> ne	blua	<input type="radio"/> jes <input type="radio"/> ne
tenera	<input type="radio"/> jes <input type="radio"/> ne	sofo	<input type="radio"/> jes <input type="radio"/> ne	traktoro	<input type="radio"/> jes <input type="radio"/> ne	distri	<input type="radio"/> jes <input type="radio"/> ne	ludi	<input type="radio"/> jes <input type="radio"/> ne
inspiri	<input type="radio"/> jes <input type="radio"/> ne	puso	<input type="radio"/> jes <input type="radio"/> ne	velo	<input type="radio"/> jes <input type="radio"/> ne	surprizi	<input type="radio"/> jes <input type="radio"/> ne	iel	<input type="radio"/> jes <input type="radio"/> ne
luli	<input type="radio"/> jes <input type="radio"/> ne	kaĝo	<input type="radio"/> jes <input type="radio"/> ne	frambo	<input type="radio"/> jes <input type="radio"/> ne	larmo	<input type="radio"/> jes <input type="radio"/> ne	ĝentila	<input type="radio"/> jes <input type="radio"/> ne
indulgi	<input type="radio"/> jes <input type="radio"/> ne	maleolo	<input type="radio"/> jes <input type="radio"/> ne	veŝto	<input type="radio"/> jes <input type="radio"/> ne	fojno	<input type="radio"/> jes <input type="radio"/> ne	konfesi	<input type="radio"/> jes <input type="radio"/> ne
sorbi	<input type="radio"/> jes <input type="radio"/> ne	dianto	<input type="radio"/> jes <input type="radio"/> ne	felo	<input type="radio"/> jes <input type="radio"/> ne	ĉapo	<input type="radio"/> jes <input type="radio"/> ne	pafi	<input type="radio"/> jes <input type="radio"/> ne
prosperi	<input type="radio"/> jes <input type="radio"/> ne	apro	<input type="radio"/> jes <input type="radio"/> ne	profesio	<input type="radio"/> jes <input type="radio"/> ne	tombo	<input type="radio"/> jes <input type="radio"/> ne	gvidi	<input type="radio"/> jes <input type="radio"/> ne
dresi	<input type="radio"/> jes <input type="radio"/> ne	groto	<input type="radio"/> jes <input type="radio"/> ne	listo	<input type="radio"/> jes <input type="radio"/> ne	deklivo	<input type="radio"/> jes <input type="radio"/> ne	kapro	<input type="radio"/> jes <input type="radio"/> ne
ekspuati	<input type="radio"/> jes <input type="radio"/> ne	himno	<input type="radio"/> jes <input type="radio"/> ne	ŝelo	<input type="radio"/> jes <input type="radio"/> ne	nevo	<input type="radio"/> jes <input type="radio"/> ne	kondiĉo	<input type="radio"/> jes <input type="radio"/> ne
atomo	<input type="radio"/> jes <input type="radio"/> ne	kapo	<input type="radio"/> jes <input type="radio"/> ne	tendo	<input type="radio"/> jes <input type="radio"/> ne	vosto	<input type="radio"/> jes <input type="radio"/> ne	oriento	<input type="radio"/> jes <input type="radio"/> ne
korvo	<input type="radio"/> jes <input type="radio"/> ne	tempo	<input type="radio"/> jes <input type="radio"/> ne	kordo	<input type="radio"/> jes <input type="radio"/> ne	mastro	<input type="radio"/> jes <input type="radio"/> ne	pomo	<input type="radio"/> jes <input type="radio"/> ne
licenso	<input type="radio"/> jes <input type="radio"/> ne	sama	<input type="radio"/> jes <input type="radio"/> ne	petrolo	<input type="radio"/> jes <input type="radio"/> ne	truo	<input type="radio"/> jes <input type="radio"/> ne	valo	<input type="radio"/> jes <input type="radio"/> ne

Abbildung 30: Fragebogen D

ĉirkaŭ	<input type="radio"/> jes <input type="radio"/> ne	paraŝuto	<input type="radio"/> jes <input type="radio"/> ne	sidi	<input type="radio"/> jes <input type="radio"/> ne	apetito	<input type="radio"/> jes <input type="radio"/> ne	plaĉi	<input type="radio"/> jes <input type="radio"/> ne
tri	<input type="radio"/> jes <input type="radio"/> ne	perono	<input type="radio"/> jes <input type="radio"/> ne	fini	<input type="radio"/> jes <input type="radio"/> ne	iel	<input type="radio"/> jes <input type="radio"/> ne	bani	<input type="radio"/> jes <input type="radio"/> ne
tia	<input type="radio"/> jes <input type="radio"/> ne	planedo	<input type="radio"/> jes <input type="radio"/> ne	nenies	<input type="radio"/> jes <input type="radio"/> ne	vigla	<input type="radio"/> jes <input type="radio"/> ne	des	<input type="radio"/> jes <input type="radio"/> ne
neniam	<input type="radio"/> jes <input type="radio"/> ne	pruno	<input type="radio"/> jes <input type="radio"/> ne	kruela	<input type="radio"/> jes <input type="radio"/> ne	rezulti	<input type="radio"/> jes <input type="radio"/> ne	freneza	<input type="radio"/> jes <input type="radio"/> ne
plu	<input type="radio"/> jes <input type="radio"/> ne	nilono	<input type="radio"/> jes <input type="radio"/> ne	melki	<input type="radio"/> jes <input type="radio"/> ne	kisi	<input type="radio"/> jes <input type="radio"/> ne	sterni	<input type="radio"/> jes <input type="radio"/> ne
se	<input type="radio"/> jes <input type="radio"/> ne	ironio	<input type="radio"/> jes <input type="radio"/> ne	sorĉi	<input type="radio"/> jes <input type="radio"/> ne	minaci	<input type="radio"/> jes <input type="radio"/> ne	kroĉi	<input type="radio"/> jes <input type="radio"/> ne
super	<input type="radio"/> jes <input type="radio"/> ne	kranio	<input type="radio"/> jes <input type="radio"/> ne	barakti	<input type="radio"/> jes <input type="radio"/> ne	graso	<input type="radio"/> jes <input type="radio"/> ne	grimpi	<input type="radio"/> jes <input type="radio"/> ne
ĉial	<input type="radio"/> jes <input type="radio"/> ne	logiko	<input type="radio"/> jes <input type="radio"/> ne	konfidi	<input type="radio"/> jes <input type="radio"/> ne	sorto	<input type="radio"/> jes <input type="radio"/> ne	tritiko	<input type="radio"/> jes <input type="radio"/> ne
malica	<input type="radio"/> jes <input type="radio"/> ne	fajenco	<input type="radio"/> jes <input type="radio"/> ne	ravi	<input type="radio"/> jes <input type="radio"/> ne	rando	<input type="radio"/> jes <input type="radio"/> ne	sablo	<input type="radio"/> jes <input type="radio"/> ne
abrupta	<input type="radio"/> jes <input type="radio"/> ne	galoŝo	<input type="radio"/> jes <input type="radio"/> ne	tono	<input type="radio"/> jes <input type="radio"/> ne	najbaro	<input type="radio"/> jes <input type="radio"/> ne	briko	<input type="radio"/> jes <input type="radio"/> ne
klopodi	<input type="radio"/> jes <input type="radio"/> ne	lako	<input type="radio"/> jes <input type="radio"/> ne	vergo	<input type="radio"/> jes <input type="radio"/> ne	sudo	<input type="radio"/> jes <input type="radio"/> ne	haveno	<input type="radio"/> jes <input type="radio"/> ne
ŝpruci	<input type="radio"/> jes <input type="radio"/> ne	kurento	<input type="radio"/> jes <input type="radio"/> ne	pelto	<input type="radio"/> jes <input type="radio"/> ne	lito	<input type="radio"/> jes <input type="radio"/> ne	martelo	<input type="radio"/> jes <input type="radio"/> ne
dorloti	<input type="radio"/> jes <input type="radio"/> ne	artiko	<input type="radio"/> jes <input type="radio"/> ne	simio	<input type="radio"/> jes <input type="radio"/> ne	rajto	<input type="radio"/> jes <input type="radio"/> ne	haŭto	<input type="radio"/> jes <input type="radio"/> ne
aludi	<input type="radio"/> jes <input type="radio"/> ne	aveno	<input type="radio"/> jes <input type="radio"/> ne	gloro	<input type="radio"/> jes <input type="radio"/> ne	nigra	<input type="radio"/> jes <input type="radio"/> ne	kravato	<input type="radio"/> jes <input type="radio"/> ne
oponi	<input type="radio"/> jes <input type="radio"/> ne	cikonio	<input type="radio"/> jes <input type="radio"/> ne	tajloro	<input type="radio"/> jes <input type="radio"/> ne	konsili	<input type="radio"/> jes <input type="radio"/> ne	sukero	<input type="radio"/> jes <input type="radio"/> ne
provoki	<input type="radio"/> jes <input type="radio"/> ne	cepo	<input type="radio"/> jes <input type="radio"/> ne	peniko	<input type="radio"/> jes <input type="radio"/> ne	tiri	<input type="radio"/> jes <input type="radio"/> ne	linio	<input type="radio"/> jes <input type="radio"/> ne
princo	<input type="radio"/> jes <input type="radio"/> ne	specimeno	<input type="radio"/> jes <input type="radio"/> ne	plumbo	<input type="radio"/> jes <input type="radio"/> ne	ŝtupo	<input type="radio"/> jes <input type="radio"/> ne	glaso	<input type="radio"/> jes <input type="radio"/> ne
traktato	<input type="radio"/> jes <input type="radio"/> ne	flanko	<input type="radio"/> jes <input type="radio"/> ne	ŝtalo	<input type="radio"/> jes <input type="radio"/> ne	klaso	<input type="radio"/> jes <input type="radio"/> ne	nul	<input type="radio"/> jes <input type="radio"/> ne
roso	<input type="radio"/> jes <input type="radio"/> ne	nacio	<input type="radio"/> jes <input type="radio"/> ne	pejzaĝo	<input type="radio"/> jes <input type="radio"/> ne	limo	<input type="radio"/> jes <input type="radio"/> ne	utila	<input type="radio"/> jes <input type="radio"/> ne
rubando	<input type="radio"/> jes <input type="radio"/> ne	juna	<input type="radio"/> jes <input type="radio"/> ne	brovo	<input type="radio"/> jes <input type="radio"/> ne	mola	<input type="radio"/> jes <input type="radio"/> ne	brili	<input type="radio"/> jes <input type="radio"/> ne

Abbildung 31: Fragebogen E

## 4.6 Antworten in der Befragung

Nach dem Ausfüllen und Absenden eines Fragebogens durch den Befragten wurde je eine elektronische Nachricht mit den Antworten an den Befragten und an den Autor der vorliegenden Arbeit gesendet.

Dazu wurde die in der Programmiersprache PERL erstellte Skript-Datei <enketo01.cgi> verwendet; Quellcode:

```
#!/usr/bin/perl -w

use CGI::Carp qw(fatalsToBrowser);

if(defined $ENV{'REQUEST_METHOD'})
{
  if($ENV{'REQUEST_METHOD'} eq 'GET') {$Daten = $ENV{'QUERY_STRING'}};
  else {read(STDIN, $Daten, $ENV{'CONTENT_LENGTH'}};
}
else
{
  $Daten = "Nomo=Andreas&Retposhto=ak0635\@akueckl.de"
}

@Formularfelder = split(/&/, $Daten);

$i = 0;

foreach $Feld (@Formularfelder)
{
  ($Name, $Wert) = split(/=/, $Feld);
  $Wert =~ tr/+ /;
  $Wert =~ s/%([a-fA-F0-9][a-fA-F0-9])/pack("C", hex($1))/eg;
  $Wert =~ s/</&lt;/g;
  $Wert =~ s/>/&gt;/g;
  $Formular[$i] = $Name;
  $i++;
  $Formular[$i] = $Wert;
  $i++;
}

$imax = @Formular;

my $retposhtilo = "/usr/lib/sendmail";
@ricevonto = ($Formular[3], "ak631\@gmx.net");

for($j=0; $j<=@imax; $j++)
{
  open(RETPOSHTO, "|$retposhtilo -t");
  print RETPOSHTO "To: $ricevonto[$j]\n";
  print RETPOSHTO "From: $Formular[3]\n";
  print RETPOSHTO "Subject: Enketo pri la konatecgrado de Esperanto\vortoj\n";
  print RETPOSHTO "Kara $Formular[1]!\n";
  print RETPOSHTO "Dankon por via partopreno en mia enketo.\n";
  print RETPOSHTO "Jen viaj respondoj:\n";
}
```

```

for($i=0;$i<=$imax-2;$i=$i+2)
{
  $Text = $Formular[$i];
  $Text =~ s/\+/ /g;
  $Text =~ s/%28\(/g;
  $Text =~ s/%29\) /g;
# print "<p><b>$Text:</b> $Formular[$i+1]</p>\n";
  print RETPOSHTO $Text, ":", $Formular[$i+1], "\n";
}

print RETPOSHTO "\n";
print RETPOSHTO "Amike\n\n";
print RETPOSHTO "Andreas Kueck\n";
print RETPOSHTO "ak631@gmx.net\n\n";
close(RETPOSHTO);
}

print "Content-type: text/html\n\n";
print '<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN">', "\n";
print "<html><head><title>Re-informo</title></head>\n";
print "<body><h1>Re-informo</h1>\n";
print "<p><b>Dankon por via partopreno en mia enketo. Retposhtletero kun viaj indikoj estis sendita al vi kaj al mi.</p>\n";
print "</body></html>\n";

```

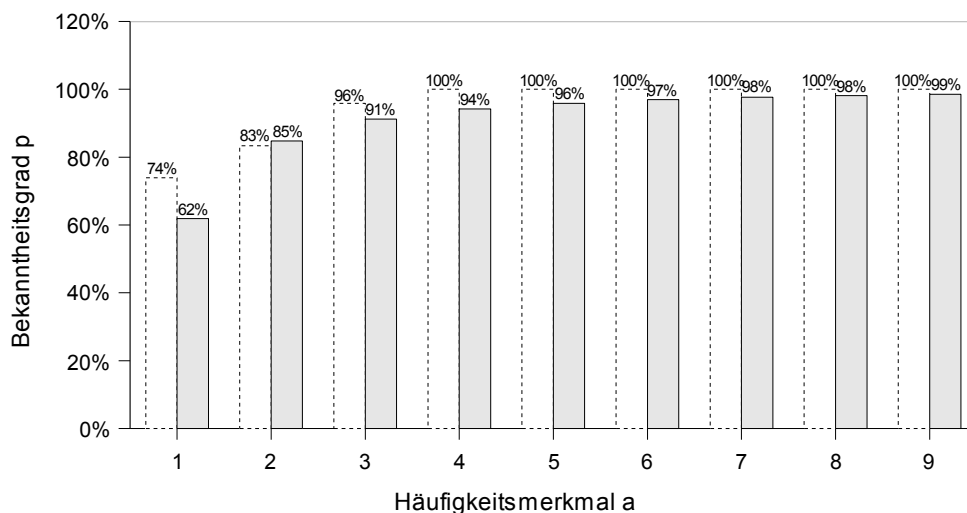
Der Autor wertete nur Antworten von solchen Befragten weiter aus, die angegeben hatten, Esperanto ausreichend gut verwenden zu können. Grund: Bei Befragten, die angegeben hatten, Esperanto nicht ausreichend gut verwenden zu können, kann die Bandbreite der Esperanto-Kenntnisse derart groß sein, dass die Ergebnisse keine wirklich verwertbare Aussagekraft im Sinne dieser Untersuchung haben. Auf diese Weise ergaben sich je sechs Befragte für jeden der fünf Fragebogen.

Nach dieser Vorauswahl wurde eine Antwort-Dartenbank nach Anhang F erstellt. Die 100 Wörter eines jeden Fragebogens wurden nach den Häufigkeitsgruppen BRO9 bis BRO1 (entsprechend Häufigkeitsmerkmal  $a = 1 \dots 9$ ) geordnet und in einer Spalte gelistet. Für jeden der Befragten wurde in einer weiteren Spalte die Kenntnis bzw. Nicht-Kenntnis der einzelnen Wörter markiert. Auf diese Weise wurde zu jedem Fragebogen A bis E die Information erhalten, wie viele Wörter pro Häufigkeitsmerkmal bekannt bzw. nicht bekannt sind. Beispiel: Fragebogen B: Die bei sechs Befragten abgefragten 30 Wörter mit Häufigkeitsmerkmal  $a = 1$  waren 159mal bekannt und 21mal nicht bekannt, die 19 Wörter mit Häufigkeitsmerkmal  $a = 2$  waren 104mal bekannt und 10mal nicht bekannt usw.

Auf diese Weise wurde für jeden der fünf Fragebogen A bis E die Ausgangsdatenbasis für eine logistische Regression nach Kapitel 4.4.3 erhalten. Diese wurde mittels des Tabellenkalkulationsprogramms OpenOffice.org Calc ausgeführt. Einzelheiten zu dem genannten Programm sind in Anhang G.

## 4.7 Auswertung der fünf Fragebogen

Mittels Gleichung 25 wurden die im Rahmen der vorliegenden Arbeit erhaltenen Stichproben-Befragungsergebnisse ausgewertet. Abbildung 32 zeigt das Ergebnis einer solchen Auswertung am Beispiel von Fragebogen E (Einzelheiten: Anhang F).



*Abbildung 32: Bekanntheitsgrad  $p$  gegen Häufigkeitsmerkmal  $a$ ; Grundlage: Fragebogen E (vgl. Abbildung 27); gestrichelte farblose Säulen: Bekanntheitsgrad nach der Vorgehensweise von Kapitel 4.4.2 (Schritt 3; also ohne Vertrauensgrad); graue Säulen: Bekanntheitsgrad bei 95% Vertrauensgrad*

Für jeden Fragebogen wurde dasjenige Häufigkeitsmerkmal  $a_{90|95}$  herausgegriffen, ab dem der Bekanntheitsgrad mindestens 90% bei einem Vertrauensgrad von 95% ist (siehe Tabelle 13). Dieses Häufigkeitsmerkmal wird in der vorliegenden Arbeit als das Häufigkeitsmerkmal mit zuverlässiger Bekanntheit definiert.

Fragebogen	Häufigkeitsmerkmal mit zuverlässiger Bekanntheit $a_{90 95}$
A	3; entspricht BRO7
B	3; entspricht BRO7
C	5; entspricht BRO5
D	2; entspricht BRO8
E	3; entspricht BRO7

*Tabelle 13: Häufigkeitsmerkmal mit zuverlässiger Bekanntheit, ermittelt aus den Fragebogen A bis E*

Aus Tabelle 13 erkennt man: Drei der fünf Fragebogen lassen darauf schließen, dass man bei Wörtern ab der Häufigkeitsgruppe BRO7 von einer zuverlässigen Bekanntheit ausgehen kann; beim vierten und fünften Fragebogen ab Häufigkeitsgruppe BRO8 bzw. BRO5.

Zwecks Zusammenfassbarkeit soll nun weiter untersucht werden, inwieweit man die Einzelergebnisse, die an den fünf Fragebogen erhalten wurden, als ähnlich ansehen kann.

#### 4.8 Ähnlichkeitsanalyse

Es soll festgestellt werden, inwieweit die an den fünf Fragebogen erhaltenen Ergebnisse überhaupt repräsentativ für die Grundgesamtheit und miteinander vergleichbar sind. Als statistischer Test bietet sich dazu der Kruskal-Wallis-Test an (Kruskal 1952). Geprüft werden soll damit im vorliegenden Fall die

*Nullhypothese  $H_0$ : Es gibt keinen signifikanten Unterschied in den  $m = 5$  Stichproben (Signifikanzniveau: 5%).*

Die Funktionsweise dieses statistischen Tests wird anhand der vorliegenden Aufgabenstellung erläutert:

- (a) Für jeden der  $m = 5$  Fragebogen A bis E wird für jedes der  $n_a = 9$  Häufigkeitsmerkmale  $a = 1 \dots 9$  der jeweilige Bekanntheitsgrad  $p_o$  nach Gleichung 15 berechnet (mit  $\mu = \mu_o$  und  $\sigma = \sigma_o$ ).
- (b) Die so erhaltenen  $N_R = 45$   $p_o$ -Werte ( $N_R$  ist die Anzahl der Rangwerte) werden in eine aufsteigende Rangordnung mit Rangwerten von 1 bis 45 gebracht (Einzelheiten: Tabelle 14).

- (c) Die einzelnen Rangwerte werden denjenigen Fragebogen zugeordnet, aus deren Ergebnissen die Rangwerte stammen. Beispiel: Der Rangwert 1 (mit  $p_0 = 0,263$  als dem kleinsten auftretenden  $p_0$ -Wert) wird dem Fragebogen E zugeordnet, weil sich dieser  $p_0$ -Wert aus den Ergebnissen von Fragebogen E ergab.
- (d) Die jedem Fragebogen zugeordneten Rangwerte werden summiert. Man erhält folgende Rangsummen  $R_A$  bis  $R_E$ : Für Fragebogen A:  $R_A = 183$ ; für Fragebogen B:  $R_B = 224$ ; für Fragebogen C:  $R_C = 140$ ; für Fragebogen D:  $R_D = 261$ ; für Fragebogen E:  $R_E = 227$ .
- (e) Es wird die Prüfgröße  $H$  nach Gleichung 96 gebildet:

$$H = \frac{12}{N_R(N_R+1)} \left( \frac{R_A^2}{n_a} + \frac{R_B^2}{n_a} + \frac{R_C^2}{n_a} + \frac{R_D^2}{n_a} + \frac{R_E^2}{n_a} \right) - 3(N_R+1) \quad (96)$$

- (f) Im vorliegenden Fall ergibt sich für die Prüfgröße  $H = 5,58$ .
- (g) Es wird das Quantil der  $\chi^2$ -Verteilung mit dem gewählten Signifikanzniveau (hier: 5%) und  $m-1$  Freiheitsgraden (hier:  $m-1 = 4$ ) ermittelt:  $\chi^2(5\%; 4) = 9,49$ .
- (h) Die Nullhypothese wird angenommen, wenn  $H < \chi^2(5\%; 4)$ . Im vorliegenden Fall trifft das zu.

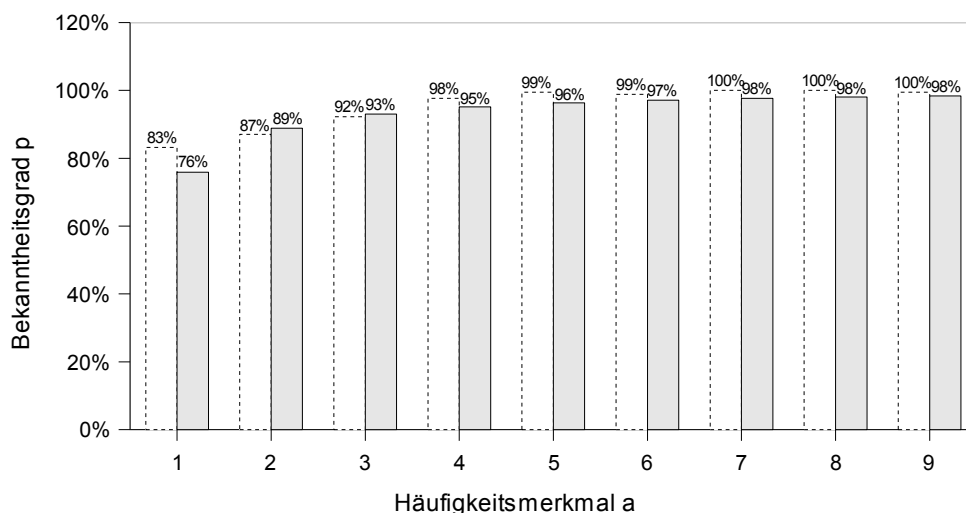
Die o. g. Nullhypothese wird also angenommen.

a	P <sub>0</sub>	Fragebogen (1=A usw.)	Rangwert	Rangwert-Zuordnung zu Fragebogen				
				A	B	C	D	E
1	0,26334990	5	1	0	0	0	0	1
1	0,31143242	1	2	2	0	0	0	0
1	0,32352431	3	3	0	0	3	0	0
1	0,33897844	4	4	0	0	0	4	0
1	0,34952875	2	5	0	5	0	0	0
2	0,85221041	3	6	0	0	6	0	0
2	0,90446670	1	7	7	0	0	0	0
2	0,93086728	5	8	0	0	0	0	8
2	0,94160426	2	9	0	9	0	0	0
3	0,96114936	3	10	0	0	10	0	0
2	0,96132430	4	11	0	0	0	11	0
3	0,98248312	1	12	12	0	0	0	0
4	0,98582096	3	13	0	0	13	0	0
3	0,99118768	5	14	0	0	0	0	14
3	0,99159147	2	15	0	15	0	0	0
5	0,99358833	3	16	0	0	16	0	0
4	0,99497940	1	17	17	0	0	0	0
3	0,99586165	4	18	0	0	0	18	0
6	0,99665878	3	19	0	0	19	0	0
4	0,99793755	2	20	0	20	0	0	0
4	0,99803208	5	21	0	0	0	0	21
7	0,99807664	3	22	0	0	22	0	0
5	0,99810805	1	23	23	0	0	0	0
8	0,99880853	3	24	0	0	24	0	0
6	0,99914898	1	25	25	0	0	0	0
4	0,99917066	4	26	0	0	0	26	0
9	0,99921922	3	27	0	0	27	0	0
5	0,99930910	2	28	0	28	0	0	0
5	0,99938728	5	29	0	0	0	0	29
7	0,99956710	1	30	30	0	0	0	0
6	0,99971749	2	31	0	31	0	0	0
8	0,99975900	1	32	32	0	0	0	0
5	0,99976210	4	33	0	0	0	33	0
6	0,99976401	5	34	0	0	0	0	34
9	0,99985625	1	35	35	0	0	0	0
7	0,99986739	2	36	0	36	0	0	0
7	0,99989469	5	37	0	0	0	0	37
6	0,99991427	4	38	0	0	0	38	0
8	0,99993113	2	39	0	39	0	0	0
8	0,99994765	5	40	0	0	0	0	40
9	0,99996136	2	41	0	41	0	0	0
7	0,99996383	4	42	0	0	0	42	0
9	0,99997174	5	43	0	0	0	0	43
8	0,99998288	4	44	0	0	0	44	0
9	0,99999114	4	45	0	0	0	45	0
H =		5,58	Σ	183	224	140	261	227
X <sup>2</sup> (95%;5-1)		9,49						

Tabelle 14: Hilfstabelle für Kruskal-Wallis-Test auf Ähnlichkeit

## 4.9 Gesamtauswertung

Aufgrund ihrer statistisch nachgewiesenen Ähnlichkeit wird die logistische Regression auch an den zusammengefassten Antworten aller fünf Fragebogen durchgeführt. Abbildung 33 zeigt den Bekanntheitsgrad in Abhängigkeit vom Häufigkeitsmerkmal bei 95% Vertrauensgrad (es wurden 3000 einzelne Aussagen zur Bekanntheit berücksichtigt).



*Abbildung 33: Bekanntheitsgrad p gegen Häufigkeitsmerkmal a; Grundlage: Fragebogen A bis E; gestrichelte farblose Säulen: Bekanntheitsgrad nach der Vorgehensweise von Kapitel 4.4.2 (Schritt 3; also ohne Vertrauensgrad); graue Säulen: Bekanntheitsgrad bei 95% Vertrauensgrad*

Eine wichtige Erkenntnis ist: Es ist  $a_{90|95} = 3$  das Häufigkeitsmerkmal mit zuverlässiger Bekanntheit, also die BRO-Häufigkeitsgruppe 7. Man kann somit folgern, dass die Esperanto-Wörter, die zur Gruppe der siebthäufigsten gehören, zuverlässig bekannt sind.

## 4.10 Auswertung nach einzelnen Sprachfamilien und -zweigen

### 4.10.1 Motivation

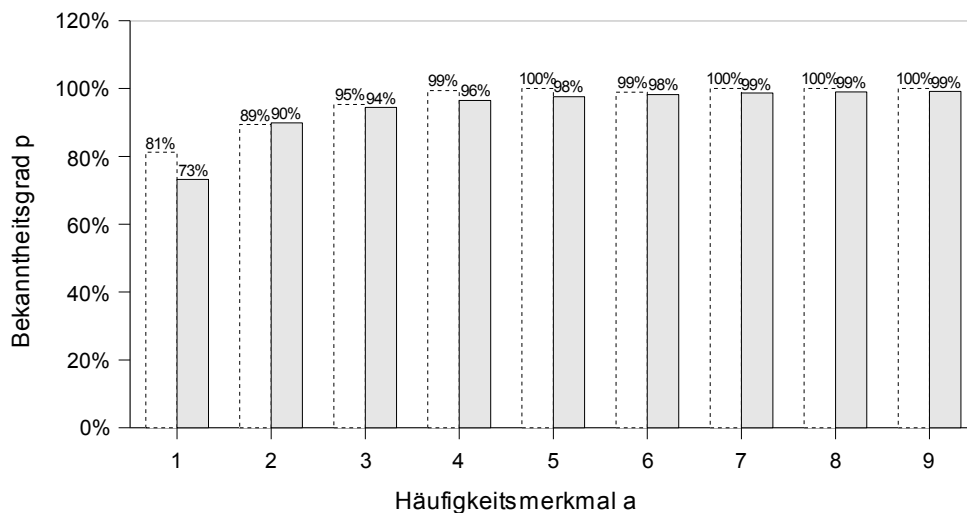
Wie bereits in Kapitel 2.1 erwähnt, sind die meisten Esperanto-Wörter aus romanischen und germanischen Sprachen abgeleitet. Es liegt daher nahe zu vermuten, dass Esperanto-Sprechern mit einer Muttersprache aus den genannten Sprachzweigen der Esperanto-Wortschatz geläufiger ist als



Anwendern mit einer Muttersprache aus einem anderen Sprachzweig oder gar einer anderen Sprachfamilie als der indoeuropäischen. Die folgenden Auswertungen sollen dieser Vermutung nachgehen. Es werden daher im Folgenden die Häufigkeitsmerkmale mit zuverlässiger Bekanntheit auf der Grundlage von solchen Befragungsergebnissen ermittelt, die nach der Zugehörigkeit der Muttersprache der Befragten zu einer bestimmten Sprachfamilie oder einem bestimmten Sprachzweig aussortiert werden. Diese Aussortierungen wurden sowohl unter sprachtypisierenden Gesichtspunkten als auch unter der (statistisch motivierten) Prämisse gebildet, in jeder Aussortierung mindestens fünf Befragte auszuwerten.

#### 4.10.2 Muttersprache aus nicht-indoeuropäischer Sprachfamilie

Unter den Befragten waren 16 Menschen mit einer Muttersprache, die nicht zur indoeuropäischen Sprachfamilie gehört. Es folgen die betreffenden Sprachen (Befragte): Chinesisch (2), Estnisch (3), Finnisch (1), Japanisch (8), Ungarisch (2). Abbildung 34 zeigt den Bekanntheitsgrad in Abhängigkeit vom Häufigkeitsmerkmal.



*Abbildung 34: Bekanntheitsgrad  $p$  gegen Häufigkeitsmerkmal  $a$ ; bei Muttersprache aus nicht-indogermanischer Sprachfamilie; gestrichelte farblose Säulen: Bekanntheitsgrad nach der Vorgehensweise von Kapitel 4.4.2 (Schritt 3; also ohne Vertrauensgrad); graue Säulen: Bekanntheitsgrad bei 95% Vertrauensgrad*

Das Häufigkeitsmerkmal mit zuverlässiger Bekanntheit ist 2, also die BRO-Häufigkeitsgruppe 8.

### 4.10.3 Muttersprache aus indoeuropäischer Sprachfamilie

Die anderen 14 Befragten haben eine Muttersprache, die zur indoeuropäischen Sprachfamilie gehört. Es folgen die betreffenden Sprachen (Befragte): Dänisch (1), Deutsch (2), Farsi (1), Französisch (3), Lettisch (1), Litauisch (2), Polnisch (2), Portugiesisch (1), Rumänisch (1). Abbildung 35 zeigt den Bekanntheitsgrad in Abhängigkeit vom Häufigkeitsmerkmal.

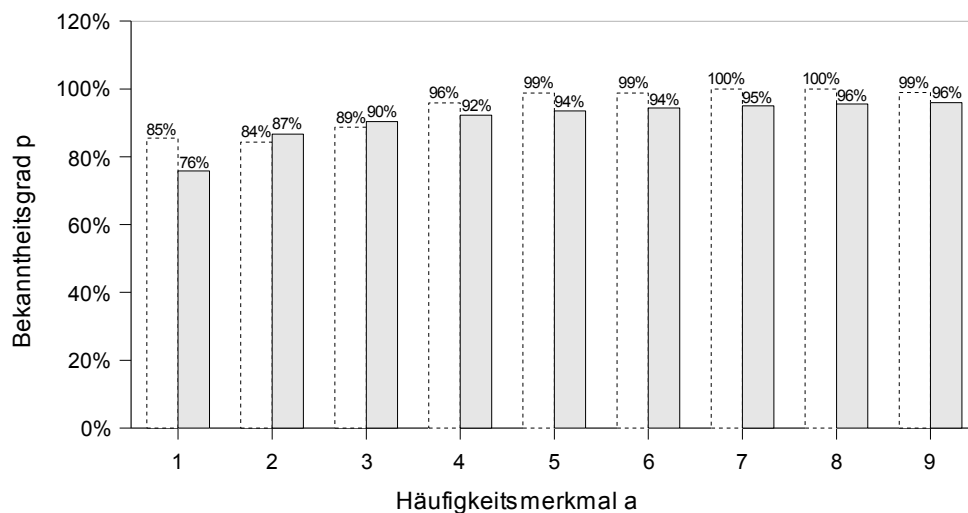
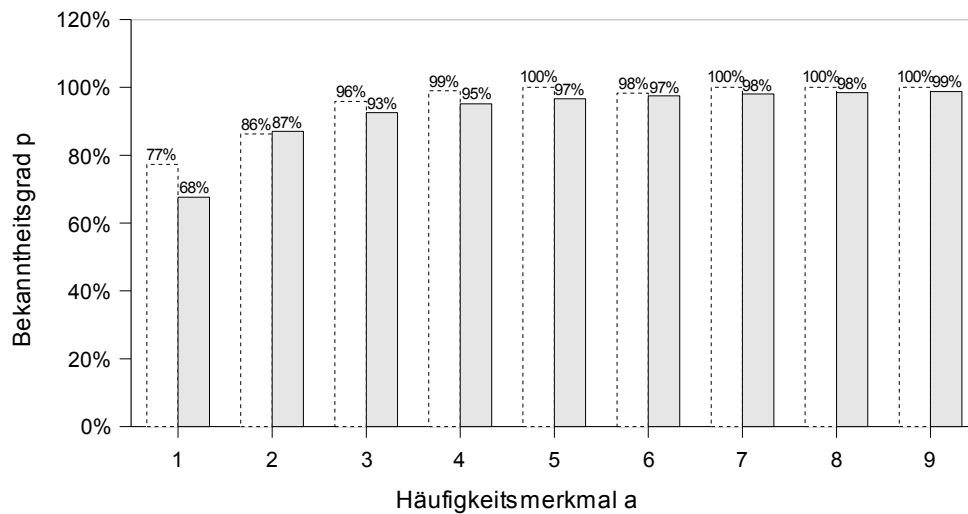


Abbildung 35: Bekanntheitsgrad  $p$  gegen Häufigkeitsmerkmal  $a$ ; bei Muttersprache aus indogermanischer Sprachfamilie; gestrichelte farblose Säulen: Bekanntheitsgrad nach der Vorgehensweise von Kapitel 4.4.2 (Schritt 3; also ohne Vertrauensgrad); graue Säulen: Bekanntheitsgrad bei 95% Vertrauensgrad

Das Häufigkeitsmerkmal mit zuverlässiger Bekanntheit ist 3, also die BRO-Häufigkeitsgruppe 7.

### 4.10.4 Muttersprache Chinesisch oder Japanisch

Die Muttersprache von 10 Befragten ist Chinesisch (2) oder Japanisch (8). Abbildung 36 zeigt den Bekanntheitsgrad in Abhängigkeit vom Häufigkeitsmerkmal.

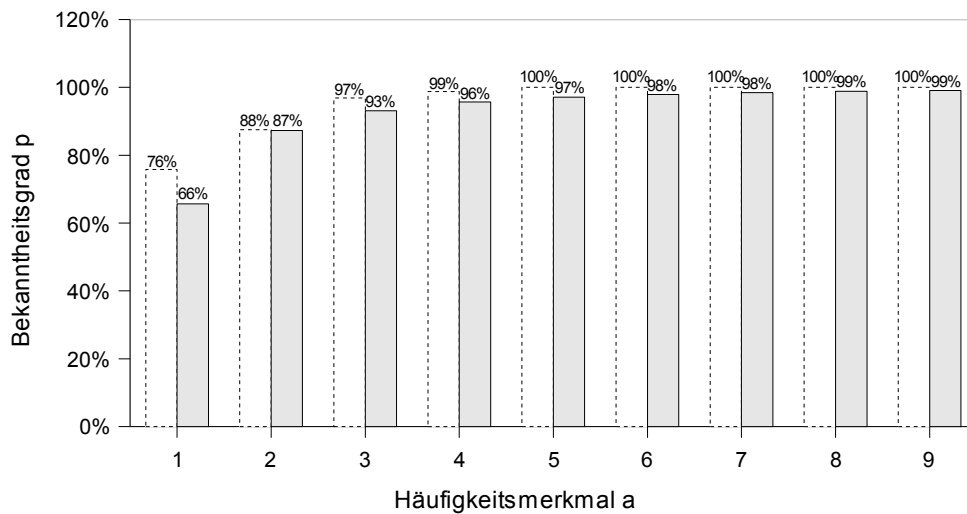


*Abbildung 36: Bekanntheitsgrad p gegen Häufigkeitsmerkmal a; bei Muttersprache Chinesisch oder Japanisch; gestrichelte farblose Säulen: Bekanntheitsgrad nach der Vorgehensweise von Kapitel 4.4.2 (Schritt 3; also ohne Vertrauensgrad); graue Säulen: Bekanntheitsgrad bei 95% Vertrauensgrad*

Das Häufigkeitsmerkmal mit zuverlässiger Bekanntheit ist 3, also die BRO-Häufigkeitsgruppe 7.

#### **4.10.5 Muttersprache Japanisch**

Da die Muttersprache von 8 Befragten Japanisch ist, wurde gesondert nach dieser Sprache aussortiert. Abbildung 37 zeigt den Bekanntheitsgrad in Abhängigkeit vom Häufigkeitsmerkmal.

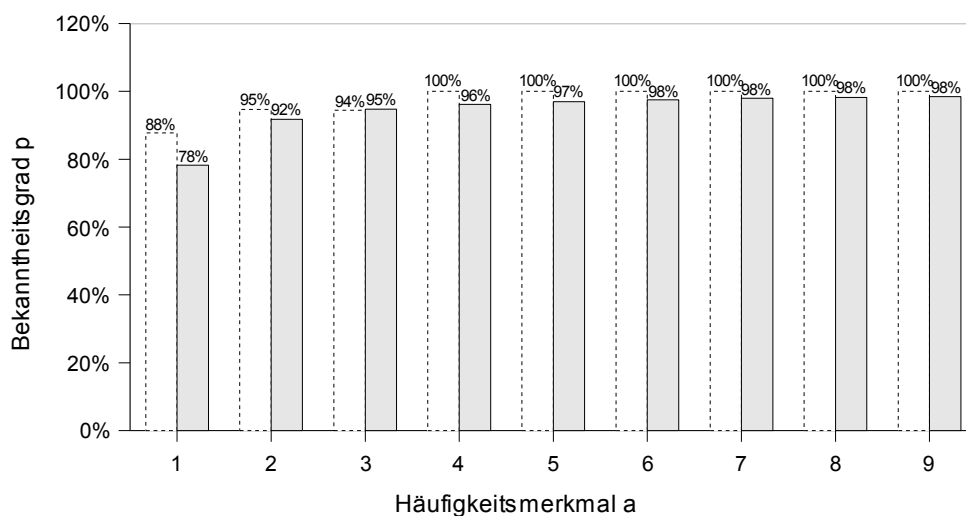


*Abbildung 37: Bekanntheitsgrad p gegen Häufigkeitsmerkmal a; bei Muttersprache Japanisch; gestrichelte farblose Säulen: Bekanntheitsgrad nach der Vorgehensweise von Kapitel 4.4.2 (Schritt 3; also ohne Vertrauensgrad); graue Säulen: Bekanntheitsgrad bei 95% Vertrauensgrad*

Das Häufigkeitsmerkmal mit zuverlässiger Bekanntheit ist 3, also die BRO-Häufigkeitsgruppe 7.

#### **4.10.6 Muttersprache aus uralischer Sprachfamilie**

Unter den Befragten waren 6 Menschen mit einer Muttersprache, die zur uralischen Sprachfamilie gehört. Es folgen die betreffenden Sprachen (Befragte): Estnisch (3), Finnisch (1), Ungarisch (2). Abbildung 38 zeigt den Bekanntheitsgrad in Abhängigkeit vom Häufigkeitsmerkmal.



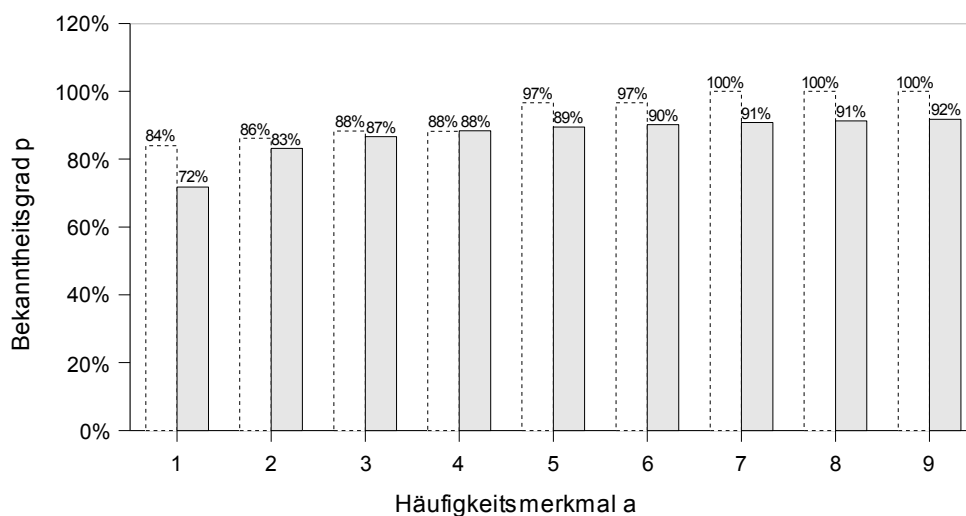
*Abbildung 38: Bekanntheitsgrad p gegen Häufigkeitsmerkmal a; bei Muttersprache aus uralischer Sprachfamilie; gestrichelte farblose Säulen: Bekanntheitsgrad nach der Vorgehensweise von Kapitel 4.4.2 (Schritt 3; also ohne Vertrauensgrad); graue Säulen: Bekanntheitsgrad bei 95% Vertrauensgrad*

Das Häufigkeitsmerkmal mit zuverlässiger Bekanntheit ist 2, also die BRO-Häufigkeitsgruppe 8.

#### **4.10.7 Muttersprache aus dem baltischen oder slawischen Sprachzweig**

Unter den Befragten waren 5 Menschen mit einer Muttersprache, die zum baltischen oder slawischen<sup>61</sup> Sprachzweig der indoeuropäischen Sprachfamilie gehört. Es folgen die betreffenden Sprachen (Befragte): Lettisch (1), Litauisch (2), Polnisch (2). Abbildung 39 zeigt den Bekanntheitsgrad in Abhängigkeit vom Häufigkeitsmerkmal.

<sup>61</sup> Laut Bußmann (2002: 300) bilden Baltisch und Slawisch sogar „möglicherweise als Balto-Slawisch eine genetische Einheit“. In der vorliegenden Arbeit ist die Zusammenlegung statistisch motiviert (Ermöglichung dieser Auswertung mit zumindest fünf Befragten).



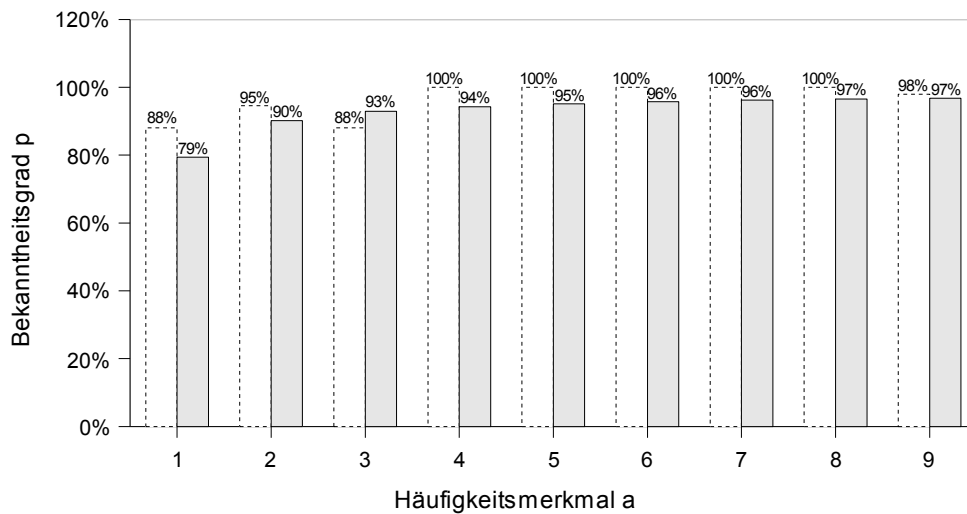
*Abbildung 39: Bekanntheitsgrad p gegen Häufigkeitsmerkmal a; bei Muttersprache aus baltischem oder slawischem Sprachzweig; gestrichelte farblose Säulen: Bekanntheitsgrad nach der Vorgehensweise von Kapitel 4.4.2 (Schritt 3; also ohne Vertrauensgrad), graue Säulen: Bekanntheitsgrad bei 95% Vertrauensgrad*

Das Häufigkeitsmerkmal mit zuverlässiger Bekanntheit ist 6, also die BRO-Häufigkeitsgruppe 4.

#### **4.10.8 Muttersprache aus dem germanischen oder romanischen Sprachzweig**

Unter den Befragten waren 8 Menschen mit einer Muttersprache, die zum germanischen oder romanischen Sprachzweig<sup>62</sup> der indoeuropäischen Sprachfamilie gehört. Es folgen die betreffenden Sprachen (Befragte): Dänisch (1), Deutsch (2), Französisch (3), Portugiesisch (1), Rumänisch (1). Abbildung 40 zeigt den Bekanntheitsgrad in Abhängigkeit vom Häufigkeitsmerkmal.

<sup>62</sup> Hiermit soll keinesfalls eine Hypothese aufgestellt werden, dass der germanische und romanische Sprachzweig eine genetische Einheit o. Ä. bilden. Es handelt sich um jedoch diejenigen Sprachzweige, aus denen die meisten Esperanto-Lexeme abgeleitet wurden (vgl. Kapitel 2.1).

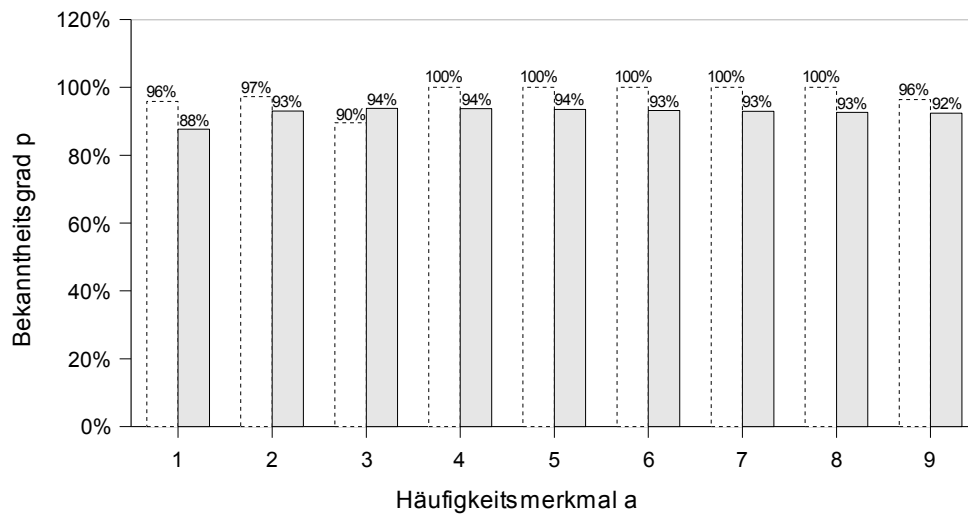


*Abbildung 40: Bekanntheitsgrad  $p$  gegen Häufigkeitsmerkmal  $a$ ; bei Muttersprache aus germanischem oder romanischem Sprachzweig; gestrichelte farblose Säulen: Bekanntheitsgrad nach der Vorgehensweise von Kapitel 4.4.2 (Schritt 3; also ohne Vertrauensgrad), graue Säulen: Bekanntheitsgrad bei 95% Vertrauensgrad*

Das Häufigkeitsmerkmal mit zuverlässiger Bekanntheit ist 2, also die BRO-Häufigkeitsgruppe 8.

#### **4.10.9 Muttersprache aus dem romanischen Sprachzweig**

Unter den Befragten waren 5 Menschen mit einer Muttersprache, die zum romanischen Sprachzweig der indoeuropäischen Sprachfamilie gehört. Es folgen die betreffenden Sprachen (Befragte): Französisch (3), Portugiesisch (1), Rumänisch (1). Abbildung 41 zeigt den Bekanntheitsgrad in Abhängigkeit vom Häufigkeitsmerkmal.



*Abbildung 41: Bekanntheitsgrad  $p$  gegen Häufigkeitsmerkmal  $a$ ; bei Muttersprache aus romanischem Sprachzweig; gestrichelte farblose Säulen: Bekanntheitsgrad nach der Vorgehensweise von Kapitel 4.4.2 (Schritt 3; also ohne Vertrauensgrad), graue Säulen: Bekanntheitsgrad bei 95% Vertrauensgrad*

Das Häufigkeitsmerkmal mit zuverlässiger Bekanntheit ist 2, also die BRO-Häufigkeitsgruppe 8.



#### 4.10.10 Zusammenfassung und Kommentierung der Auswertungen

Tabelle 15 fasst die Gesamtauswertung und die Einzelauswertungen zusammen.

Muttersprache	Häufigkeitsmerkmal mit zuverlässiger Bekanntheit
Aus indoeuropäischer oder nicht-indoeuropäischer Sprachfamilie	3, also BRO-Häufigkeitsgruppe 7
Aus nicht-indoeuropäischer Sprachfamilie	2, also BRO-Häufigkeitsgruppe 8
Aus indoeuropäischer Sprachfamilie	3, also BRO-Häufigkeitsgruppe 7
Chinesisch oder Japanisch	3, also BRO-Häufigkeitsgruppe 7
Japanisch	3, also BRO-Häufigkeitsgruppe 7
Aus uralischer Sprachfamilie	2, also BRO-Häufigkeitsgruppe 8
Aus baltischem oder slawischem Sprachzweig	6, also BRO-Häufigkeitsgruppe 4
Aus germanischem oder romanischem Sprachzweig	2, also BRO-Häufigkeitsgruppe 8
Aus romanischem Sprachzweig	2, also BRO-Häufigkeitsgruppe 8

*Tabelle 15: Zusammenfassung von Gesamt- und Einzelauswertungen*

Eine wichtige Erkenntnis ist, dass das Häufigkeitsmerkmal mit zuverlässiger Bekanntheit durch eine der neun BRO-Häufigkeitsgruppen wiedergegeben wird. Das trifft sowohl für die einzelnen Auswertungen als auch für die Gesamtauswertung zu (hier BRO7, also die Gruppe mit den siebthäufigsten Wörtern). Offensichtlich wurde der durch die ersten sieben BRO-Häufigkeitsgruppen repräsentierte Wortschatz des Esperanto von dessen Anwendern zuverlässig aufgenommen.

Vielleicht ein wenig überraschend ist, dass es keinen Zusammenhang zwischen dem Häufigkeitsmerkmal mit zuverlässiger Bekanntheit und der Zugehörigkeit der Muttersprache zu einer bestimmten Sprachfamilie oder einem bestimmten Sprachzweig gibt, obgleich der Großteil der Esperanto-Wörter von Wörtern aus germanischen und romanischen Sprachen abgeleitet ist<sup>63</sup>. Vielleicht glichen und gleichen Esperanto-Sprecher mit nicht-indoeuropäischer Muttersprache durch

<sup>63</sup> In Sikosek (2003<sup>2</sup>: 65) wird darauf hingewiesen, dass „der alte Glaube, die Wortschatzzusammensetzung einer Plansprache sei mit ihrer Verbreitung korreliert, nicht erklären kann, dass Länder mit nicht-indoeuropäischer Sprache (wie Ungarn und Finnland) eine überproportional starke Esperanto-Bewegung haben“.

besonders intensives Lernen den mangelnden Vorteil bezüglich des Esperanto-Wortschatzes hinsichtlich der Wortähnlichkeit aus.

Das einzige auffallend abweichende Einzelergebnis ist das der Befragten mit einer Muttersprache aus dem baltischen oder slawischen Sprachzweig: Erst Wörter aus BRO-Häufigkeitsgruppe 4 (also der vierthäufigsten Wortgruppe) wurden zuverlässig gekannt. Zum Vergleich: Die Gesamtauswertung und nach anderen Sprachzweigen oder -familien gefilterten Auswertungen ergaben BRO-Häufigkeitsgruppe 7 oder 8 (also bereits die „nur siebt- oder achthäufigste“ Wortgruppe).

Erwartungsgemäß „gut“ sind die Einzelergebnisse der Befragten mit einer Muttersprache aus dem germanischen oder romanischen Sprachzweig und auch aus dem romanischen Sprachzweig für sich genommen: Die BRO-Häufigkeitsgruppe 8, also die „achthäufigste Wortgruppe“ wird bereits zuverlässig gekannt. Es spricht für die Stärke des Auswerteverfahrens, dass das Nicht-Kennen des Zahlwortes „ok“ („acht“; ein Wort mit  $a = 9$ ) bei einem Befragten mit Muttersprache Französisch sich offensichtlich nicht signifikant auf das Endergebnis auswirkte. Eine Auswertung nach der 29/29-Methode (Kapitel 4.4.2) würde hier ergeben, dass das Häufigkeitsmerkmal mit zuverlässiger Bekanntheit  $a_{90|95} = 9$  ist (also BRO1, die Gruppe der häufigsten Wörter). Grund: Erst bei den Häufigkeitsmerkmalen  $a = 8$  und  $a = 9$  ergibt sich mit (zusammengenommen) 47 Bekannt- und einer Nicht-bekannt-Aussage eine Konstellation, mit der nach Gleichung 14 das Häufigkeitsmerkmal mit zuverlässiger Bekanntheit bestimmbar ist; hier ist  $a_{90|95} = 9$  (es ist der konservative Wert zu nehmen, also BRO1). Zur Erinnerung: Die 29/29-Methode „braucht“ Paare wie 29/29, 46/45, 61/59 usw. (vgl. Kapitel 4.4.2).

#### **4.11 Selbsteinschätzung des eigenen Esperanto-Anwendungsvermögens**

In der Befragung wurden die Teilnehmer auch gebeten anzugeben, in welchem Jahr sie begannen, Esperanto zu lernen, und welche der folgenden beiden Aussagen zur Selbsteinschätzung des eigenen Esperanto-Anwendungsvermögens zutrifft:

- (a) „Ich kann Esperanto ausreichend gut verwenden“ oder

(b) „Ich kann Esperanto nicht ausreichend gut verwenden“

Auch diese Befragungsergebnisse wurden mittels logistischer Regression ausgewertet ( $y = 1$  oder  $y = 0$  in Abhängigkeit von der Esperanto-Lernzeit  $j$  in Jahren;  $j$ : Jahr der Aussage [hier: 2007] minus Jahr des Esperanto-Lernbeginns). Abbildung 42 zeigt das Ergebnis. Einzelheiten sind in Anhang F, Abbildungen 59 und 60.

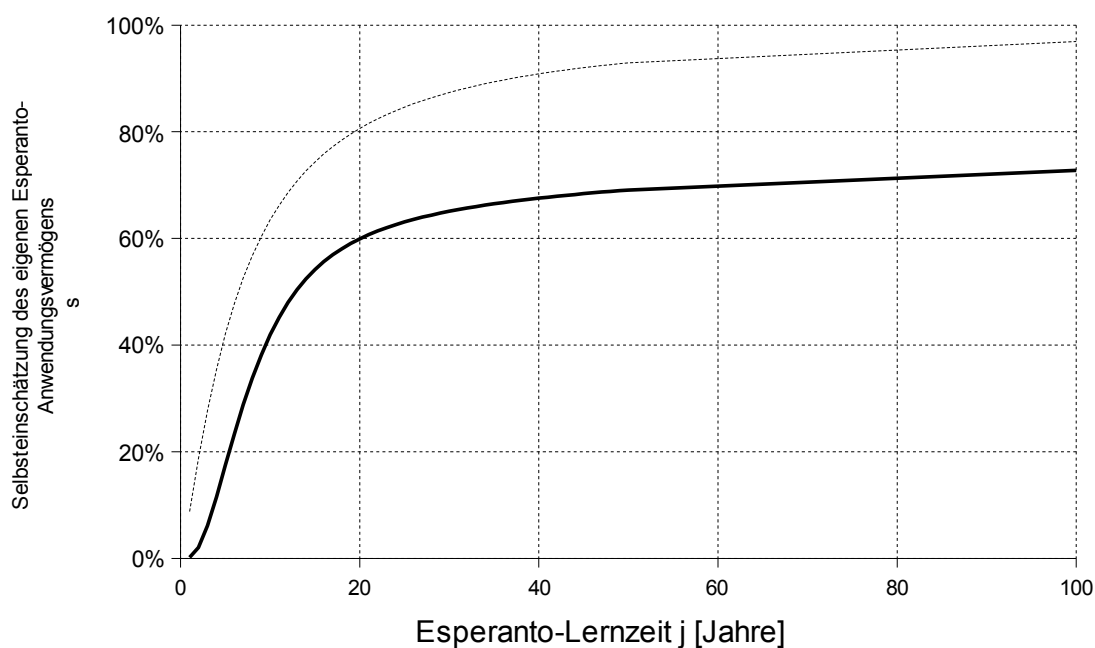


Abbildung 42: Selbsteinschätzung des eigenen Esperanto-Anwendungsvermögens  $s$  gegen Esperanto-Lernzeit  $j$ ; durchgezogene Linie: Vertrauensgrad 95%; gestrichelte Linie: ohne Vertrauensgrad

Bei einer statistischen Sicherheit von 95% ist die Selbsteinschätzung des eigenen Esperanto-Anwendungsvermögens durch seine Sprecher auch nach 100 Lernjahren 73%. Interessant ist, dass keine Lernzeit für eine zuverlässige Anwendbarkeit (also 90% bei 95% Vertrauensgrad) ermittelt werden kann, dass also „ewig“ gelernt muss: Auf der Grundlage der beschriebenen statistisch untersuchten Stichprobe besteht die Esperanto-Sprecherschaft also aus Leuten, die sich in Bezug auf Esperanto als „lebenslang Lernende“ sehen (was übrigens auch der subjektiven Erfahrung des Autors nicht widerspricht).

## 4.12 Prüfung der Modell-Tauglichkeit

### 4.12.1 Devianz-Test

Es soll festgestellt werden, inwieweit das Berens-Modell (also Gleichung 15) mit den empirischen Werten übereinstimmt. Als statistischer Test bietet sich der Devianz-Test von Tiede (1995: 22) an. Geprüft werden soll damit im vorliegenden Fall die

*Nullhypothese  $H_0$ : Zwischen den empirisch erhaltenen Stichproben-Befragungsergebnissen und dem für ihre Beschreibung verwendeten Berens-Modell (also Gleichung 15) gibt es keinen signifikanten Unterschied (Signifikanzniveau: 5%).*

Die Funktionsweise dieses statistischen Tests wird anhand der vorliegenden Aufgabenstellung erläutert:

- (a) Es wird zu jeder durchgeführten logistischen Regression als Prüfgröße die Devianz nach Gleichung 97 gebildet. ( $L_0$  ist die Mutmaßlichkeit nach Gleichung 24 mit den Größen  $\mu_0$  und  $\sigma_0$ .)

$$D = -2 \ln(L_0) \quad (97)$$

- (b) Es wird das Quantil der  $\chi^2$ -Verteilung mit dem gewählten Signifikanzniveau (hier: 5%) und  $N-2$  Freiheitsgraden gebildet. Die Größe  $N$  ist der Stichprobenumfang. Die davon abgezogene 2 ist die Anzahl der Modellparameter (also hier  $\mu$  und  $\sigma$ ).
- (c) Die Nullhypothese wird angenommen, wenn  $D < \chi^2(5\%; N-2)$ . Für alle im Rahmen der vorliegenden Arbeit durchgeführten logistischen Regressionen trifft das zu.

Die o. g. Nullhypothese wird angenommen.

### 4.12.2 Vergleich mit einem alternativen Modell

Wie bereits in Kapitel 4.4.3 erwähnt, wird in Köhler, Rapp (2007) folgendes Modell zwischen Wortvertrautheit  $y$  und Worthäufigkeit  $x$  angegeben ( $A$ : Konstante;  $B$ : Proportional-Operator):

$$y = Ax^B \quad (98)$$

Es wird eine Anpassung des Modells von Gleichung 98 an die mit Gleichung 15 für  $p_0$  erhaltene Gesamtauswertung vorgenommen. Für Gleichung 15 werden die Parameter  $\mu = \mu_0$  und  $\sigma = \sigma_0$  entsprechend der Gesamtauswertung angesetzt, also  $\mu_0 = -1.0996$  und  $\sigma_0 = 1,3671$  (Einzelheiten: Anhang F, Abbildung 50). Für Gleichung 98 werden angesetzt: die Wortvertrautheit  $y$  durch den Bekanntheitsgrad (in diesem Fall  $p_{0,K}$  genannt); die Worthäufigkeit  $x$  durch den Mittelwert  $d$  der Häufigkeiten im Dietze-Korpus (ausgedrückt durch Gleichung 12). Dies führt zu Gleichung 99:

$$p_{0,K} = A(111,0696(10 - a)^{-1,5011})^B \quad (99)$$

Die Anpassungsrechnung ergibt  $A = 0,8686$ ,  $B = 0,0344$  bei einem Bestimmtheitsmaß  $D = 0,41$  (Einzelheiten: Anhang F, Abbildung 61). Abbildung 43 zeigt  $p_0$  nach Gleichung 15 und  $p_{0,K}$  nach Gleichung 99. Die nicht zu stark ausgeprägte Abweichung wird als weitere Bestätigung der Tauglichkeit des durch Gleichung 15 für den Bekanntheitsgrad gegebenen Modells gewertet.

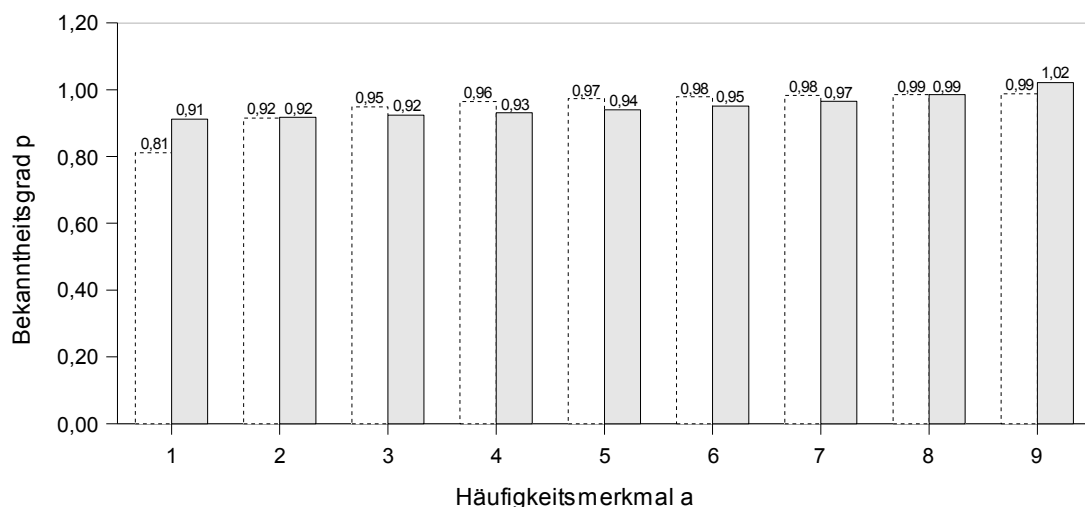


Abbildung 43: Vergleich der mit Gleichung 15 (gestrichelte farblose Säulen) und Gleichung 99 (graue Säulen) gegebenen Modelle

## 5 Zusammenfassung

### 5.1 Allgemeines zu dieser Arbeit

Es wurden quantitative Studien zur Plansprache Esperanto zu ausgewählten Fragestellungen der Phonologie, Morphologie, Syntax, Lexik, Semantik, Worthäufigkeit und – als Schwerpunkt – Wortbekanntheit durchgeführt. Zur Untersuchung des Zusammenhangs zwischen Wortbekanntheit und Worthäufigkeit sowie zur Untersuchung der Selbsteinschätzung des eigenen Esperanto-Anwendungsvermögens durch seine Sprecher wurde das Verfahren der logistischen Regression ausführlich dargestellt und als eine innovative Option auch zur Untersuchung analoger Fragestellungen angeboten. In den folgenden Kapiteln 5.2 bis 5.8 sind die Ergebnisse zusammengefasst; in Kapitel 5.9 wird ein Ausblick gegeben.

Die Relevanz des Schwerpunkt-Themas Wortbekanntheit besteht auf zwei Ebenen:

- (a) *Auf Esperanto bezogen*: Die Untersuchung lieferte unmittelbar statistisch fundierte Erkenntnisse darüber, inwieweit das Lexikon des Esperanto von dessen Sprechern beherrscht wird. Zudem wurden Erkenntnisse darüber gewonnen, inwieweit die Tatsache, dass das Esperanto-Lexikon größtenteils aus germanischen und romanischen Sprachen abgeleitet wurde (also gewissermaßen „europäischlastig“ ist), nachteilig ist für Sprecher mit einer Muttersprache aus anderen Sprachfamilien und -zweigen. Weitere Erkenntnisse wurden bezüglich der Selbsteinschätzung des eigenen Esperanto-Anwendungsvermögens durch seine Sprecher gewonnen.
- (b) *Allgemein*: Das Modell der logistischen Regression (mit Berücksichtigung der statistischen Sicherheit) wird der quantitativen Linguistik als innovative Option zur Datenanalyse (und entsprechender Auslegung von Erhebungen) angeboten.

## 5.2 Phonologie

Das Inventar der Phoneme des Esperanto und ihre Häufigkeitsverteilung wurden dargestellt. An Letztere wurden verschiedene Modellfunktionen angepasst, die sich bereits bei Phonem-Häufigkeitsverteilungen anderer Sprachen bewährt haben: Tuldava (1988); Zipf (1929, 1935, 1949); Yule (1924); Naranan, Balasubrahmanyam (1992a, b, 2000); Altmann (1993); Sigurd (1968); Good (1969); Altmann, Popescu<sup>64</sup>. Damit wurde für die Plansprache Esperanto die in Strauss, Altmann, Best (2007) aufgestellte Hypothese überprüft, nach der die geordnete Reihenfolge der Phonemhäufigkeiten einer regulären Wahrscheinlichkeitsfunktion oder regulären monoton abnehmenden Funktion folgt. Folgendes ist festzuhalten:

- (a) Esperanto verfügt über 28 Phoneme (23 Konsonanten; fünf Vokale); sie werden phonemisch verschriftet.
- (b) Die absteigende Rangfolge der relativen Phonemhäufigkeit sieht wie folgt aus (nach Haszpra 2001; Korpus-Umfang: 496.196 Phoneme; aus Haszpra 1998): 1. a (12,59%); 2. e (9,36%); 3. o (8,99%); 4. i (8,80%); 5. n (7,79%); 6. l (6,24%); 7. r (5,97%); 8. s (5,91%); 9. t (5,50%); 10. k (4,22%); 11. u (3,36%); 12. m (3,06%); 13. d (3,04%); 14. p (2,82%); 15. j (2,71%); 16. v (1,87%); 17. g (1,28%); 18. b (1,15%); 19. f (1,09%); 20. c (0,85%); 21. ĝ (0,69%); 22. ĉ (0,67%); 23. z (0,52%); 24. ŭ (0,51%); 25. h (0,50%); 26. ŝ (0,35%); 27. ĵ (0,15%); 28. ĥ (0,01%). Andere Untersuchungen (Jung 1926; Stancliff 1933; Sadler 1959; Harry 1967; Dominte 2001) ergaben ähnliche Ergebnisse.
- (c) Nach Dominte (2001) kann Esperanto mit seinem durchschnittlichen Vokal-zu-Konsonant-Verhältnis von 43 zu 57 als relativ vokalische Sprache betrachtet werden (mit einem Platz zwischen den vorwiegend vokalischen Sprachen Finnisch, Türkisch, Italienisch, Neugriechisch, Serbisch, Kroatisch, Rumänisch, Litauisch, Latein, Slowakisch, Französisch und den eher konsonantischen Sprachen Ungarisch, Schwedisch, Tschechisch, Deutsch und Amerikanisches Englisch).
- (d) Mit dem Zipf- und Naranan/Balasubrahmanyam-Modell ergibt sich eine verhältnismäßig schlechte Anpassung (Bestimmtheitsmaß: 0,82 bzw.

---

<sup>64</sup> Siehe Fußnote 2.

0,92). Ein mit dem letztgenannten Wert vergleichbarer Wert hatte sich mit diesen Modellen auch bezüglich des Hawaiischen ergeben (nach Strauss, Altmann, Best 2007). Mit den anderen Modellen ergeben sich bezüglich des Esperanto Anpassungen mit Bestimmtheitsmaßen von 0,96 und 0,98. Die beste Anpassung ergibt sich mit dem Yule-Modell (Bestimmtheitsmaß: 0,98; Abweichungsquadratsumme: 5,85): relative Phonemhäufigkeit (in %)  $y = 13,3814 x^{0,0361} - 0,8805x$  (x: Rang). Auch bezüglich des Hawaiischen hatte sich die beste Anpassung mit dem Yule-Modell ergeben (Bestimmtheitsmaß: 0,94). Das Ergebnis ist ebenfalls in Einklang mit dem Ergebnis einer aktuellen Untersuchung von 95 Sprachen (Tambovtsev, Martindale 2007), der zufolge das Yule-Modell die beste Anpassung an die Phonemhäufigkeitsverteilung liefert.

- (e) Unter Zugrundelegung des Yule-Modells und der Phonemhäufigkeiten von Haszpra (2001) ergibt sich bei biphonematischer Wertung der drei Affrikaten *c*, *ĉ* und *ĝ* eine etwas schlechtere Anpassung als bei monophonematischer Wertung.

Die o. g. Hypothese wurde bestätigt. Dies war zu erwarten, weil es sich bei Esperanto um eine aposteriorische Plansprache handelt und es somit naheliegend ist, dass Gesetzmäßigkeiten, die für Ethnosprachen gelten, auch für Esperanto gelten.

### 5.3 Morphologie und Sprachtypologie

Esperanto weist in erster Linie synthetischen Sprachbau auf. Dabei dominiert Agglutination im Vergleich zu inkorporierenden, polysynthetischen und kaum vorhandenen typisch flektierenden Merkmalen. Merkmale isolierenden Sprachbaus sind festzustellen. Diese Aussagen wurden durch eine quantitative morphologische Untersuchung einer kurzen Esperanto-Textprobe (Zamenhof 1991<sup>10</sup>: 107-108) untermauert: Es ergab sich unter Verwendung des klassifikatorischen Schlüssels von Altmann, Lehfeldt (1973: 47) und Kempgen, Lehfeldt (2004: 1243) eine Einordnung als Sprache der Klasse D, dessen Repräsentanten mündliches Türkisch, Bengali, Neuenglisch, Neupersisch und Neugriechisch sind. – Alles in allem wird das Resümee von Sikosek (2003<sup>2</sup>: 204) bestätigt, dem zufolge nach Meinung der meisten Autoren



und in passendster Verallgemeinerung Esperanto agglutinierend ist und einige Merkmale einer isolierenden Sprache aufweist.

Des Weiteren wurde die Modellierbarkeit der Häufigkeitsverteilung der Morphemstrukturen des Esperanto mittels des Zipfschen Gesetzes detailliert.

## 5.4 Syntax

Die Rechts- und Linkseinbettung von verschiedenartigen Phrasen und Relativsätzen im Esperanto wurde dargestellt. Dazu wurden einige Beispielsätze in X-Bar-Theorie analysiert. Beachtenswert ist, dass Esperanto als Plansprache zwar eine präskriptive Grundlage in Zamenhof (1991<sup>10</sup>) hat, es darin jedoch keine unmittelbaren expliziten Regeln zur Einbettung von Phrasen und Relativsätzen gibt. Etwaige diesbezügliche Regeln oder Empfehlungen können nur mittelbar aus Beispielsätzen hergeleitet werden.

Fazit:

- (a) Subjekt-Nominalphrasen sind überwiegend links in der Flexionsphrase, Objekt-Nominalphrasen überwiegend rechts in der Verbalphrase eingebettet. Daher kann die Reihenfolge Subjekt-Verb-Objekt als Basis-Wortstellung im Esperanto identifiziert werden (vgl. Jansen 2007: 81, Gledhill 2000<sup>2</sup>: 87 und Wells 1978: 41). Bei einem Kopulativverb ist die Linkseinbettung der Nominalphrase praktisch obligatorisch (Wells 1978: 47).
- (b) Der Artikel ist links in der Nominalphrase eingebettet (identifizierte Ausnahme: Rechtseinbettung nach Personennamen vor Ordnungszahl/Adjektiv [Wells 1978: 47]).
- (c) Adjektivphrasen sind überwiegend links in Nominalphrasen eingebettet (Rechtseinbettung vor allem bei Emphase, Euphonie, Korrelativwort, Komplexität [Jansen 2007: 112]).
- (d) Präpositionalphrasen sind überwiegend links in Verbalphrasen eingebettet.
- (e) Keine Vorzugsrichtung bezüglich Einbettung konnte für Adverbphrasen in Verbalphrasen gefunden werden.

(f) Negationen ohne Kontrast sind (ausnahmslos) links eingebettet; mit Kontrast dagegen überwiegend rechts.

(g) Relativsätze sind rechts eingebettet.

## 5.5 Lexik

Die Zahl der grundlegenden und offiziellen Wörter des Esperanto in Abhängigkeit vom Rang in der Offizialisierungsfolge kann mit dem Zipfschen Gesetz und auch in Anlehnung an das Piotrowski-Gesetz beschrieben werden.

## 5.6 Semantik

Charakteristische Merkmale der lexikalischen Semantik der Plansprache Esperanto wurden dargestellt. Eine Besonderheit des Esperanto besteht darin, dass die Zuordnung von Bedeutungen zu den einzelnen Lexemen, die weitestgehend aposteriorisch aus Lexemen von Ethnosprachen abgeleitet wurden, zumindest am Anfang zunächst präskriptiv erfolgte, und zwar durch die Grundlagerschrift *Fundamento de Esperanto* (Zamenhof 1991<sup>10</sup>). Die bei anderen Sprachen festgestellte Gesetzmäßigkeit zwischen der Länge von Lexemen und Polysemie gilt offensichtlich auch für Esperanto (dies wurde anhand einer Untersuchung einer Stichprobe von 100 Esperanto-Lexemen gezeigt).

## 5.7 Worthäufigkeit

Es wurde gezeigt:

Das Häufigkeitsmerkmal eines Wortes, das durch seine Zugehörigkeit zu einer der neun Häufigkeitsgruppen BRO1 bis BRO9 (resultierend aus einer Metaanalyse von bereits bestehenden Worthäufigkeitslisten nach Akademio de Esperanto [Hrsg.] 2007<sup>2</sup>: 23ff.) definiert ist, korreliert mit einer Häufigkeitsverteilung, die mit dem Zipfschen Gesetz modellierbar ist.

## 5.8 Wortbekanntheit

Es wurden untersucht:

(a) der Zusammenhang zwischen Bekanntheit und Häufigkeit von Esperanto-Wörtern,

(b) die Selbsteinschätzung des eigenen Esperanto-Anwendungsvermögens durch seine Sprecher.

Dazu wurde eine statistische Erhebung in Form einer Befragung durchgeführt und mit dem mathematischen Verfahren der logistischen Regression ausgewertet. Dabei wurde aus Daten dichotomer Ausprägung (z. B. „Wort bekannt“ oder „Wort nicht bekannt“) eine graduelle Ausprägung der (a) Wortbekanntheit in Abhängigkeit von der Worthäufigkeit und (b) Selbsteinschätzung des eigenen Esperanto-Anwendungsvermögens in Abhängigkeit von der Esperanto-Lernzeit ermittelt. Dabei wurde auch das Häufigkeitsmerkmal der Wörter mit zuverlässiger Bekanntheit bestimmt. Für solche Wörter gilt:

Ihr Bekanntheitsgrad ist 90% bei einer statistischen Sicherheit von 95%. Das bedeutet: Bei mindestens 95 von 100 Texten sind je mindestens 90% dieser Wörter solchen Esperanto-Sprechern bekannt, die nach eigener Einschätzung Esperanto ausreichend gut verwenden.

Ein analoges Merkmal wurde bezüglich der Selbsteinschätzung des eigenen Esperanto-Anwendungsvermögens durch seine Sprecher ermittelt – und zwar die Esperanto-Lernzeit für ein zuverlässiges Anwendungsvermögen (also 90% bei 95% Vertrauensgrad).

Die Methode zum Ermitteln des gewünschten Zusammenhangs wurde dargestellt. Es handelt sich um ein logistisches Regressionsmodell. Es war zuvor in der Werkstoffprüftechnik zur Bestimmung von Riss-Auffindwahrscheinlichkeiten (bei vorgegebener statistischer Sicherheit) aus dichotomen Stichprobendaten („Riss gefunden“ oder „Riss nicht gefunden“) angewendet worden (Berens 1989). In innovativer Weise wurde dieses Modell auf die o. g. dichotomen Fragestellungen der Sprachwissenschaft übertragen.

Zu den nachfolgend beschriebenen Haupthypothesen kann auf Grundlage der Untersuchung Folgendes ausgesagt werden (es sei angemerkt, dass die Haupthypothesen (a) und (b) dabei ausschließlich für solche Esperanto-Sprecher gelten, die ihr Esperanto-Anwendungsvermögen als ausreichend gut einschätzen; die Haupthypothese (c) gilt auch für solche Esperanto-Sprecher, die ihr Anwendungsvermögen als nicht ausreichend gut einschätzen):

- (a) Es wurde bestätigt: Die Häufigkeit der zuverlässig bekannten Wörter ist nicht zufällig, sondern vorhersagbar. Sie entspricht genau einer wohl definierten Häufigkeitsgruppe, nämlich der Häufigkeitsgruppe Sieben auf einer neunstufigen Skala von Häufigkeitsgruppen der von der Esperanto-Akademie entwickelten Statistik („offizielle Basis-Wortwurzelsammlung“, BRO).
- (b) Sei A die Häufigkeit der Esperanto-Wörter, die Menschen mit einer Muttersprache aus dem romanischen oder germanischen Sprachzweig zuverlässig bekannt sind. Sei B die Häufigkeit der Esperanto-Wörter, die Menschen mit Muttersprache Chinesisch, Japanisch oder aus der finnougri-schen Sprachfamilie zuverlässig bekannt sind. Die Untersuchung zeigte, dass kein signifikanter Unterschied zwischen A und B besteht. Das heißt: Die naheliegende Vermutung, dass Sprecher mit einer Muttersprache aus dem romanischen oder germanischen Sprachzweig einen Vorteil haben und seltenere Esperanto-Wörter vergleichsweise besser kennen als Sprecher mit anderen Muttersprachen, wurde durch die vorliegende empirische Überprüfung nicht bestätigt.
- (c) Es wurde bestätigt: Bei einer statistischen Sicherheit von 95% ist die Selbsteinschätzung des eigenen Esperanto-Anwendungsvermögens durch seine Sprecher auch nach 100 Lernjahren 73%. Interessant ist, dass keine Lernzeit für eine zuverlässige Anwendbarkeit (also 90% bei 95% Vertrauensgrad) ermittelt werden kann, also „ewig“ gelernt muss: Auf der Grundlage der wie beschrieben statistisch untersuchten Stichprobe besteht die Esperanto-Sprecherschaft also aus Leuten, die sich in Bezug auf Esperanto als „lebenslang Lernende“ sehen (was übrigens auch der subjektiven Erfahrung des Autors nicht widerspricht).

## 5.9 Ausblick

Da die Tauglichkeit des hier herangezogenen Modells der logistischen Regression statistisch fundiert gezeigt werden konnte (durch den Devianz-Test und einen Vergleich mit einer alternativen Beschreibung), wird es hiermit als eine innovative Option zur Untersuchung von Fragestellungen wie der vorliegenden oder ähnlichen angeboten; Beispiele:

- (a) Wortvertrautheit und -häufigkeit bei anderen Sprachen als Esperanto
- (b) Vertrautheit von Phraseologismen
- (c) Selbsteinschätzung des eigenen Anwendungsvermögens von anderen Sprachen als Esperanto

Der ausführlich dargestellte Rechenalgorithmus, der in der vorliegenden Arbeit durch ein Standard-Tabellenkalkulationsprogramm umgesetzt wurde, kann für andere Programmierumgebungen leicht angepasst werden.

## Anhang A: Anpassung des Yule-Modells (monophonematische Wertung von Affrikaten)

Siehe Tabelle 16.

Rang x	Phonem	f <sub>x</sub>	Modell y	Matrix R		Matrix J (Differenzenquot.)			Diff. für Diffe- renzenquot.	(f <sub>x</sub> - <f <sub>x</sub> >) <sup>2</sup>
				r = f <sub>x</sub> · y	r <sup>2</sup>	dr/da	dr/db	dr/dd		
1	a	12,59	11,78	0,81	0,65	-0,88	0,00	-13,38	0,000001	81,33
2	e	9,36	10,64	-1,28	1,63	-0,79	7,37	-24,16		33,51
3	i	8,99	9,50	-0,51	0,26	-0,71	10,44	-32,38		29,36
4	o	8,80	8,46	0,34	0,12	-0,63	11,72	-38,42		27,34
5	n	7,79	7,51	0,28	0,08	-0,56	12,08	-42,62		17,80
6	l	6,24	6,65	-0,41	0,17	-0,50	11,92	-45,34		7,12
7	r	5,97	5,89	0,08	0,01	-0,44	11,46	-46,83		5,75
8	s	5,91	5,21	0,70	0,49	-0,39	10,84	-47,35		5,47
9	t	5,50	4,61	0,89	0,79	-0,34	10,13	-47,11		3,72
10	k	4,22	4,07	0,15	0,02	-0,30	9,38	-46,26		0,42
11	u	3,36	3,60	-0,24	0,06	-0,27	8,63	-44,96		0,04
12	m	3,06	3,18	-0,12	0,01	-0,24	7,90	-43,33		0,26
13	d	3,04	2,81	0,23	0,05	-0,21	7,20	-41,45		0,28
14	p	2,82	2,48	0,34	0,12	-0,19	6,54	-39,41		0,56
15	j	2,71	2,19	0,52	0,27	-0,16	5,92	-37,27		0,74
16	v	1,87	1,93	-0,06	0,00	-0,14	5,35	-35,09		2,89
17	g	1,28	1,70	-0,42	0,18	-0,13	4,83	-32,90		5,25
18	b	1,15	1,50	-0,35	0,12	-0,11	4,35	-30,73		5,86
19	f	1,09	1,33	-0,24	0,06	-0,10	3,91	-28,62		6,16
20	c	0,85	1,17	-0,32	0,10	-0,09	3,51	-26,58		7,41
21	ğ	0,69	1,03	-0,34	0,12	-0,08	3,14	-24,61		8,30
22	ç	0,67	0,91	-0,24	0,06	-0,07	2,81	-22,74		8,42
23	z	0,52	0,80	-0,28	0,08	-0,06	2,52	-20,97		9,31
24	ž	0,51	0,71	-0,20	0,04	-0,05	2,25	-19,30		9,37
25	h	0,50	0,62	-0,12	0,02	-0,05	2,01	-17,73		9,43
26	š	0,35	0,55	-0,20	0,04	-0,04	1,79	-16,25		10,38
27	ž	0,15	0,49	-0,34	0,11	-0,04	1,60	-14,88		11,71
28	ñ	0,01	0,43	-0,42	0,17	-0,03	1,43	-13,61		12,68
		3,57			5,85				320,89	
									0,98	

Parameter	K = (J transp. mal J) <sup>-1</sup>		L = J transp. mal R			E = K mal L		Neues V ist V minus E
Bez.	Start S	Verbessert V						
a	13,3486	13,3814	0,86	0,05	0,00	0,0000000	0,0000014	13,3814
b	0,6000	-0,0361	0,05	0,01	0,00	0,0000775	-0,0000006	-0,0361
d	1,0000	0,8805	0,00	0,00	0,00	-0,0007773	-0,0000002	0,8805

Tabelle 16: Tabellenkalkulationsblatt zur Anpassung des Yule-Modells an die empirisch erhaltene relative Phonemhäufigkeit mit Hilfe des Gauß-Newton-Verfahrens (monophonematische Wertung von Affrikaten)

## Anhang B: Anpassung des Yule-Modells (biphonematische Wertung von Affrikaten)

Siehe Tabelle 17.

Rang x	Phonem	$f_x$	Modell y	Matrix R		Matrix J (Differenzenquot.)			Diff. für Diffe- renzenquot.	$(f_x - \langle f_x \rangle)^2$
				$r = \frac{f_x - y}{f_x}$	$r^2$	dr/da	dr/db	dr/dd		
1	a	12,59	11,59	1,00	1,01	-0,88	0,00	-13,19	0,000001	72,28
2	e	9,36	10,74	-1,38	1,90	-0,81	7,44	-24,45		27,79
3	i	8,99	9,73	-0,74	0,55	-0,74	10,69	-33,24		24,03
4	o	8,80	8,74	0,06	0,00	-0,66	12,12	-39,80		22,20
5	n	7,79	7,81	-0,02	0,00	-0,59	12,57	-44,46		13,70
6	t	7,02	6,96	0,06	0,00	-0,53	12,47	-47,52		8,59
7	s	6,76	6,18	0,58	0,33	-0,47	12,03	-49,28		7,14
8	l	6,24	5,49	0,75	0,56	-0,42	11,41	-49,99		4,63
9	r	5,97	4,86	1,11	1,22	-0,37	10,69	-49,85		3,54
10	k	4,22	4,31	-0,09	0,01	-0,33	9,92	-49,05		0,02
11	d	3,73	3,81	-0,08	0,01	-0,29	9,14	-47,74		0,13
12	u	3,36	3,37	-0,01	0,00	-0,26	8,38	-46,05		0,53
13	m	3,06	2,98	0,08	0,01	-0,23	7,64	-44,09		1,06
14	p	2,82	2,63	0,19	0,04	-0,20	6,95	-41,95		1,61
15	j	2,71	2,32	0,39	0,15	-0,18	6,29	-39,69		1,90
16	v	1,87	2,05	-0,18	0,03	-0,16	5,69	-37,37		4,92
17	g	1,28	1,81	-0,53	0,28	-0,14	5,13	-35,04		7,89
18	b	1,15	1,60	-0,45	0,20	-0,12	4,62	-32,73		8,63
19	f	1,09	1,41	-0,32	0,10	-0,11	4,15	-30,48		8,99
20	š	1,02	1,24	-0,22	0,05	-0,09	3,72	-28,29		9,42
21	ǰ	0,84	1,10	-0,26	0,07	-0,08	3,34	-26,19		10,55
22	z	0,52	0,97	-0,45	0,20	-0,07	2,99	-24,19		12,73
23	ü	0,51	0,85	-0,34	0,12	-0,06	2,67	-22,29		12,80
24	h	0,50	0,75	-0,25	0,06	-0,06	2,38	-20,50		12,88
25	ñ	0,01	0,66	-0,65	0,42	-0,05	2,13	-18,81		16,63
		4,09			7,32				294,59	
									0,98	

Parameter	K = (J transp. mal J) <sup>-1</sup>		Bestimmtheitsmaß		
Bez.	Start S	Verbessert V	L = J transp. mal R	E = K mal L	Neues V ist V minus E
a	13,3486	13,1917	0,85	0,05	0,00
b	0,6000	-0,0772	0,05	0,01	0,00
d	1,0000	0,8784	0,00	0,00	0,00

L = J transp. mal R		E = K mal L		Neues V ist V minus E
0,0000000	-0,0000001	13,1917		
-0,0000054	0,0000000	-0,0772		
0,0000535	0,0000000	0,8784		

**Tabelle 17: Tabellenkalkulationsblatt zur Anpassung des Yule-Modells an die empirisch erhaltene relative Phonemhäufigkeit mit Hilfe des Gauß-Newton-Verfahrens (biphonematische Wertung von Affrikaten)**

## Anhang C: Anpassung des Zipfschen Gesetzes an die Häufigkeitsverteilung von Morphemstrukturen des Esperanto

Siehe Tabelle 18.

Rang x	Häufigk. y	f	Matrix R		Matrix J (Differenzenquot.)		Differenz	(y-<y>)^2
			r=y-f	r^2	dr/da	dr/db		
1	12970,00	12969,46	0,54	0	-129681639	-129681638	0,000100	134158939
2	6005,00	4462,42	1542,58	2379563	-44619706	-44616612		21323153
3	1386,00	2390,71	-1004,71	1009440	-23904699	-23902073		2
4	800,00	1535,39	-735,39	540797	-15352352	-15350223		344921
5	795,00	1089,06	-294,06	86472	-10889525	-10887773		350819
6	642,00	822,57	-180,57	32607	-8224916	-8223443		555472
7	631,00	648,83	-17,83	318	-6487622	-6486360		571990
8	619,00	528,28	90,72	8230	-5282301	-5281202		590285
9	505,00	440,69	64,31	4136	-4406442	-4405474		778453
10	417,00	374,71	42,29	1788	-3746771	-3745908		941482
11	400,00	323,58	76,42	5839	-3235524	-3234748		974761
12	376,00	283,02	92,98	8645	-2829956	-2829253		1022728
13	359,00	250,22	108,78	11834	-2501918	-2501277		1057401
14	318,00	223,24	94,76	8979	-2232203	-2231614		1143402
15	317,00	200,75	116,25	13514	-2007307	-2006763		1145542
16	311,00	181,77	129,23	16701	-1817487	-1816983		1158422
17	254,00	165,57	88,43	7819	-1655561	-1655092		1284369
18	220,00	151,63	68,37	4675	-1516129	-1515691		1362589
19	215,00	139,52	75,48	5697	-1395063	-1394652		1374287
20	206,00	128,93	77,07	5940	-1289156	-1288770		1395470
	1387,30			4152993				171534488
							Bestmaß	0,98

P.	Start	Verbessert	K = (J transp. mal J)^-1		D = J transp. mal R	E = K mal D	Verbessert minus E
a	12970,0000	12969,4608	0,00	0,00	-31627274720,61420	-0,0000010	12969,4608
b	1,0000	1,5392	0,00	0,00	-31626801920,62490	-0,0000005	1,5392

**Tabelle 18: Tabellenkalkulationsblatt zur Anpassung des Zipfschen Gesetzes an die Häufigkeitsverteilung von Morphemstrukturen des Esperanto**



## Anhang D: Länge von Lexemen und Polysemie

Siehe Tabelle 19 bis 22 und Abbildung 44.

Nr.	Lexem	wesentl. Bedeutung im Dt.	Länge L	Polysemie P
1	ENTUZIASMO	Enthusiasmus	5	1
2	HEZITI	zögern	3	1
3	INTERPELACII	interpellieren	6	1
4	KARPENO	Hainbuche	3	1
5	KOKETA	kokett	3	1
6	KULPA	schuldig	2	1
7	MELEAGRO	Truthahn	4	1
8	PROFESIO	Beruf	4	1
9	POLENO	Pollen	3	1
10	RODODENDRO	Rhododendron	4	1
		Mittelwerte:	3,700	1,0
11	ARBITRACIO	Schiedspruch	5	2
12	DO	also	1	2
13	LACA	müde	2	2
14	OBSTRUKCO	Verstopfung	3	2
15	PEKI	sündigen	2	2
16	PLU	weiter	1	2
17	RIVERO	Fluss	3	2
18	SEN	ohne	1	2
19	STARTI	starten	2	2
20	VERKO	Werk	2	2
		Mittelwerte:	2,200	2,0
21	ARANGHI	anordnen	3	3
22	FOR	fort	1	3
23	HARO	Haar	2	3
24	INTERNO	Inneres	3	3
25	JEN	sieh da!	1	3
26	KIES	wessen	2	3
27	NAGHI	schwimmen	2	3
28	OFERI	opfern	3	3
29	PLUS	plus	1	3
30	POMO	Apfel	2	3
31	PRI	über	1	3
32	PROKSIMA	nah	3	3
33	SELO	Sattel	2	3
34	SHATI	schätzen	2	3
		Mittelwerte:	2,000	3,0

*Tabelle 19: Länge und Polysemie von 100 Lexemen  
(Teil 1 von 3)*

35	BRECHO	Bresche	2	4
36	DENTO	Zahn	2	4
37	GUSTO	Geschmack	2	4
38	KADRO	Rahmen	2	4
39	LABORI	arbeiten	3	4
40	METI	stellen	2	4
41	MEZO	Mitte	2	4
42	OKAZO	Ereignis	3	4
43	OL	als	1	4
44	PENTRI	malen	2	4
45	PLUMO	Feder	2	4
46	RILATI	sich beziehen	3	4
47	ROKO	Fels	2	4
48	RONDO	Kreis	2	4
49	SHARGI	laden	2	4
50	TERO	Erde	2	4
		Mittelwerte:	2,125	4,0
51	HOMO	Mensch	2	5
52	KORPO	Körper	2	5
53	KRUSTO	Kruste	2	5
54	LIGI	verbinden	2	5
55	MALVASTA	eng	3	5
56	OKULO	Auge	3	5
57	PERDI	verlieren	2	5
58	SEKSO	Geschlecht	2	5
59	SUFERI	leiden	3	5
60	VARI	variieren	3	5
		Mittelwerte:	2,400	5,0
61	BRULI	brennen	2	6
62	EL	aus	1	6
63	FORTA	stark	2	6
64	GHIRI	überweisen	2	6
65	HAVI	haben	2	6
66	HORO	Stunde	2	6
67	JE	{allg. Pröp.}	1	6
68	KUN	mit	1	6
69	KUSHI	liegen	2	6
70	LEVI	heben	2	6
71	NAZO	Nase	2	6
72	PER	mittels	1	6
73	SENDI	senden	2	6
74	SUR	auf	1	6
		Mittelwerte:	1,643	6,0

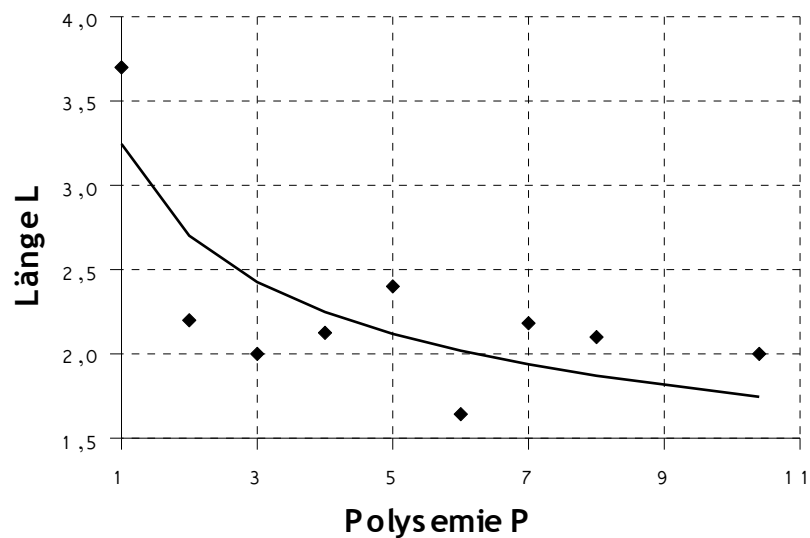
Tabelle 20: Länge und Polysemie von 100 Lexemen  
(Teil 2 von 3)

75	KOLEKTI	sammeln	3	7
76	KOVRI	bedecken	2	7
77	LUDI	spielen	2	7
78	OMBRO	Schatten	2	7
79	PELI	treiben	2	7
80	PREZENTI	präsentieren	3	7
81	REKTA	gerade	2	7
82	ROMPI	brechen	2	7
83	SEKA	trocken	2	7
84	VIDI	sehen	2	7
85	VERDA	grün	2	7
		Mittelwerte:	2,182	7,0
86	EN	in	1	8
87	IRI	gehen	2	8
88	KOLONO	Säule	3	8
89	KRUCO	Kreuz	2	8
90	LIBRO	Buch	2	8
91	POR	für	1	8
92	RADIKO	Wurzel	3	8
93	RADIO	Strahl	3	8
94	RINGO	Ring	2	8
95	STELO	Stern	2	8
		Mittelwerte:	2,100	8,0
96	PIEDO	Fuß	3	9
97	AL	zu	1	10
98	LIBERA	frei	3	10
99	ORDO	Ordnung	2	10
100	DE	von	1	13
		Mittelwerte:	2,000	10,4

*Tabelle 21: Länge und Polysemie von 100 Lexemen  
(Teil 3 von 3)*

Nr.	L (Mittelwerte)	P (Mittelwerte)	$L_{th} = C \cdot P^b$	C	b
				3,2450	0,2649
1	3,700	1,0	3,245	0,207	2,071
2	2,200	2,0	2,701	0,251	0,004
3	2,000	3,0	2,426	0,181	0,068
4	2,125	4,0	2,248	0,015	0,019
5	2,400	5,0	2,119	0,079	0,019
6	1,643	6,0	2,019	0,141	0,382
7	2,182	7,0	1,938	0,059	0,006
8	2,100	8,0	1,871	0,053	0,026
9	2,000	10,4	1,745	0,065	0,068
Summe:				1,051	2,663
Best'maß D:				0,61	

*Tabelle 22: Länge und Polysemie von 100 Lexemen;  
Auswertung; die Anpassung erfolgte mittels  
OpenOffice Calc 2.3, Zusatzprogramm („Extension“)  
EuroSolver Optimierer 1.6 (Option: „Das nichtlineare Verfahren“)*



*Abbildung 44: Länge von Lexemen als Funktion der  
Polysemie im Esperanto*

## Anhang E: HTML-Quellcode des Fragebogens A

Der nachfolgend auszugsweise wiedergegebene HTML-Quellcode erzeugt Fragebogen A (vgl. Kapitel 4.5, Abbildung 27).

```
<!doctype html public "-//W3C//DTD HTML 4.01 Transitional//EN">
<head>
<meta http-equiv="Content-Type"
content="text/html; charset=utf-8">
<meta name="Author" content="Andreas Kueck">
<meta name="generator" content="Ulli Meybohms HTML EDITOR">
<title>Enketo</title>
</head>
<body bgcolor="#FFFFFF" text="#000000" link="#0000EE" vlink="#551A8B" alink="#0000EE">

<div align="center"><a href="ak.htm">&#264;efpa&#285;o</a></div>
<p>
<div align="center"><h2>Enketo</h2>
<h3>Por eltrovi la konatecgradon de Esperanto-vortoj</h3>
<h4>Andreas K&#252;ck, Rechtenfleth, Germanujo</h4>
</div>
<b>Enkonduko</b>
<p>
Per tiu &#265;i enketo la a&#365;toro celas eltrovi la konatecgradon de Esperanto-vortoj la&#365;
speciala statistika metodo. Tiucele via helpo estas bezonata. Bonvole plenigu jenan enketilon kaj &#285;in
sendu al la a&#365;toro, premante la butonon &#8220;Sendu&#8221; en la fino de tiu &#265;i
interretpa&#285;o. Viaj nomo kaj retpo&#349;tadreso estos nek pludonataj nek publikigataj. La
a&#365;toro anticipe dankas.
<p>
<b>Enketilo</b>
<p>
<form action="http://www.akueck.de/cgi-bin/enketo01.cgi" method="post">

Mia nomo estas <input size="20" maxlength="80" name="Mia nomo estas">. Mia retpo&#349;tadreso
estas <input size="20" maxlength="80" name="Mia retposhtadreso estas">.<br>
Mia gepatra lingvo estas <input size="20" maxlength="80" name="Mia gepatra lingvo estas">. Mi
komencis lerni Esperanton en la jaro <input size="4" maxlength="4" name="Mi komencis lerni Esperanton
en la jaro">.
<p>
Kiu el la du sekvaj ebla&#309;oj trafas la&#365; vi pri via uzo de Esperanto?<br>
<input type="radio" name="Jeno trafas lau mi" value="Mi scipovas uzi Esperanton sufiche bone.">Mi
scipovas uzi Esperanton sufi&#265;e bone.<br>
<input type="radio" name="Jeno trafas lau mi" value="Mi ne scipovas uzi Esperanton sufiche bone.">Mi ne
scipovas uzi Esperanton sufi&#265;e bone.
<p>
Jen cent Esperanto-vortoj. Bonvole marku apud &#265;iu el ili, &#265;u vi scias &#285;ian signifon
(a&#365; &#8220;jes&#8221; a&#365; &#8220;ne&#8221;).<br>
Bonvole ne rigardu en vortaron kaj ne uzu la helpon de io alia a&#365; de iu alia, sed decidu tuj la&#365;
via unua impresio. (Vi ne bezonas doni tradukojn a&#365; difinojn.)
<p>
<table border="1">
<tr>
<td>
<table border="0">
<tr>
<td><b>alia</b></td><td><input type="radio" name="alia" value="jes">jes</td><td><input
type="radio" name="alia" value="ne">ne</td>
</tr>
<tr>
<td><b>fantazio</b></td><td><input type="radio" name="fantazio" value="jes">jes</td><td><input
type="radio" name="fanatazio" value="ne">ne</td>
</tr>
</table>
</td>
</tr>
</table>
```

[und entsprechend für die weiteren Wörter dieser Spalte; vgl. Abbildung 27]

```

</td>
<td>
<table border="0">
<tr>
<td><b>formulo</b></td><td><input type="radio" name="formulo" value="jes">jes</td><td><input
type="radio" name="formulo" value="ne">ne</td>
</tr>
[und entsprechend für die weiteren Wörter dieser Spalte; vgl. Abbildung 27]
<tr>
<td><b>perlo</b></td><td><input type="radio" name="perlo" value="jes">jes</td><td><input
type="radio" name="perlo" value="ne">ne</td>
</tr>
</table>
</td>
<td>
<table border="0">
<tr>
<td><b>leporo</b></td><td><input type="radio" name="leporo" value="jes">jes</td><td><input
type="radio" name="leporo" value="ne">ne</td>
</tr>
[und entsprechend für die weiteren Wörter dieser Spalte; vgl. Abbildung 27]
<tr>
<td><b>kial</b></td><td><input type="radio" name="kial" value="jes">jes</td><td><input
type="radio" name="kial" value="ne">ne</td>
</tr>
</table>
</td>
<td>
<table border="0">
<tr>
<td><b>reno</b></td><td><input type="radio" name="reno" value="jes">jes</td><td><input
type="radio" name="reno" value="ne">ne</td>
</tr>
[und entsprechend für die weiteren Wörter dieser Spalte; vgl. Abbildung 27]
<tr>
<td><b>kolektiva</b></td><td><input type="radio" name="kolektiva" value="jes">jes</td><td><input
type="radio" name="kolektiva" value="ne">ne</td>
</tr>
</table>
</td>
<td>
<table border="0">
<tr>
<td><b>ies</b></td><td><input type="radio" name="ies" value="jes">jes</td><td><input type="radio"
name="ies" value="ne">ne</td>
</tr>
[und entsprechend für die weiteren Wörter dieser Spalte; vgl. Abbildung 27]
<tr>
<td><b>agento</b></td><td><input type="radio" name="agento" value="jes">jes</td><td><input
type="radio" name="agento" value="ne">ne</td>
</tr>
<tr>
<td><b>ami</b></td><td><input type="radio" name="ami" value="jes">jes</td><td><input
type="radio" name="ami" value="ne">ne</td>
</tr>
</table>
</td>
</tr>
</table>
<p>
Mi havas jenan komenton: <input size="100" maxlength="400" name="Mi havas jenan komenton">.
<p>
<input type="submit" value="Sendu">
</form>
</body>
</html>

```

## Anhang F: Einzelheiten der Befragungsergebnisse

Die nachfolgenden Tabellen 23 bis 37 zeigen Einzelheiten der Befragungsergebnisse.

Hinweise:

- (a) In der Spalte „Häufigkeitsmerkmal“ steht unterhalb der Zahl, die das Häufigkeitsmerkmal angibt, die Anzahl der abgefragten Wörter des betreffenden Häufigkeitsmerkmals.
- (b) Die Esperanto-Wörter sind (anders als in den Fragebogen) im grundlegenden Ersatzschreibsystem<sup>65</sup> nach Zamenhof (1991<sup>10</sup>) geschrieben, d. h. der Zirkumflex ist durch ein nachgestelltes h ersetzt, und der Brevis ist weggelassen.
- (c) Die angegebenen deutschen Übersetzungen entsprechen nicht unbedingt den Übersetzungen, die ggf. in den Offizialisierungsdokumenten stehen, und es können durchaus andere oder weitere Übersetzungen zutreffen. Der Autor der vorliegenden Arbeit griff hier teilweise auf Wörterbücher (Krause 1999, Langenscheidt-Verlag [Hrsg.] 1959) zurück. Die Übersetzungen sind als Anhaltspunkte für die Bedeutung der abgefragten Esperanto-Wörter zu sehen.

Die sich an die genannten Tabellen anschließenden Abbildungen 45 bis 58 geben die Auswertung der Fragebogen mittels des Tabellenkalkulationsprogramms OpenOffice.org Calc wieder. Die Bedeutung der Spalten ist anhand der Spaltenüberschriften weitgehend selbsterklärend. (Die programmierten mathematischen Zusammenhänge sind in Anhang G ausführlich dargestellt.) Die letzten drei Spalten der jeweiligen Auswertungen in den Abbildungen 45 bis 58 enthalten folgende Größen:

Spalte  $p_7$ : Anteil der gekannten Wörter (berechnet nach Kapitel 4.4.2, Schritt 3)

Spalte  $p_{95\%}$ : Bekanntheitsgrad bei Vertrauensgrad 95% im Modell der logistischen Regression (also nach Gleichung 25 berechnet)

---

<sup>65</sup> Der Begriff „Fundamenta anstataua skribsistemo“ wird in Corsetti et al. (2007) verwendet.

Spalte  $p_0$ : Bekanntheitsgrad ohne Vertrauensgrad im Modell der logistischen Regression (also nach Gleichung 15 berechnet)

In den einzelnen graphischen Darstellungen sind  $p_1$  durch gestrichelte farblose und  $p_{95\%}$  durch graue Säulen wiedergegeben.



	Befragter		A1j	A2j	A3j	A4j	A5j	A6j				
	Mutterspr.		da	fr	fr	hu	pl	ro				
	Lernbeginn		1998	2003	1985	1966	1989	1986				
Häufigkeitsmerkmal	Esperanto-Wort	Deutsche Übersetzung	Eintrag 1 wenn nicht bekannt, sonst kein Eintrag									
1	agento	Agent										
30	avantagho	Vorteil										
	bulbo	Zwiebel	1									
	chelo	Zelle		1								
	daktilo	Dattel		1								
	ekvatoro	Äquator						1				
	fantazio	Phantasie										
	formulo	Formel	1									
	generacio	Generation										
	kategorio	Kategorie										
	kerno	Kern										
	malvo	Malve	1	1		1	1					
	periodo	Periode										
	perlo	Perle										
	reno	Niere		1								
	sago	Pfeil				1						
	singularo	Singular	1	1				1				
	teksto	Text										
	turismo	Tourismus										
	vulturo	Geier	1			1	1					
	zoologio	Zoologie										
	kolektiva	kollektiv										
	kaduka	gebrechlich				1	1					
	agiti	agitieren	1	1								
	dikti	diktieren	1									
	jungu	spannen	1					1				
	komponi	komponieren		1								
	kompromiti	bloßstellen	1	1								
	shpini	spinnen	1									
	chies	jedermanns								Bekannt	Nicht bek.	
			10	8	0	4	6	0		152	28	

Tabelle 23: Fragebogen A; Inhalt (Teil 1 von 3)

2	araneo	Spinne									
18	cerbo	Gehirn									
	ezoko	Hecht	1	1	1	1	1	1	1		
	foiro	Messe	1								
	kasó	Kasse									
	krepusko	Dämmerung				1					
	kreto	Kreide		1							
	masto	Mast						1	1		
	parko	Park									
	pozicio	Position									
	angulo	Winkel									
	identa	identisch		1						1	
	dependi	abhängen									
	komisii	beauftragen	1								
	remi	rudern									
	segi	sägen			1			1	1		
	shuti	schütten	1								
	ju	je (...desto)								1	
			4	3	2	2	3	5	89	19	
3	balono	Ballon									
12	brasiko	Kohl									
	direktoro	Direktor									
	komitato	Ausschuss									
	ledo	Leder								1	
	leporo	Hase									
	pizo	Erbse								1	
	acida	sauer									
	diveni	(er)raten									
	forghi	schmieden	1							1	
	petoli	übermütig sein	1							1	
	malgrau	trotz									
			2	0	0	0	0	4	66	6	

Tabelle 24: Fragebogen A; Inhalt (Teil 2 von 3)

4	kajero	Heft									
11	okcidento	Westen									
	pinto	Spitze									
	premio	Prämie									
	publiko	Öffentlichkeit									
	amaso	Masse									
	densa	dicht									
	anonci	ankündigen									
	imagi	sich vorstellen									
	informi	informieren									
	ies	jemandes									
			0	0	0	0	0	0	0	66	0
5	Julio	Juli									
6	jhaudo	Donnerstag									
	vendredo	Freitag									
	sagha	weise									
	atingi	erreichen									
	klini	neigen									
			0	0	0	0	0	0	0	36	0
6	celo	Ziel									
6	minuto	Minute									
	sango	Blut									
	aparta	getrennt									
	vendi	verkaufen									
	nul	null									
			0	0	0	0	0	0	0	36	0
7	domo	Haus									
5	urbo	Stadt									
	vera	wahr									
	kuri	laufen									
	lasi	lassen									
			0	0	0	0	0	0	0	30	0
8	resto	Rest									
5	tago	Tag									
	alia	anderer									
	ami	lieben									
	demandi	fragen									
			0	0	0	0	0	0	0	30	0
9	ok	acht							1		
7	ankau	auch									
	apenau	kaum									
	preskau	fast									
	kial	warum									
	chia	jeglicher									
	neniu	keiner									
			0	0	0	0	0	0	1	41	1

Tabelle 25: Fragebogen A; Inhalt (Teil 3 von 3)

	Befragter		B1j	B2j	B3j	B4j	B5j	B6j	
	Mutterspr.		de	de	fa	lt	lv	pt	
	Lernbeginn		1982	2003	1983	1990	1984	1999	
Häufigkeitsmerkmal	Esperanto-Wort	Deutsche Übersetzung	Eintrag 1 wenn nicht bekannt, sonst kein Eintrag						
1	alarmo	Alarm							
30	beto	Beete		1					
	figo	Feige		1					
	grego	Herde			1	1			
	kolombo	Taube							
	lino	Leinen		1					
	obstaklo	Hindernis							
	palto	Paletot, Überzieher	1	1		1			
	papilio	Schmetterling							
	pendolo	Pendel							
	plago	Plage				1			
	pruno	Pflaume							
	rimeno	Riemen		1					1
	romano	Roman							
	sino	Busen		1					
	spinaco	Spinat						1	
	splito	Splitter							
	tradicio	Tradition							
	tragedio	Tragödie							
	tubero	Knolle	1	1					
	vertico	Scheitelpunkt	1	1		1	1		
	amara	bitter							
	pia	fromm			1				
	aserti	behaupten							
	grumbli	nörgeln							
	inundi	überschwemmen							
	palpi	tasten		1					
	tinti	klirren							
	venghi	rächen							
	po	je							Bekannt
			3	9	2	4	2	1	159
									21

Tabelle 26: Fragebogen B; Inhalt (Teil 1 von 3)

2	abelo	Biene								
19	drapo	Tuch (Gewebe)	1	1			1			
	ekvilibro	Gleichgewicht								
	hajlo	Hagel								
	kliento	Kunde								
	kolono	Säule								
	krucho	Krug		1						
	pato	Pfanne		1				1		
	proceso	Vorgang								
	ripo	Rippe			1			1		
	trezoro	Schatz								
	trikoto	Trikot						1		
	sobra	nüchtern	1		1					
	dungi	anheuern								
	fleksi	biegen								
	konjekti	vermuten	1							
	perfidi	verraten	1							
	regali	bewirten	1							
	minus	minus								
			4	1	2	0	3	0	104	10
3	chifono	Lappen	1							
12	nebulo	Nebel								
	rizo	Reis								
	sezono	Jahreszeit								
	shvito	Schweiß								
	tolo	Leinwand	1	1			1	1		
	violo	Geige								
	milda	mild								
	esplori	erforschen								
	masoni	mauern	1	1						
	ornami	schmücken								
	malgrau	trotz								
			3	2	0	1	1	0	65	7

Tabelle 27: Fragebogen B; Inhalt (Teil 2 von 3)

4	kelo	Keller								
10	legomo	Gemüse								
	poto	Topf								
	rado	Rad								
	shuo	Schuh								
	stranga	seltsam								
	flegi	pflegen								
	posedi	besitzen								
	tremi	zittern								
	chiel	auf jede Weise						1		
			0	0	0	0	1	0	59	1
5	krajono	Bleistift								
6	ofico	Amt								
	stelo	Stern								
	bruna	braun								
	pendi	hängen								
	shlosi	zuschließen								
			0	0	0	0	0	0	36	0
6	fero	Eisen								
6	negho	Schnee								
	onklo	Onkel								
	sana	gesund								
	tushi	berühren								
	nul	null								
			0	0	0	0	0	0	36	0
7	akvo	Wasser								
5	tablo	Tisch								
	pura	rein								
	bati	schlagen								
	permesi	erlauben								
			0	0	0	0	0	0	30	0
8	chambro	Zimmer								
5	koro	Herz								
	lasta	letzter								
	fari	machen								
	helpi	helfen								
			0	0	0	0	0	0	30	0
9	anstatau	anstatt								
7	au	oder								
	cent	hundert								
	chiam	immer								
	ekster	außerhalb (von)								
	kiam	wann, wenn, als								
	kun	mit								
			0	0	0	0	0	0	42	0

Tabelle 28: Fragebogen B; Inhalt (Teil 3 von 3)

	Befragter		C1j	C2j	C3j	C4j	C5j	C6j				
	Mutterspr.		hu	ja	ja	ja	lt	zh				
	Lernbeginn		1988	1970	1963	1964	1985	1980				
Häufigkeitsmerkmal	Esperanto-Wort	Deutsche Übersetzung	Eintrag 1 wenn nicht bekannt, sonst kein Eintrag									
1	angilo	Aal						1	1			
30	bendo	Band						1				
	cirklo	Kreis										
	dando	Modenarr	1					1	1			
	etapo	Etappe										
	funelo	Trichter			1			1				
	gramo	Gramm										
	gvardio	Garde					1					
	hepato	Leber										
	karpo	Karpfen						1				
	kolumno	Spalte (Tabelle)										
	konkordo	Eintracht										
	lado	Blech						1				
	limako	Schnecke	1		1							
	mielo	Honig										
	profeto	Prophet										
	rafano	Rettich	1					1				
	ribo	Johannisbeere			1	1	1	1	1			
	sekcio	Abteilung										
	strigo	Eule						1				
	trupo	Truppe										
	graveda	schwanger										
	solena	feierlich										
	indulgi	verschonen										
	liveri	liefern						1				
	mencii	erwähnen										
	mughi	zwischen	1					1				
	sarki	jäten	1	1	1			1	1			
	vagi	herumschweifen										
	chiom	alles								Bekannt	Nicht bek.	
			5	1	4	2	12	4		152	28	

Tabelle 29: Fragebogen C; Inhalt (Teil 1 von 3)

2	arko	Bogen									
19	fulgo	Blitz			1	1	1	1			
	kastelo	Burg									
	komo	Komma									
	krado	Gitter	1					1			
	kubuto	Ellbogen						1	1		
	kupro	Kupfer	1								
	poezio	Poesie									
	rango	Rang						1			
	revuo	Zeitschrift									
	sojlo	Schwelle	1					1			
	vulpo	Fuchs									
	avida	gierig						1			
	boji	bellen						1			
	jhuri	schwören									
	machi	kauen						1			
	pretendi	beanspruchen									
	skui	schütteln						1			
	nenies	niemandes									
			3	0	1	1	9	2	98	16	
3	aventuro	Abenteuer									
12	nepo	Enkel									
	pajlo	Stroh						1			
	pinglo	Stecknadel						1			
	pugno	Faust						1			
	shultro	Schulter									
	vato	Watte	1	1				1			
	frandi	naschen	1					1			
	klera	gebildet									
	koncerni	betreffen									
	nutri	ernähren									
	adiau	leb wohl									
			2	1	0	0	5	0	64	8	

Tabelle 30: Fragebogen C; Inhalt (Teil 2 von 3)



4	frunto	Stirn					1		
10	kuvo	Wanne			1		1		
	nevo	Neffe					1		
	pilko	Ball							
	vango	Wange					1		
	honesto	ehrlich							
	balai	fegen							
	konduti	sich benehmen							
	shiri	reißen					1		
	ial	aus irgend-							
		einem Grund							
			0	0	1	0	5	0	54   6
5	azeno	Esel							
6	merkredo	Mittwoch							
	sako	Sack							
	serioza	ernst							
	kompati	bemitleiden					1		
	sukcesi	Erfolg haben							
			0	0	0	0	1	0	35   1
6	doloro	Schmerz							
6	nubo	Wolke					1		
	vagono	Waggon							
	komuna	gemeinsam							
	frapi	klopfen							
	nul	null							
			0	0	0	0	1	0	35   1
7	mondo	Welt							
5	sinjoro	Herr							
	multa	viel							
	ghoji	sich freuen							
	kompreni	verstehen							
			0	0	0	0	0	0	30   0
8	okulo	Auge							
5	valoro	Wert							
	afabla	freundlich							
	kosti	kosten							
	manki	fehlen							
			0	0	0	0	0	0	30   0
9	antau	vor							
7	da	von, an (Menge)							
	jen	sieh da, hier ist							
	kio	was							
	mil	tausend							
	tamen	allerdings							
	tiom	so viel							
			0	0	0	0	0	0	42   0

Tabelle 31: Fragebogen C; Inhalt (Teil 3 von 3)

	Befragter		D1j	D2j	D3j	D4j	D5j	D6j				
	Mutterspr.		et	et	et	fi	ja	zh				
	Lernbeginn		1967	1976	1956	1978	1960	1983				
Häufigkeitsmerkmal	Esperanto-Wort	Deutsche Übersetzung	Eintrag 1 wenn nicht bekannt, sonst kein Eintrag									
1	apro	Wildschwein						1				
30	atomo	Atom										
	dianto	Nelke					1	1				
	groto	Grotte										
	himno	Hymne										
	kagho	Käfig								1		
	korvo	Rabe										
	licenso	Lizenz								1		
	maleolo	Knöchel								1		
	menso	Geist										
	paliso	Pfahl			1		1	1	1	1		
	pasero	Spatz										
	pigo	Elster										
	puso	Eiter	1	1	1		1	1				
	sledo	Schlitten								1		
	sofo	Sofa					1					
	stadiono	Stadion										
	stano	Zinn					1	1				
	talio	Taille										
	vipa	Peitsche										
	volumeno	Volumen					1	1				
	tenera	zärtlich						1				
	vanta	eitel										
	dresi	dressieren										
	ekspluati	ausbeuten										
	indulgi	verschonen			1		1					
	luli	wiegen (kind)										
	prosperi	gedeihen										
	sorbi	schlüpfen					1	1	1	1		
	ties	dessen, deren										
			1	3	1	8	8	6	Bekannt	Nicht bek.		
									153	27		

Tabelle 32: Fragebogen D; Inhalt (Teil 1 von 3)

2	enigmo	Rätsel											
19	felo	Fell										1	
	frambo	Himbeere							1			1	
	kordo	Saite			1								
	listo	Liste											
	petrolo	Erdöl											
	profesio	Beruf											
	shelo	Schale										1	
	tendo	Zelt											
	traktoro	Traktor											
	velo	Segel		1									
	veshto	Weste											
	jhaluza	eifersüchtig											
	beni	segnen											
	oscedi	gähnen			1							1	
	shmiri	schmieren											
	shovi	schieben											
	velki	welken										1	
	plus	plus											
				1	2	0	0	1			5	105	9
3	chapo	Mütze											
12	deklivo	Abhang										1	
	fojno	Heu											
	larmo	Träne											
	nevo	Neffe						1					
	tombo	Grab											
	vosto	Schwanz											
	lama	lahm			1								
	distri	zerstreuen										1	
	peki	sündigen											
	surprizi	überraschen											
	des	desto											
				0	1	0	1	0			2	68	4

Tabelle 33: Fragebogen D; Inhalt (Teil 2 von 3)

4	kapro	Ziegenbock									
10	kondicho	Bedingung									
	oriento	Osten									
	pomo	Apfel									
	valo	Tal									
	ghentila	nett									
	gvidi	leiten									
	konfesi	gestehen									
	pafi	schießen									
	iel	irgendwie									
			0	0	0	0	0	0	60	0	
5	brusto	Brust									
6	mastro	Wirt									
	truo	Loch									
	ordinara	gewöhnlich									
	promeni	spazieren									
	pushi	stoßen									
			0	0	0	0	0	0	36	0	
6	avo	Großvater									
6	besto	Tier									
	stacio	Station									
	blua	blau									
	ludi	spielen									
	nul	null						1			
			0	0	0	0	0	1	35	1	
7	bildo	Bild									
5	monto	Berg									
	dolcha	süß									
	dormi	schlafen									
	rigardi	schauen									
			0	0	0	0	0	0	30	0	
8	kapo	Kopf									
5	tempo	Zeit									
	sama	gleicher									
	preni	nehmen									
	ridi	lachen									
			0	0	0	0	0	0	30	0	
9	dum	während									
7	el	aus									
	ie	irgendwo									
	inter	zwischen									
	kaj	und									
	kontrau	gegen									
	nau	neun									
			0	0	0	0	0	0	42	0	

Tabelle 34: Fragebogen D; Inhalt (Teil 3 von 3)

	Befragter		E1j	E2j	E3j	E4j	E5j	E6j		
	Mutterspr.		ja	ja	ja	pl	ja	fr		
	Lernbeginn		1957	1994	40?	1958	1994	2000		
Häufigkeitsmerkmal	Esperanto-Wort	Deutsche Übersetzung	Eintrag 1 wenn nicht bekannt, sonst kein Eintrag							
1	artiko	Gelenk	1		1				1	
30	aveno	Hafer	1	1	1			1	1	
	cepo	Zwiebel	1		1					
	cikonio	Storch		1				1		
	fajenco	Fayence	1	1	1			1		
	galosho	Galosche	1	1	1			1		
	ironio	Ironie								
	kranio	Schädel	1	1	1			1		
	kurento	Strom								1
	lako	Lack	1	1	1			1		
	logiko	Logik								
	nilono	Nylon			1					
	parashuto	Fallschirm								
	perono	Freitreppe	1	1	1			1	1	
	planedo	Planet								
	princo	Fürst								
	pruno	Pflaume	1	1				1		
	roso	Tau (Niedersch.)		1				1		
	rubando	Band								
	specimeno	Probe(körper)								
	traktato	Abhandlung								
	abrupta	schroff								
	malica	heimtückisch								
	aludi	anspielen								
	dortoti	verhätscheln								
	klopodi	sich bemühen								
	oponi	entgegensetzen	1	1				1		
	provoki	provizieren	1		1					
	shpruci	spritzen		1				1		
	chial	aus jedem Grund								
			11	11	10	0	11	4	Bekannt	Nicht bek.
									133	47

Tabelle 35: Fragebogen E; Inhalt (Teil 1 von 3)

2	apetito	Appetit								
19	brovo	Augenbraue								
	gloro	Ruhm								
	pejzagho	Landschaft								
	pelto	Pelz	1	1				1	1	
	peniko	Pinsel			1					
	plumbo	Blei	1	1				1		
	shtalo	Stahl		1				1		
	simio	Affe								
	tajloro	Schneider								
	tono	Ton								
	vergo	Rute	1	1	1	1	1	1		
	kruela	grausam								
	barakti	zappeln								
	konfidi	vertrauen								
	melki	melken		1	1			1	1	
	ravi	entzücken								
	sorchi	zaubern								
	nenies	niemandes								
			3	5	3	1	5	2	95	19
3	briko	Ziegel								
12	hauto	Haut								
	haveno	Hafen								
	kravato	Krawatte								
	martelo	Hammer								
	sablo	Sand								
	tritiko	Weizen								
	freneza	wahnsinnig								
	grimpi	klettern								
	krochi	anklammern			1					
	sterni	betten			1				1	
	des	desto								
			0	0	2	0	0	1	69	3

Tabelle 36: Fragebogen E; Inhalt (Teil 2 von 3)

4	graso	Fett								
10	najbaro	Nachbar								
	rando	Rand								
	sorto	Schicksal								
	sudo	Süden								
	vigla	rege								
	kisi	küssen								
	minaci	bedrohen								
	rezulti	resultieren								
	iel	irgendwie								
			0	0	0	0	0	0	60	0
5	klaso	Klasse								
6	limo	Grenze								
	shtupo	Stufe								
	mola	weich								
	bani	baden								
	plachi	gefallen								
			0	0	0	0	0	0	36	0
6	glaso	Glas (Gefäß)								
6	linio	Linie								
	sukero	Zucker								
	utila	nützlich								
	brili	glänzen								
	nul	null								
			0	0	0	0	0	0	36	0
7	lito	Bett								
5	rajto	Recht								
	nigra	schwarz								
	konsili	(be)raten								
	tiri	ziehen								
			0	0	0	0	0	0	30	0
8	flanko	Seite								
5	nacio	Nation								
	juna	jung								
	fini	(be)enden								
	sidi	sitzen								
			0	0	0	0	0	0	30	0
9	chirkau	um; gegen								
7	neniam	nie								
	plu	weiter								
	se	falls								
	super	über								
	tia	solcher								
	tri	drei								
			0	0	0	0	0	0	42	0

Tabelle 37: Fragebogen E; Inhalt (Teil 3 von 3)

**Bestimmung des Bekanntheitsgrades (Fragebogen A)**

Daten-Nr.	Häufigkeitsmerkmal a	Bekannt	Nicht bekannt	ln a	Hilfsgröße 1	Hilfsgröße 2	Für k <sub>0</sub>	Für k <sub>1</sub>	Für k <sub>2</sub>	ln L	∂lnL/∂μ	∂lnL/∂σ	ln L <sub>max</sub> (Summe)	-164,43	a	p <sub>1</sub>	p <sub>95%</sub>	p <sub>0</sub>
													∂ln L/∂μ (Summe)	0,0000				
1	1	152	28	0,00			-90,3	40,2	58,9	-78,4	5,8	0,0	Mittelwert μ <sub>1</sub> (Startwert)	0,1808	1	84%	72%	81%
2	2	89	19	0,69	1,2	0,4	-29,0	20,4	47,4	-54,2	-9,3	-6,5	Steilheit σ <sub>1</sub> (Startwert)	0,4134	2	82%	86%	91%
3	3	66	6	1,10	0,7	0,6	-12,6	10,8	30,4	-21,1	-1,9	-2,1	μ <sub>0</sub> (max. Mutmaßl.)	-1,1896	3	92%	90%	94%
4	4	66	0	1,39	0,6	0,7	-8,4	8,1	25,7	-2,7	2,7	3,7	σ <sub>0</sub> (max. Mutmaßl.)	1,4734	4	100%	92%	96%
5	5	36	0	1,61	0,4	0,7	-3,5	3,7	12,8	-1,1	1,1	1,8	Anzahl N d. Einzelgr.	600	5	100%	93%	97%
6	6	36	0	1,79	0,4	0,6	-2,9	3,2	11,7	-0,9	0,9	1,6	Vertrauensgrad V	95%	6	100%	94%	98%
7	7	30	0	1,95	0,3	0,6	-2,0	2,3	9,0	-0,6	0,6	1,2	γ(N; V)	5,138	7	100%	95%	98%
8	8	30	0	2,08	0,3	0,5	-1,7	2,1	8,4	-0,5	0,5	1,1	k <sub>0</sub>	0,2542	8	100%	96%	98%
9	9	41	1	2,20	0,2	0,5	-2,1	2,6	10,9	-4,8	-0,4	-0,8	k <sub>1</sub>	-0,2826	9	98%	96%	98%
Σ		546	54		4,0	4,6	-152,5	93,5	215,2	-164,4	0,0	0,0	k <sub>2</sub>	0,3587				

Anz. untersch. Einzelgr.	9
Devianz D	328,86
χ <sup>2</sup> (1-V; N-2)	656
Modell gut lt. D-Test?	ja

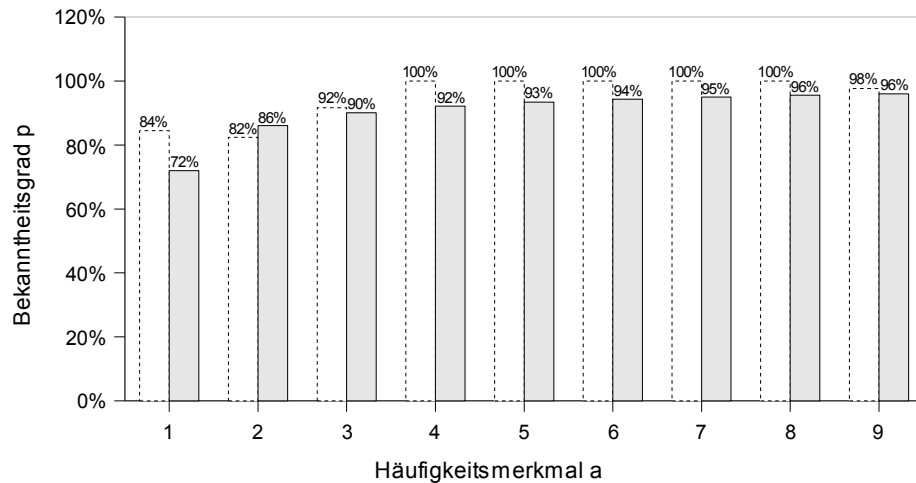


Abbildung 45: Auswertung von Fragebogen A



**Bestimmung des Bekanntheitsgrades (Fragebogen B)**

Da- ten- Nr.	Häufig- keits- merk- mal a	Be- kannt	Nicht be- kannt	ln a	Hilfs- grö- ße 1	Hilfs- grö- ße 2	Für k <sub>0</sub>	Für k <sub>1</sub>	Für k <sub>2</sub>	ln L	∂lnL/∂ μ	∂lnL/∂ σ	ln L <sub>max</sub> (Summe)		a	p <sub>1</sub>	p <sub>95%</sub>	p <sub>0</sub>	
													-132,69	0,0000					
													∂ln L/∂μ (Summe)						
													0,0000						
													∂ln L/∂σ (Summe)						
													0,0000						
1	1	159	21	0,00			-68,9	39,0	72,7	-65,1	3,2	0,0	Mittelwert μ <sub>1</sub> (Startwert)		0,1266	1	88%	78%	87%
2	2	104	10	0,69	1,2	0,4	-22,5	18,3	49,2	-34,4	-2,7	-1,9	Steilheit σ <sub>1</sub> (Startwert)		0,3696	2	91%	89%	94%
3	3	65	7	1,10	0,7	0,7	-9,3	8,9	28,1	-25,1	-4,1	-4,5	μ <sub>0</sub> (max. Mutmaßl.)		-1,5790	3	90%	92%	96%
4	4	59	1	1,39	0,5	0,7	-5,6	6,0	20,9	-5,3	0,8	1,1	σ <sub>0</sub> (max. Mutmaßl.)		1,5372	4	98%	94%	97%
5	5	36	0	1,61	0,4	0,7	-2,6	3,0	11,3	-0,8	0,8	1,3	Anzahl N d. Einzelgr.		600	5	100%	95%	98%
6	6	36	0	1,79	0,4	0,6	-2,1	2,6	10,3	-0,7	0,7	1,2	Vertrauensgrad V		95%	6	100%	95%	98%
7	7	30	0	1,95	0,3	0,6	-1,5	1,9	7,9	-0,5	0,5	0,9	γ(N; V)		5,138	7	100%	96%	98%
8	8	30	0	2,08	0,3	0,5	-1,3	1,7	7,3	-0,4	0,4	0,8	k <sub>0</sub>		0,1923	8	100%	96%	99%
9	9	42	0	2,20	0,2	0,5	-1,6	2,1	9,5	-0,5	0,5	1,1	k <sub>1</sub>		-0,2524	9	100%	97%	99%
Σ		561	39		4,1	4,7	-115,4	83,5	217,0	-132,7	0,0	0,0	k <sub>2</sub>		0,3617				
													Anz. untersch. Einzelgr.		9				
													Devianz D		265,39				
													χ <sup>2</sup> (1-V; N-2)		656				
													Modell gut lt. D-Test?		ja				

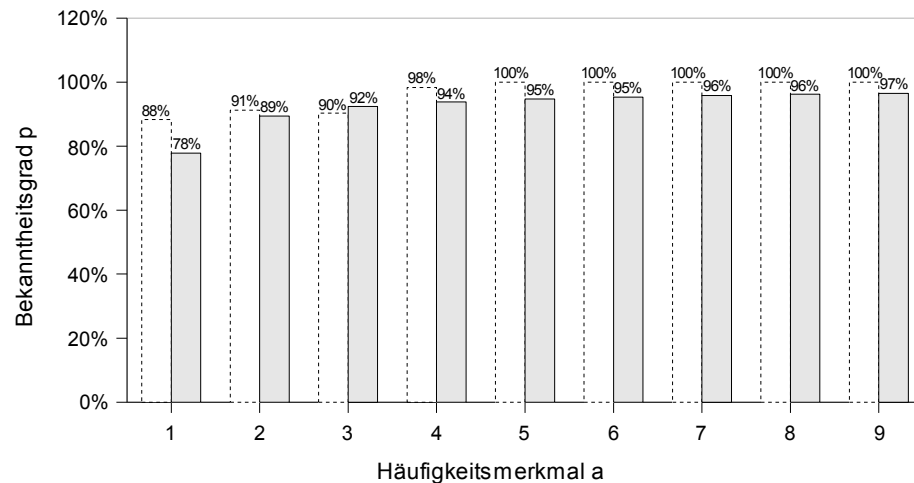


Abbildung 46: Auswertung von Fragebogen B

**Bestimmung des Bekanntheitsgrades (Fragebogen C)**

Da- ten- Nr.	Häufig- keits- merk- mal a	Be- kannt	Nicht be- kannt	ln a	Hilfs- grö- ße 1	Hilfs- grö- ße 2	Für k <sub>0</sub>	Für k <sub>1</sub>	Für k <sub>2</sub>	ln L	∂lnL/∂ μ	∂lnL/∂ σ	ln L <sub>max</sub> (Summe)	-183,87	a	p <sub>1</sub>	p <sub>95%</sub>	p <sub>0</sub>
													∂ln L/∂μ (Summe)	0,0000				
1	1	152	28	0,00			-87,5	40,3	61,0	-78,2	4,5	0,0	Mittelwert μ <sub>1</sub> (Startwert)	0,2054	1	84%	73%	82%
2	2	98	16	0,69	1,2	0,4	-35,3	23,0	49,2	-46,9	-4,0	-2,8	Steilheit σ <sub>1</sub> (Startwert)	0,5050	2	86%	84%	89%
3	3	64	8	1,10	0,7	0,6	-16,5	12,6	31,5	-25,7	-2,6	-2,8	μ <sub>0</sub> (max. Mutmaßl.)	-1,6763	3	89%	88%	92%
4	4	54	6	1,39	0,5	0,6	-11,0	9,2	25,6	-20,3	-2,5	-3,4	σ <sub>0</sub> (max. Mutmaßl.)	2,0080	4	90%	89%	94%
5	5	35	1	1,61	0,4	0,6	-5,5	5,0	14,7	-4,8	0,8	1,2	Anzahl N d. Einzelgr.	600	5	97%	91%	95%
6	6	35	1	1,79	0,4	0,6	-4,7	4,5	14,1	-4,7	0,5	0,9	Vertrauensgrad V	95%	6	97%	91%	96%
7	7	30	0	1,95	0,3	0,6	-3,5	3,5	11,3	-1,1	1,1	2,1	γ(N; V)	5,138	7	100%	92%	96%
8	8	30	0	2,08	0,3	0,5	-3,1	3,2	10,9	-1,0	1,0	2,0	k <sub>0</sub>	0,2852	8	100%	93%	97%
9	9	42	0	2,20	0,2	0,5	-3,9	4,2	14,6	-1,3	1,2	2,7	k <sub>1</sub>	-0,3186	9	100%	93%	97%
Σ		540	60		4,0	4,5	-171,1	105,4	233,0	-183,9	0,0	0,0	k <sub>2</sub>	0,3883				
														Anz. untersch. Einzelgr.	9			
														Devianz D	367,73			
														χ <sup>2</sup> (1-V; N-2)	656			
														Modell gut lt. D-Test?	ja			

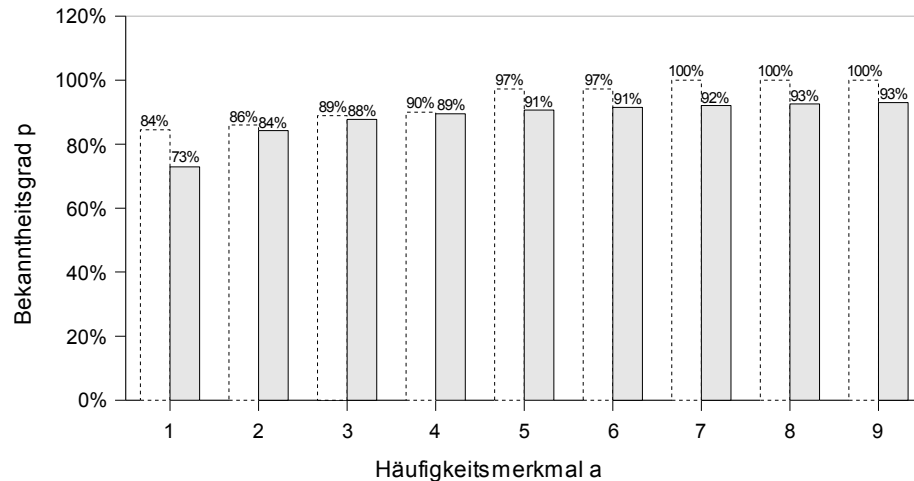


Abbildung 47: Auswertung von Fragebogen C

**Bestimmung des Bekanntheitsgrades (Fragebogen D)**

Da- ten- Nr.	Häufig- keits- merk- mal a	Be- kannt	Nicht be- kannt	ln a	Hilfs- grö- ße 1	Hilfs- grö- ße 2	Für k <sub>0</sub>	Für k <sub>1</sub>	Für k <sub>2</sub>	ln L	∂lnL/∂ μ	∂lnL/∂ σ	ln L <sub>max</sub> (Summe)		a	p <sub>1</sub>	p <sub>95%</sub>	p <sub>0</sub>	
													-131,34	0,0000					
													∂ln L/∂μ (Summe)						
													0,0000						
													∂ln L/∂σ (Summe)						
													0,0000						
1	1	153	27	0,00			-79,3	40,1	66,7	-76,1	1,7	0,0	Mittelwert μ <sub>1</sub> (Startwert)		0,1193	1	85%	75%	84%
2	2	105	9	0,69	1,2	0,4	-21,4	17,8	48,8	-31,8	-2,1	-1,4	Steilheit σ <sub>1</sub> (Startwert)		0,3239	2	92%	90%	94%
3	3	68	4	1,10	0,8	0,7	-7,6	7,8	26,3	-15,9	-1,6	-1,8	μ <sub>0</sub> (max. Mutmaßl.)		-1,0715	3	94%	93%	97%
4	4	60	0	1,39	0,6	0,7	-4,2	4,8	18,4	-1,3	1,3	1,8	σ <sub>0</sub> (max. Mutmaßl.)		1,1682	4	100%	95%	98%
5	5	36	0	1,61	0,4	0,7	-1,8	2,3	9,4	-0,6	0,6	0,9	Anzahl N d. Einzelgr.		600	5	100%	96%	98%
6	6	35	1	1,79	0,4	0,6	-1,4	1,8	8,2	-4,9	-0,6	-1,0	Vertrauensgrad V		95%	6	97%	97%	99%
7	7	30	0	1,95	0,3	0,6	-0,9	1,3	6,0	-0,3	0,3	0,5	γ(N; V)		5,138	7	100%	97%	99%
8	8	30	0	2,08	0,3	0,5	-0,7	1,1	5,3	-0,2	0,2	0,5	k <sub>0</sub>		0,1968	8	100%	98%	99%
9	9	42	0	2,20	0,2	0,5	-0,9	1,3	6,7	-0,3	0,3	0,6	k <sub>1</sub>		-0,2367	9	100%	98%	99%
Σ		559	41		4,2	4,7	-118,1	78,3	195,7	-131,3	0,0	0,0	k <sub>2</sub>		0,3262				
													Anz. untersch. Einzelgr.		9				
													Devianz D		262,68				
													χ <sup>2</sup> (1-V; N-2)		656				
													Modell gut lt. D-Test?		ja				

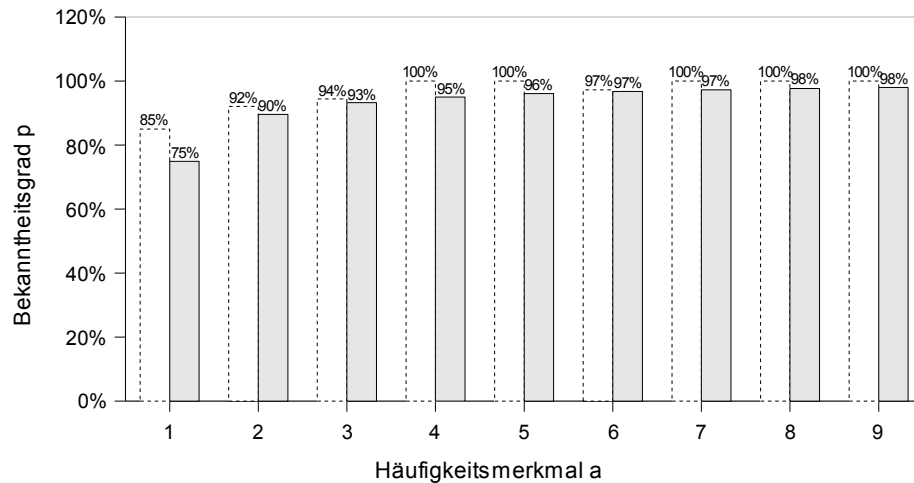


Abbildung 48: Auswertung von Fragebogen D

**Bestimmung des Bekanntheitsgrades (Fragebogen E)**

Da- ten- Nr.	Häufig- keits- merk- mal a	Be- kannt	Nicht be- kannt	ln a	Hilfs- grö- ße 1	Hilfs- grö- ße 2	Für k <sub>0</sub>	Für k <sub>1</sub>	Für k <sub>2</sub>	ln L	∂lnL/∂ μ	∂lnL/∂ σ	ln L <sub>max</sub> (Summe)	-173,78	a	p <sub>1</sub>	p <sub>95%</sub>	p <sub>0</sub>
													∂ln L/∂μ (Summe)	0,0000				
1	1	133	47	0,00			-121,2	33,5	30,5	-103,7	4,6	0,0	Mittelwert μ <sub>1</sub> (Startwert)	0,1965	1	74%	62%	71%
2	2	95	19	0,69	1,1	0,4	-32,1	22,0	49,8	-54,3	-8,2	-5,7	Steilheit σ <sub>1</sub> (Startwert)	0,3465	2	83%	85%	91%
3	3	69	3	1,10	0,7	0,7	-10,2	9,5	28,9	-12,5	0,3	0,3	μ <sub>0</sub> (max. Mutmaßl.)	-0,4673	3	96%	91%	95%
4	4	60	0	1,39	0,6	0,7	-5,1	5,5	20,0	-1,6	1,6	2,2	σ <sub>0</sub> (max. Mutmaßl.)	0,9309	4	100%	94%	97%
5	5	36	0	1,61	0,4	0,7	-2,0	2,5	10,0	-0,6	0,6	1,0	Anzahl N d. Einzelgr.	600	5	100%	96%	98%
6	6	36	0	1,79	0,4	0,6	-1,4	1,9	8,3	-0,4	0,4	0,8	Vertrauensgrad V	95%	6	100%	97%	99%
7	7	30	0	1,95	0,3	0,6	-0,9	1,3	5,9	-0,3	0,3	0,5	γ(N; V)	5,138	7	100%	98%	99%
8	8	30	0	2,08	0,3	0,5	-0,7	1,0	5,1	-0,2	0,2	0,4	k <sub>0</sub>	0,2904	8	100%	98%	99%
9	9	42	0	2,20	0,2	0,5	-0,8	1,2	6,2	-0,2	0,2	0,5	k <sub>1</sub>	-0,2371	9	100%	99%	99%
Σ		531	69		4,0	4,7	-174,2	78,4	164,8	-173,8	0,0	0,0	k <sub>2</sub>	0,2747				

Anz. untersch. Einzelgr.	9
Devianz D	347,55
χ <sup>2</sup> (1-V; N-2)	656
Modell gut lt. D-Test?	ja

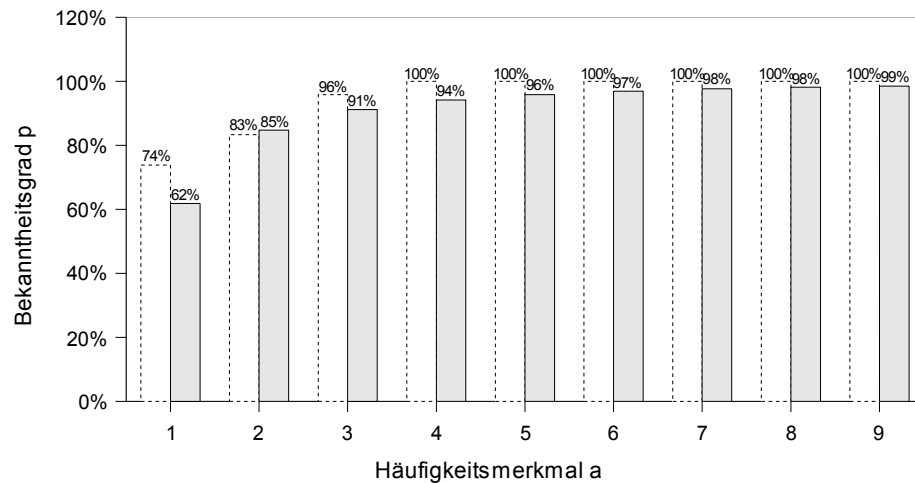


Abbildung 49: Auswertung von Fragebogen E

**Bestimmung des Bekanntheitsgrades (Fragebogen A bis E)**

Da- ten- Nr.	Häufig- keits- merk- mal a	Be- kannt	Nicht be- kannt	ln a	Hilfs- grö- ße 1	Hilfs- grö- ße 2	Für k <sub>0</sub>	Für k <sub>1</sub>	Für k <sub>2</sub>	ln L	∂lnL/∂ μ	∂lnL/∂ σ	ln L <sub>max</sub> (Summe)	-798,68	a	p <sub>1</sub>	p <sub>95%</sub>	p <sub>0</sub>
													∂ln L/∂μ (Summe)	0,0000				
1	1	749	151	0,00			-453,2	201,0	293,2	-408,4	18,8	0,0	Mittelwert μ <sub>1</sub> (Startwert)	0,1653	1	83%	76%	81%
2	2	491	73	0,69	1,2	0,4	-144,0	104,1	247,7	-223,6	-25,2	-17,4	Steilheit σ <sub>1</sub> (Startwert)	0,3978	2	87%	89%	92%
3	3	332	28	1,10	0,7	0,7	-57,7	51,1	149,1	-100,6	-9,5	-10,5	μ <sub>0</sub> (max. Mutmaßl.)	-1,0996	3	92%	93%	95%
4	4	299	7	1,39	0,5	0,7	-34,6	34,7	114,4	-34,2	3,9	5,4	σ <sub>0</sub> (max. Mutmaßl.)	1,3671	4	98%	95%	96%
5	5	179	1	1,61	0,4	0,7	-15,4	16,8	60,5	-8,5	3,8	6,1	Anzahl N d. Einzelgr.	3000	5	99%	96%	97%
6	6	178	2	1,79	0,4	0,6	-12,2	14,3	54,8	-11,5	1,8	3,2	Vertrauensgrad V	95%	6	99%	97%	98%
7	7	150	0	1,95	0,3	0,6	-8,4	10,3	41,6	-2,6	2,6	5,0	γ(N; V)	5,138	7	100%	98%	98%
8	8	150	0	2,08	0,3	0,5	-7,1	9,1	38,2	-2,2	2,2	4,5	k <sub>0</sub>	0,2470	8	100%	98%	99%
9	9	209	1	2,20	0,2	0,5	-8,5	11,3	49,4	-7,0	1,6	3,5	k <sub>1</sub>	-0,2737	9	100%	98%	99%
Σ		2737	263		4,1	4,6	-741,1	452,7	1048,8	-798,7	0,0	0,0	k <sub>2</sub>	0,3496				
													Anz. untersch. Einzelgr.	9				
													Devianz D	1597,36				
													χ <sup>2</sup> (1-V; N-2)	3126,49				
													Modell gut lt. D-Test?	ja				

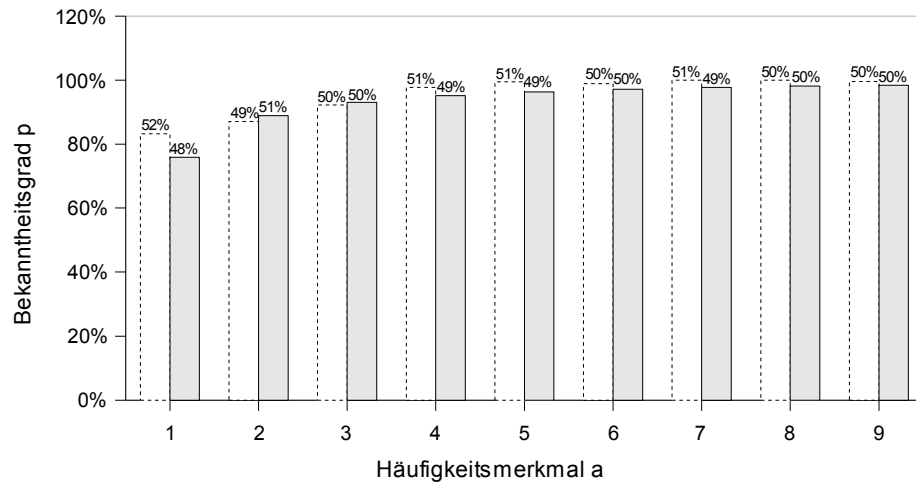


Abbildung 50: Auswertung der Fragebogen A bis E

**Bestimmung des Bekanntheitsgrades (Muttersprache aus nicht-indoeuropäischer Sprachfamilie)**

Da- ten- Nr.	Häufig- keits- merk- mal a	Be- kannt	Nicht be- kannt	ln a	Hilfs- grö- ße 1	Hilfs- grö- ße 2	Für k <sub>0</sub>	Für k <sub>1</sub>	Für k <sub>2</sub>	ln L	∂lnL/∂ μ	∂lnL/∂ σ	ln L <sub>max</sub> (Summe)		a	p <sub>1</sub>	p <sub>95%</sub>	p <sub>0</sub>	
													-389,54	0,0000					
													∂ln L/∂μ (Summe)						
													0,0000						
													∂ln L/∂σ (Summe)						
													0,0000						
1	1	390	90	0,00			-254,5	106,3	146,1	-232,0	6,9	0,0	Mittelwert μ <sub>1</sub> (Startwert)		0,1426	1	81%	73%	80%
2	2	271	32	0,69	1,2	0,4	-65,1	51,1	132,0	-104,7	-10,7	-7,4	Steilheit σ <sub>1</sub> (Startwert)		0,3263	2	89%	90%	93%
3	3	183	9	1,10	0,7	0,7	-21,9	21,9	72,0	-36,6	-2,1	-2,3	μ <sub>0</sub> (max. Mutmaßl.)		-0,7880	3	95%	94%	96%
4	4	160	1	1,39	0,6	0,7	-11,4	13,2	49,9	-7,4	2,6	3,5	σ <sub>0</sub> (max. Mutmaßl.)		1,0402	4	99%	96%	98%
5	5	96	0	1,61	0,4	0,7	-4,7	6,0	24,9	-1,5	1,4	2,3	Anzahl N d. Einzelgr.		1600	5	100%	98%	98%
6	6	95	1	1,79	0,4	0,6	-3,4	4,7	21,1	-5,6	0,1	0,1	Vertrauensgrad V		95%	6	99%	98%	99%
7	7	80	0	1,95	0,3	0,6	-2,2	3,2	15,2	-0,7	0,7	1,3	γ(N; V)		5,138	7	100%	99%	99%
8	8	80	0	2,08	0,3	0,5	-1,7	2,7	13,3	-0,5	0,5	1,1	k <sub>0</sub>		0,2294	8	100%	99%	99%
9	9	112	0	2,20	0,2	0,5	-2,0	3,2	16,5	-0,6	0,6	1,3	k <sub>1</sub>		-0,2405	9	100%	99%	99%
Σ		1467	133		4,1	4,7	-367,0	212,1	491,0	-389,5	0,0	0,0	k <sub>2</sub>		0,3069				

Anz. untersch. Einzelgr.	9
Devianz D	779,07
χ <sup>2</sup> (1-V; N-2)	1692,11
Modell gut lt. D-Test?	ja

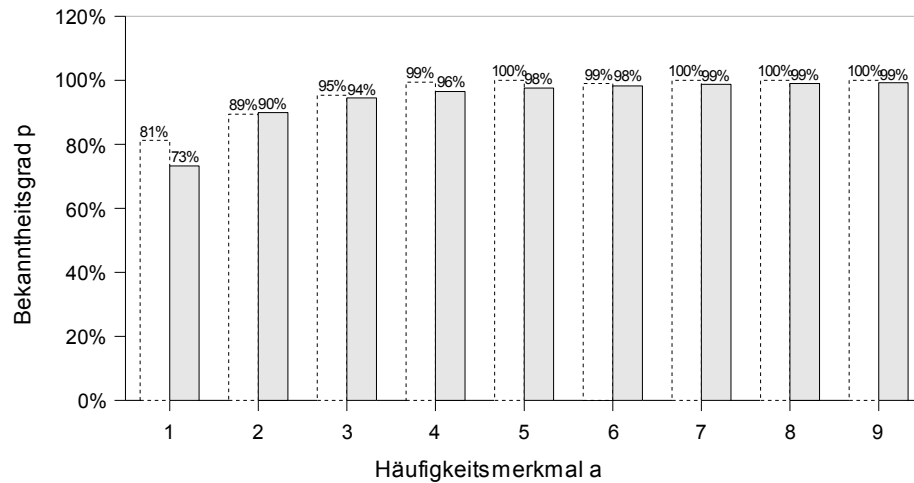


Abbildung 51: Auswertung, Muttersprache aus nicht-indoeuropäischer Sprachfamilie

**Bestimmung des Bekanntheitsgrades (Muttersprache aus indoeuropäischer Sprachfamilie)**

Da- ten- Nr.	Häufig- keits- merk- mal a	Be- kannt	Nicht be- kannt	ln a	Hilfs- grö- ße 1	Hilfs- grö- ße 2	Für k <sub>0</sub>	Für k <sub>1</sub>	Für k <sub>2</sub>	ln L	∂lnL/∂ μ	∂lnL/∂ σ	ln L <sub>max</sub> (Summe)					
													∂ln L/∂μ (Summe)	∂ln L/∂σ (Summe)	a	p <sub>1</sub>	p <sub>95%</sub>	p <sub>0</sub>
													-403,47					
													0,0000					
													0,0000					
1	1	359	61	0,00			-200,5	94,0	145,1	-175,5	13,0	0,0	Mittelwert μ <sub>1</sub> (Startwert)	0,1914	1	85%	76%	82%
2	2	220	41	0,69	1,2	0,4	-74,8	50,9	113,9	-118,2	-15,9	-11,0	Steilheit σ <sub>1</sub> (Startwert)	0,4648	2	84%	87%	90%
3	3	149	19	1,10	0,7	0,6	-34,2	27,5	72,8	-61,8	-7,9	-8,6	μ <sub>0</sub> (max. Mutmaßl.)	-1,5377	3	89%	90%	93%
4	4	139	6	1,39	0,5	0,7	-22,9	20,4	59,9	-25,1	1,3	1,8	σ <sub>0</sub> (max. Mutmaßl.)	1,8076	4	96%	92%	95%
5	5	83	1	1,61	0,4	0,7	-10,8	10,4	32,8	-6,7	2,4	3,9	Anzahl N d. Einzelgr.	1400	5	99%	94%	96%
6	6	83	1	1,79	0,4	0,6	-9,1	9,3	31,0	-6,3	1,9	3,4	Vertrauensgrad V	95%	6	99%	94%	97%
7	7	70	0	1,95	0,3	0,6	-6,6	7,0	24,4	-2,1	2,1	4,0	γ(N; V)	5,138	7	100%	95%	97%
8	8	70	0	2,08	0,3	0,5	-5,8	6,4	23,2	-1,8	1,8	3,8	k <sub>0</sub>	0,2656	8	100%	96%	97%
9	9	97	1	2,20	0,2	0,5	-7,3	8,3	31,0	-6,0	1,3	2,8	k <sub>1</sub>	-0,3033	9	99%	96%	98%
Σ		1270	130		4,0	4,6	-371,9	234,1	533,9	-403,5	0,0	0,0	k <sub>2</sub>	0,3814				

Anz. untersch. Einzelgr.	9
Devianz D	806,95
χ <sup>2</sup> (1-V; N-2)	1486,1
Modell gut lt. D-Test?	ja

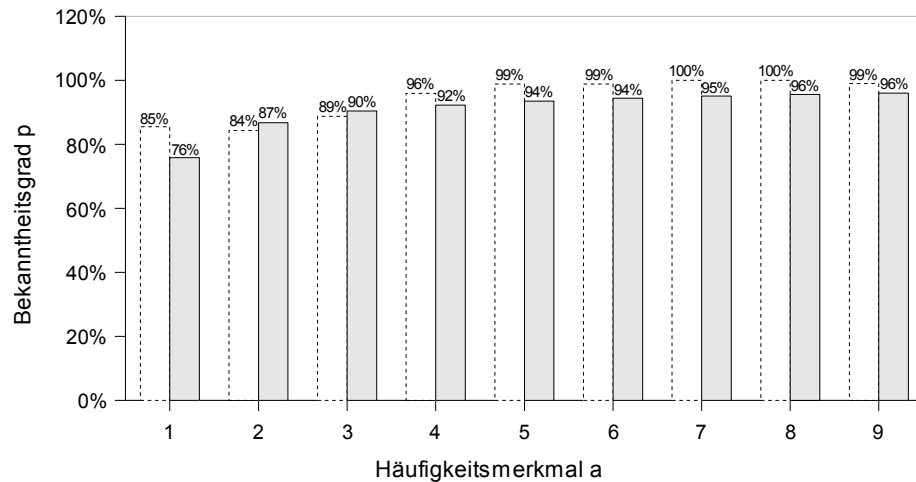


Abbildung 52: Auswertung, Muttersprache aus indoeuropäischer Sprachfamilie

**Bestimmung des Bekanntheitsgrades (Muttersprache Chinesisch oder Japanisch)**

Da- ten- Nr.	Häufig- keits- merk- mal a	Be- kannt	Nicht be- kannt	ln a	Hilfs- grö- ße 1	Hilfs- grö- ße 2	Für k <sub>0</sub>	Für k <sub>1</sub>	Für k <sub>2</sub>	ln L	∂lnL/∂ μ	∂lnL/∂ σ	ln L <sub>max</sub> (Summe)	-274,01	a	p <sub>1</sub>	p <sub>95%</sub>	p <sub>0</sub>
													∂ln L/∂μ (Summe)	0,0000				
1	1	232	68	0,00			-182,6	62,4	70,2	-160,8	5,5	0,0	Mittelwert μ <sub>1</sub> (Startwert)	0,1735	1	77%	68%	75%
2	2	164	26	0,69	1,1	0,4	-49,2	35,3	83,4	-78,5	-9,6	-6,7	Steilheit σ <sub>1</sub> (Startwert)	0,3506	2	86%	87%	91%
3	3	115	5	1,10	0,7	0,7	-16,5	15,5	47,7	-20,8	0,2	0,3	μ <sub>0</sub> (max. Mutmaßl.)	-0,6299	3	96%	93%	96%
4	4	99	1	1,39	0,6	0,7	-8,5	9,3	33,6	-6,3	1,7	2,3	σ <sub>0</sub> (max. Mutmaßl.)	1,0158	4	99%	95%	97%
5	5	60	0	1,61	0,4	0,7	-3,5	4,2	17,0	-1,1	1,1	1,7	Anzahl N d. Einzelgr.	1000	5	100%	97%	98%
6	6	59	1	1,79	0,4	0,6	-2,5	3,3	14,5	-5,1	-0,2	-0,4	Vertrauensgrad V	95%	6	98%	97%	99%
7	7	50	0	1,95	0,3	0,6	-1,6	2,3	10,4	-0,5	0,5	1,0	γ(N; V)	5,138	7	100%	98%	99%
8	8	50	0	2,08	0,3	0,5	-1,3	1,9	9,1	-0,4	0,4	0,8	k <sub>0</sub>	0,2672	8	100%	98%	99%
9	9	70	0	2,20	0,2	0,5	-1,5	2,2	11,3	-0,4	0,4	1,0	k <sub>1</sub>	-0,2476	9	100%	99%	99%
Σ		899	101		4,0	4,7	-267,2	136,5	297,2	-274,0	0,0	0,0	k <sub>2</sub>	0,2972				

Anz. untersch. Einzelgr.	9
Devianz D	548,01
χ <sup>2</sup> (1-V; N-2)	1072,61
Modell gut lt. D-Test?	ja

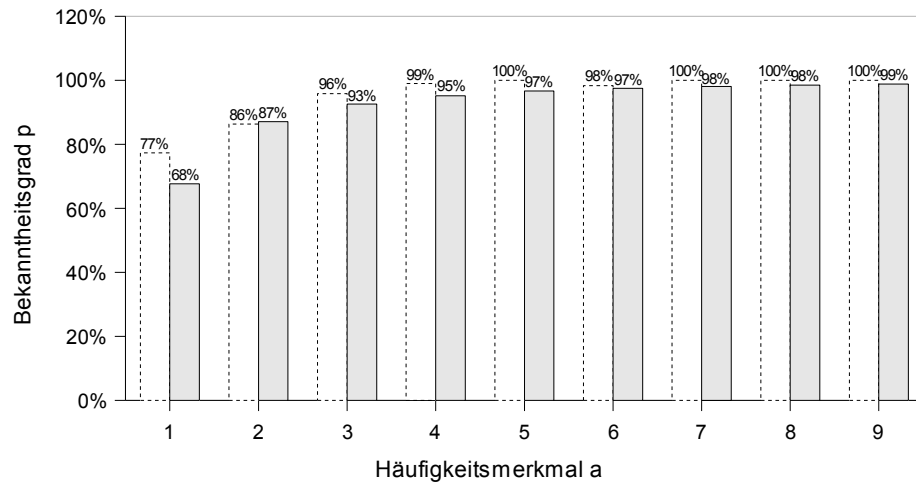


Abbildung 53: Auswertung, Muttersprache Chinesisch oder Japanisch



**Bestimmung des Bekanntheitsgrades (Muttersprache Japanisch)**

Daten-Nr.	Häufigkeitsmerkmal a	Bekannt	Nicht bekannt	ln a	Hilfsgröße 1	Hilfsgröße 2	Für k <sub>0</sub>	Für k <sub>1</sub>	Für k <sub>2</sub>	ln L	∂lnL/∂μ	∂lnL/∂σ	ln L <sub>max</sub> (Summe)					
													-212,78	a	p <sub>1</sub>	p <sub>95%</sub>	p <sub>0</sub>	
													∂ln L/∂μ (Summe)	0,0000				
													∂ln L/∂σ (Summe)	0,0000				
1	1	182	58	0,00			-151,2	48,5	51,2	-132,9	4,0	0,0	Mittelwert μ <sub>1</sub> (Startwert)	0,1664	1	76%	66%	74%
2	2	133	19	0,69	1,1	0,4	-36,5	27,2	66,7	-59,1	-6,9	-4,8	Steilheit σ <sub>1</sub> (Startwert)	0,3164	2	88%	87%	92%
3	3	93	3	1,10	0,7	0,7	-11,2	11,1	36,2	-13,4	0,5	0,6	μ <sub>0</sub> (max. Mutmaßl.)	-0,5240	3	97%	93%	96%
4	4	79	1	1,39	0,6	0,7	-5,4	6,3	24,2	-5,5	0,7	0,9	σ <sub>0</sub> (max. Mutmaßl.)	0,9003	4	99%	96%	98%
5	5	48	0	1,61	0,4	0,7	-2,1	2,7	11,7	-0,6	0,6	1,0	Anzahl N d. Einzelgr.	800	5	100%	97%	99%
6	6	48	0	1,79	0,4	0,6	-1,5	2,1	9,7	-0,4	0,4	0,8	Vertrauensgrad V	95%	6	100%	98%	99%
7	7	40	0	1,95	0,3	0,6	-0,9	1,4	6,7	-0,3	0,3	0,5	γ(N; V)	5,138	7	100%	98%	99%
8	8	40	0	2,08	0,3	0,5	-0,7	1,1	5,7	-0,2	0,2	0,4	k <sub>0</sub>	0,2627	8	100%	99%	99%
9	9	56	0	2,20	0,2	0,5	-0,8	1,3	6,9	-0,2	0,2	0,5	k <sub>1</sub>	-0,2304	9	100%	99%	100%
Σ		719	81		4,1	4,7	-210,1	101,6	219,2	-212,8	0,0	0,0	k <sub>2</sub>	0,2740				

Anz. untersch. Einzelgr.	9
Devianz D	425,56
χ <sup>2</sup> (1-V; N-2)	864,83
Modell gut lt. D-Test?	ja

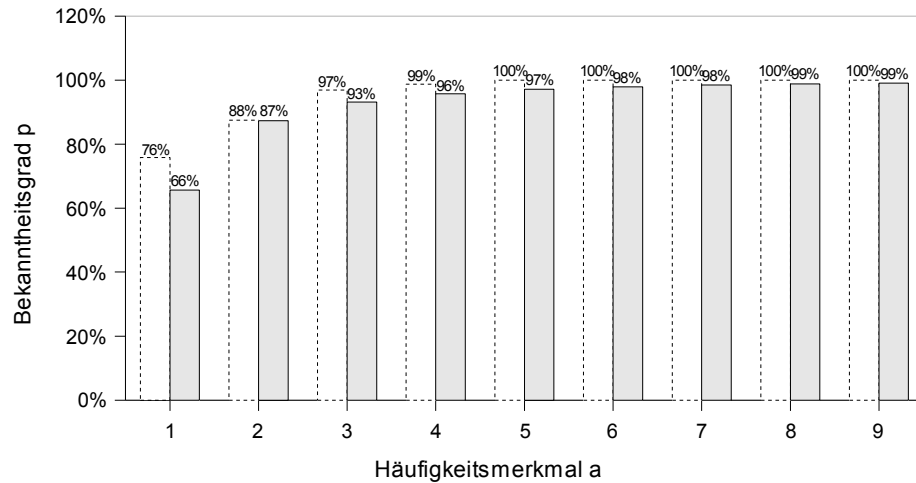


Abbildung 54: Auswertung, Muttersprache Japanisch

**Bestimmung des Bekanntheitsgrades (Muttersprache aus uralischer Sprachfamilie)**

Da- ten- Nr.	Häufig- keits- merk- mal a	Be- kannt	Nicht be- kannt	ln a	Hilfs- grö- ße 1	Hilfs- grö- ße 2	Für k <sub>0</sub>	Für k <sub>1</sub>	Für k <sub>2</sub>	ln L	∂lnL/∂ μ	∂lnL/∂ σ	ln L <sub>max</sub> (Summe)		a	p <sub>1</sub>	p <sub>95%</sub>	p <sub>0</sub>	
													-109,06	0,0000					
													∂ln L/∂μ (Summe)						
													0,0000						
													∂ln L/∂σ (Summe)						
													0,0000						
1	1	158	22	0,00			-66,8	38,7	73,6	-66,9	1,3	0,0	Mittelwert μ <sub>1</sub> (Startwert)		0,0908	1	88%	78%	87%
2	2	107	6	0,69	1,3	0,5	-15,8	14,8	45,2	-23,5	-1,0	-0,7	Steilheit σ <sub>1</sub> (Startwert)		0,2729	2	95%	92%	96%
3	3	68	4	1,10	0,8	0,7	-5,3	6,1	22,8	-16,7	-2,3	-2,6	μ <sub>0</sub> (max. Mutmaßl.)		-1,1374	3	94%	95%	98%
4	4	61	0	1,39	0,6	0,7	-2,8	3,7	15,5	-0,9	0,9	1,2	σ <sub>0</sub> (max. Mutmaßl.)		1,0833	4	100%	96%	99%
5	5	36	0	1,61	0,4	0,7	-1,2	1,6	7,5	-0,4	0,4	0,6	Anzahl N d. Einzelgr.		600	5	100%	97%	99%
6	6	36	0	1,79	0,4	0,6	-0,9	1,3	6,3	-0,3	0,3	0,5	Vertrauensgrad V		95%	6	100%	98%	99%
7	7	30	0	1,95	0,3	0,6	-0,6	0,9	4,5	-0,2	0,2	0,3	γ(N; V)		5,138	7	100%	98%	99%
8	8	30	0	2,08	0,3	0,5	-0,4	0,7	3,9	-0,1	0,1	0,3	k <sub>0</sub>		0,1573	8	100%	98%	100%
9	9	42	0	2,20	0,2	0,5	-0,5	0,9	4,9	-0,2	0,2	0,3	k <sub>1</sub>		-0,2073	9	100%	98%	100%
Σ		568	32		4,2	4,7	-94,4	68,6	184,3	-109,1	0,0	0,0	k <sub>2</sub>		0,3072				

Anz. untersch. Einzelgr.	9
Devianz D	218,13
χ <sup>2</sup> (1-V; N-2)	656
Modell gut lt. D-Test?	ja

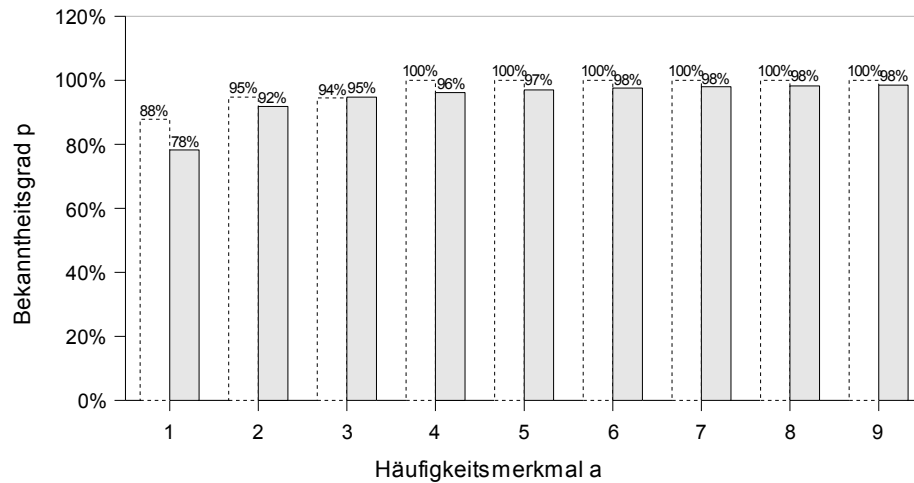


Abbildung 55: Auswertung, Muttersprache aus uralischer Sprachfamilie

**Bestimmung des Bekanntheitsgrades (Muttersprache aus baltischem oder slawischen Sprachzweig)**

Da- ten- Nr.	Häufig- keits- merk- mal a	Be- kannt	Nicht be- kannt	ln a	Hilfs- grö- ße 1	Hilfs- grö- ße 2	Für k <sub>0</sub>	Für k <sub>1</sub>	Für k <sub>2</sub>	ln L	∂lnL/∂ μ	∂lnL/∂ σ	ln L <sub>max</sub> (Summe)	-158,02	a	p <sub>1</sub>	p <sub>95%</sub>	p <sub>0</sub>
													∂ln L/∂μ (Summe)	0,0000				
1	1	126	24	0,00			-73,9	33,6	50,1	-66,2	3,5	0,0	Mittelwert μ <sub>1</sub> (Startwert)	0,2143	1	84%	72%	82%
2	2	81	13	0,69	1,2	0,4	-30,3	19,3	40,2	-38,1	-2,6	-1,8	Steilheit σ <sub>1</sub> (Startwert)	0,5229	2	86%	83%	89%
3	3	53	7	1,10	0,7	0,6	-14,6	10,8	26,3	-22,1	-2,2	-2,4	μ <sub>0</sub> (max. Mutmaßl.)	-1,7375	3	88%	87%	92%
4	4	45	6	1,39	0,5	0,6	-10,0	8,2	22,0	-19,5	-2,7	-3,8	σ <sub>0</sub> (max. Mutmaßl.)	2,1105	4	88%	88%	94%
5	5	29	1	1,61	0,4	0,6	-5,0	4,4	12,5	-4,5	0,6	1,0	Anzahl N d. Einzelgr.	500	5	97%	89%	95%
6	6	29	1	1,79	0,4	0,6	-4,3	4,0	12,1	-4,4	0,4	0,7	Vertrauensgrad V	95%	6	97%	90%	95%
7	7	25	0	1,95	0,3	0,6	-3,2	3,1	9,7	-1,0	1,0	2,0	γ(N; V)	5,141	7	100%	91%	96%
8	8	25	0	2,08	0,3	0,5	-2,9	2,9	9,4	-0,9	0,9	1,9	k <sub>0</sub>	0,2958	8	100%	91%	96%
9	9	35	0	2,20	0,2	0,5	-3,7	3,8	12,7	-1,2	1,2	2,5	k <sub>1</sub>	-0,3259	9	100%	92%	97%
Σ		448	52		4,0	4,5	-147,9	89,8	195,1	-158,0	0,0	0,0	k <sub>2</sub>	0,3902				

Anz. untersch. Einzelgr.	9
Devianz D	316,04
χ <sup>2</sup> (1-V; N-2)	551,02
Modell gut lt. D-Test?	ja

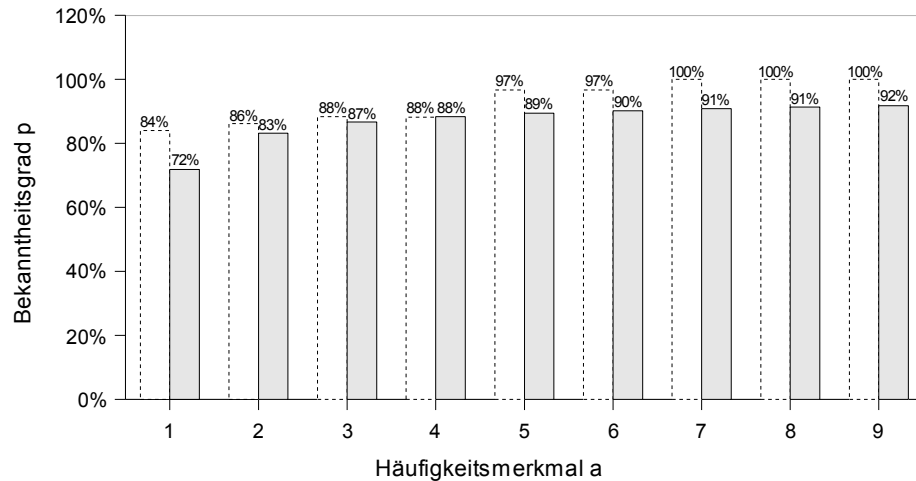


Abbildung 56: Auswertung, Muttersprache aus baltischem oder slawischem Sprachzweig

**Bestimmung des Bekanntheitsgrades (Muttersprache aus germanischem oder romanischem Sprachzweig)**

Daten-Nr.	Häufigkeitsmerkmal a	Bekannt	Nicht bekannt	ln a	Hilfsgröße 1	Hilfsgröße 2	Für k <sub>0</sub>	Für k <sub>1</sub>	Für k <sub>2</sub>	ln L	∂lnL/∂μ	∂lnL/∂σ	ln L <sub>max</sub> (Summe)		a	p <sub>1</sub>	p <sub>95%</sub>	p <sub>0</sub>	
													-149,21	0,0000					
													∂ln L/∂μ (Summe)						
													0,0000						
													∂ln L/∂σ (Summe)						
													0,0000						
1	1	185	25	0,00			-76,1	44,8	86,7	-76,7	1,5	0,0	Mittelwert μ <sub>1</sub> (Startwert)		0,1133	1	88%	79%	87%
2	2	123	7	0,69	1,3	0,5	-24,5	20,4	55,7	-27,3	0,9	0,6	Steilheit σ <sub>1</sub> (Startwert)		0,3523	2	95%	90%	94%
3	3	74	10	1,10	0,7	0,7	-10,4	10,1	32,4	-35,4	-6,7	-7,4	μ <sub>0</sub> (max. Mutmaßl.)		-1,6815	3	88%	93%	96%
4	4	73	0	1,39	0,5	0,7	-6,6	7,1	25,1	-2,1	2,1	2,9	σ <sub>0</sub> (max. Mutmaßl.)		1,5755	4	100%	94%	97%
5	5	42	0	1,61	0,4	0,7	-3,0	3,4	13,0	-0,9	0,9	1,5	Anzahl N d. Einzelgr.		700	5	100%	95%	98%
6	6	42	0	1,79	0,4	0,6	-2,4	3,0	11,9	-0,8	0,8	1,4	Vertrauensgrad V		95%	6	100%	96%	98%
7	7	35	0	1,95	0,3	0,6	-1,7	2,2	9,1	-0,5	0,5	1,0	γ(N; V)		5,138	7	100%	96%	98%
8	8	35	0	2,08	0,3	0,5	-1,5	1,9	8,4	-0,5	0,5	0,9	k <sub>0</sub>		0,1830	8	100%	97%	99%
9	9	48	1	2,20	0,2	0,5	-1,8	2,5	11,0	-0,4	-1,0	-1,0	k <sub>1</sub>		-0,2471	9	98%	97%	99%
Σ		657	43		4,2	4,7	-128,1	95,4	253,3	-149,2	0,0	0,0	k <sub>2</sub>		0,3619				

Anz. untersch. Einzelgr.	9
Devianz D	298,42
χ <sup>2</sup> (1-V; N-2)	760,57
Modell gut lt. D-Test?	ja

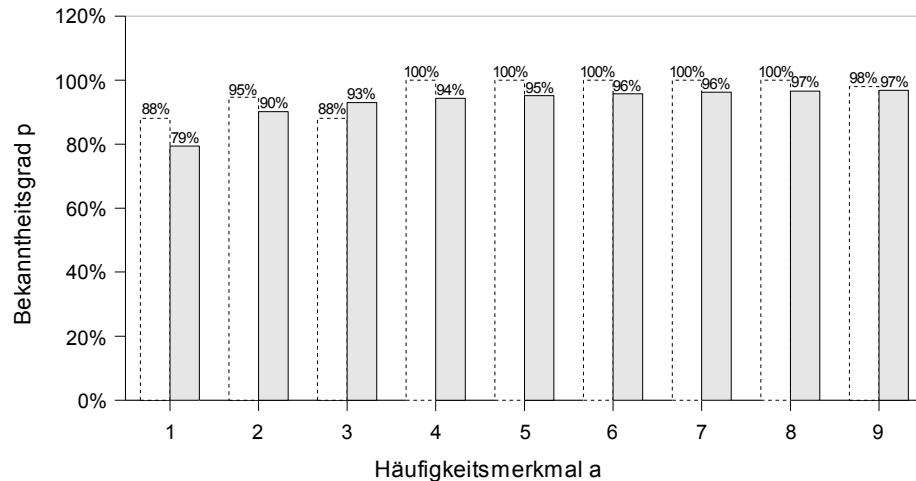


Abbildung 57: Auswertung, Muttersprache aus germanischem oder romanischem Sprachzweig

**Bestimmung des Bekanntheitsgrades (Muttersprache aus romanischem Sprachzweig)**

Da- ten- Nr.	Häufig- keits- merk- mal a	Be- kannt	Nicht be- kannt	ln a	Hilfs- grö- ße 1	Hilfs- grö- ße 2	Für k <sub>0</sub>	Für k <sub>1</sub>	Für k <sub>2</sub>	ln L	$\partial \ln L / \partial \mu$	$\partial \ln L / \partial \sigma$	ln L <sub>max</sub> (Summe)		a	p <sub>1</sub>	p <sub>95%</sub>	p <sub>0</sub>
													-56,51	0,0000				
													$\partial \ln L / \partial \mu$ (Summe)					
													0,0000					
													$\partial \ln L / \partial \sigma$ (Summe)					
													0,0000					
1	1	115	5	0,00			-18,0	16,3	48,9	-20,8	0,7	0,0	Mittelwert $\mu_1$ (Startwert)	0,0675	1	96%	88%	95%
2	2	72	2	0,69	1,3	0,5	-8,2	8,2	27,4	-9,3	0,6	0,4	Steilheit $\sigma_1$ (Startwert)	0,3234	2	97%	93%	97%
3	3	43	5	1,10	0,8	0,7	-4,4	4,7	16,6	-19,0	-3,6	-4,0	$\mu_0$ (max. Mutmaßl.)	-6,2159	3	90%	94%	97%
4	4	42	0	1,39	0,5	0,7	-3,4	3,8	13,8	-1,1	1,1	1,5	$\sigma_0$ (max. Mutmaßl.)	3,7678	4	100%	94%	97%
5	5	24	0	1,61	0,4	0,7	-1,7	2,0	7,5	-0,5	0,5	0,9	Anzahl N d. Einzelgr.	400	5	100%	94%	98%
6	6	24	0	1,79	0,4	0,6	-1,6	1,9	7,2	-0,5	0,5	0,9	Vertrauensgrad V	95%	6	100%	93%	98%
7	7	20	0	1,95	0,3	0,6	-1,2	1,5	5,8	-0,4	0,4	0,8	$\gamma(N; V)$	5,142	7	100%	93%	98%
8	8	20	0	2,08	0,3	0,5	-1,2	1,4	5,7	-0,4	0,4	0,8	k <sub>0</sub>	0,1030	8	100%	93%	98%
9	9	27	1	2,20	0,2	0,5	-1,6	1,9	7,7	-4,5	-0,5	-1,1	k <sub>1</sub>	-0,1893	9	96%	92%	98%
$\Sigma$		387	13		4,3	4,7	-41,2	41,7	140,7	-56,5	0,0	0,0	k <sub>2</sub>	0,3516				
													Anz. untersch. Einzelgr.	9				
													Devianz D	113,02				
													$\chi^2(1-V; N-2)$	445,52				
													Modell gut lt. D-Test?	ja				

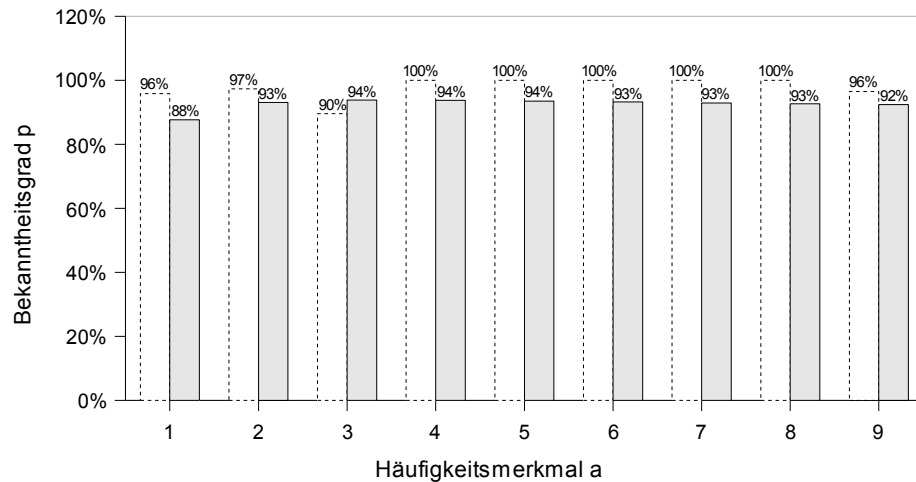


Abbildung 58: Auswertung, Muttersprache aus romanischem Sprachzweig

Da- ten- Nr.	Espe- ranto- Lern- zeit j	Be- kannt	Nicht be- kannt	ln j	Hilfs- grö- ße 1	Hilfs- grö- ße 2	Für k <sub>0</sub>	Für k <sub>1</sub>	Für k <sub>2</sub>	ln L	$\partial \ln L / \partial \mu$	$\partial \ln L / \partial \sigma$	ln L <sub>max</sub> (Summe)	-21,50	j	s <sub>95%</sub>	s <sub>0</sub>
													$\partial \ln L / \partial \mu$ (Summe)	0,0000			
1	1	0	1	0,00			-0,3	-0,2	0,4	-0,1	-0,1	0,0	Mittelwert $\mu_1$ (Startwert)	1,8749	1	0%	9%
2	2	0	1	0,69	0,0	0,0	-0,5	-0,2	0,3	-0,2	-0,2	-0,1	Steilheit $\sigma_1$ (Startwert)	1,0374	2	2%	19%
3	4	2	1	1,39	0,5	0,6	-2,3	-0,4	0,2	-2,5	0,9	1,3	$\mu_0$ (max. Mutmaßl.)	1,8626	3	6%	28%
4	7	1	2	1,95	0,6	0,9	-2,5	0,1	0,0	-2,1	-0,6	-1,1	$\sigma_0$ (max. Mutmaßl.)	1,4431	4	12%	35%
5	8	1	0	2,08	0,2	0,4	-0,8	0,1	0,0	-0,6	0,4	0,9	Anzahl N d. Einzelgr.	45	5	17%	42%
6	9	1	2	2,20	0,2	0,3	-2,4	0,3	0,1	-2,4	-0,8	-1,8	Vertrauensgrad V	95%	6	23%	48%
7	13	3	0	2,56	0,5	1,2	-2,0	0,5	0,5	-1,0	0,9	2,3	$\gamma(N; V)$	5,185	7	29%	53%
8	14	1	0	2,64	0,1	0,4	-0,7	0,2	0,2	-0,3	0,3	0,7	k <sub>0</sub>	0,5065	8	34%	57%
9	17	1	2	2,83	0,3	0,7	-1,7	0,6	0,8	-3,2	-1,3	-3,7	k <sub>1</sub>	-0,2573	9	38%	60%
10	18	0	1	2,89	0,0	0,1	-0,6	0,2	0,3	-1,5	-0,8	-2,3	k <sub>2</sub>	0,2847	10	42%	63%
11	19	1	0	2,94	0,1	0,2	-0,5	0,2	0,3	-0,2	0,2	0,6	Anz. untersch. Einzelgr.	31	11	45%	66%
12	20	1	0	3,00	0,1	0,3	-0,5	0,2	0,3	-0,2	0,2	0,6	Devianz D	43	12	48%	69%
13	21	1	0	3,04	0,1	0,3	-0,5	0,2	0,3	-0,2	0,2	0,6	$\chi^2(1-V; N-2)$	59,3	13	50%	71%
14	22	2	0	3,09	0,1	0,3	-1,0	0,4	0,7	-0,4	0,4	1,1	Modell gut lt. D-Test?	ja	14	52%	73%
15	23	1	0	3,14	0,1	0,3	-0,5	0,2	0,4	-0,2	0,2	0,5			15	54%	74%
16	24	2	0	3,18	0,1	0,3	-0,9	0,4	0,7	-0,4	0,3	1,0			16	56%	76%
17	25	1	0	3,22	0,1	0,3	-0,4	0,2	0,4	-0,2	0,2	0,5			17	57%	77%
18	27	2	0	3,30	0,2	0,5	-0,8	0,4	0,8	-0,3	0,3	0,9			18	58%	78%
19	28	1	0	3,33	0,1	0,2	-0,4	0,2	0,4	-0,1	0,1	0,5			19	59%	80%
20	29	1	1	3,37	0,1	0,2	-0,7	0,4	0,8	-2,2	-0,7	-2,5			20	60%	81%
21	31	1	0	3,43	0,1	0,3	-0,4	0,2	0,4	-0,1	0,1	0,4			21	61%	82%
22	35	0	1	3,56	0,1	0,4	-0,3	0,2	0,4	-2,2	-0,9	-3,2			22	61%	82%
23	37	1	0	3,61	0,1	0,2	-0,3	0,2	0,4	-0,1	0,1	0,4			23	62%	83%
24	39	1	0	3,66	0,1	0,4	-0,3	0,2	0,4	-0,1	0,1	0,3			24	63%	84%

Abbildung 59: Selbsteinschätzung des eigenen Esperanto-Anwendungsvermögens s und Esperanto-Lernzeit j  
(Teil 1 von 2)

25	40	1	0	3,69	0,1	0,2	-0,3	0,2	0,4	-0,1	0,1	0,3
26	41	1	0	3,71	0,0	0,2	-0,3	0,2	0,4	-0,1	0,1	0,3
27	43	1	0	3,76	0,1	0,4	-0,3	0,2	0,4	-0,1	0,1	0,3
28	44	1	0	3,78	0,0	0,2	-0,2	0,2	0,4	-0,1	0,1	0,3
29	47	1	0	3,85	0,1	0,5	-0,2	0,2	0,4	-0,1	0,1	0,3
30	50	1	0	3,91	0,1	0,5	-0,2	0,2	0,4	-0,1	0,1	0,3
31	51	1	0	3,93	0,0	0,2	-0,2	0,2	0,4	-0,1	0,1	0,3
Σ		33	12		4,1	10,7	-22,8	6,4	12,8	-21,5	0,0	0,0

25	63%	85%
26	64%	85%
27	64%	86%
28	64%	86%
29	65%	87%
30	65%	87%
31	65%	88%
32	66%	88%
33	66%	89%
34	66%	89%
35	66%	89%
36	67%	90%
37	67%	90%
38	67%	90%
39	67%	91%
40	68%	91%
41	68%	91%
42	68%	91%
43	68%	92%
44	68%	92%
45	68%	92%
46	69%	92%
47	69%	92%
48	69%	93%
49	69%	93%
50	69%	93%
100	73%	97%

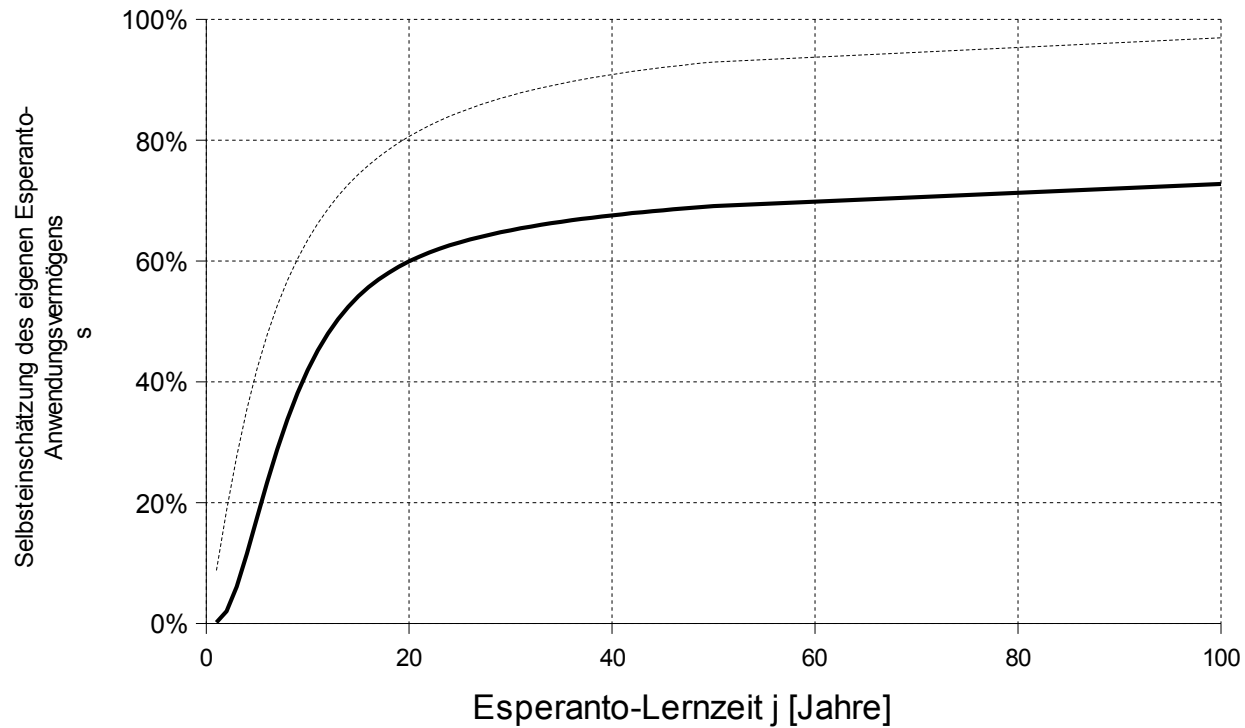


Abbildung 60: Selbsteinschätzung des eigenen Esperanto-Anwendungsvermögens  $s$  und Esperanto-Lernzeit  $j$  (Teil 2 von 2)

a	p <sub>0</sub>	f	Matrix R		Matrix J (Differenzenquot.)		Differenz	(p <sub>0</sub> - <p <sub>0</sub> >) <sup>2</sup>
			r=y-f	r <sup>2</sup>	dr/dA	dr/dB		
1	0,81	0,91	-0,10	0,01	-1,05	-1,29	0,000010	0,02
2	0,92	0,92	0,00	0,00	-1,06	-1,46		0,00
3	0,95	0,92	0,02	0,00	-1,06	-1,65		0,00
4	0,96	0,93	0,03	0,00	-1,07	-1,88		0,00
5	0,97	0,94	0,03	0,00	-1,08	-2,16		0,00
6	0,98	0,95	0,03	0,00	-1,09	-2,50		0,00
7	0,98	0,97	0,02	0,00	-1,11	-2,95		0,00
8	0,99	0,99	0,00	0,00	-1,13	-3,62		0,00
9	0,99	1,02	-0,03	0,00	-1,18	-4,81		0,00
	0,95			0,0152				0,03
							Bestmaß	0,41

P.	Start	Verbessert	K = (J transp. mal J) <sup>-1</sup>		D = J transp. mal R	E = K mal D	Verbessert minus E
A	1,3000	0,8686	0,69	-0,26	0,0000000	0,0000001	0,8686
B	-0,0500	0,0344	-0,26	0,11	-0,0000002	0,0000000	0,0344

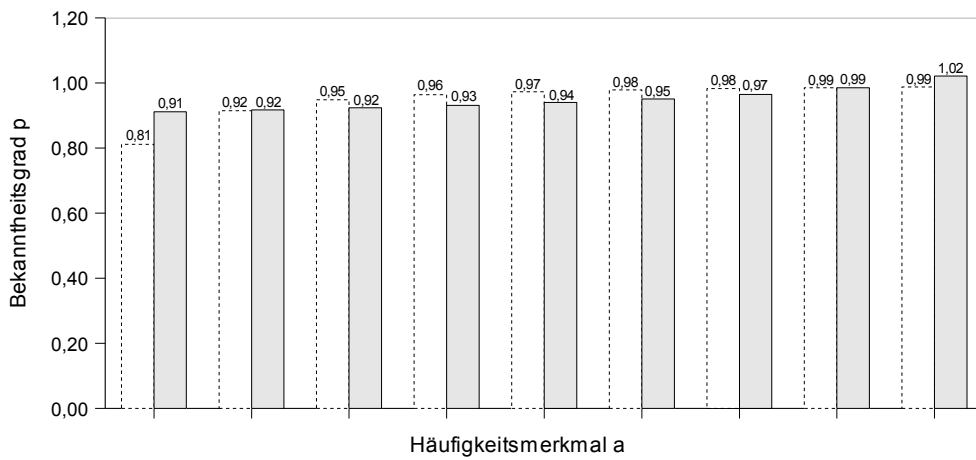


Abbildung 61: Modellvergleich



## Anhang G: Umsetzung des Auswerte-Algorithmus

Die Auswertealgorithmen der in der vorliegenden Arbeit verwendeten logistischen Regression wurden mit dem Tabellenkalkulationsprogramm OpenOffice.org Calc umgesetzt. Es kann unentgeltlich aus dem Internet bezogen werden<sup>66</sup>. In den folgenden Abbildungen 62 bis 86 wird diese Umsetzung dargestellt, sodass der Leser sie nachvollziehen kann (ausreichende Vertrautheit mit dem genannten Tabellenkalkulationsprogramm wird vorausgesetzt).

	A	B	C	D	E	F
1	<b>Bestimmung des Bekanntheitsgrades</b>					
2	Da- ten- Nr.	Häufig- keits- merk- mal a	Be- kannt	Nicht be- kannt	ln a	Hilfs- grö- ße 1
3						
4						
5	1	1	749	151	0,00	

Abbildung 62: Zelle \$E5

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	<b>Bestimmung des Bekanntheitsgrades (Fragebögen A bis E)</b>															
2	Da- ten- Nr.	Häufig- keits- merk- mal a	Be- kannt	Nicht be- kannt	ln a	Hilfs- grö- ße 1	Hilfs- grö- ße 2	Für k <sub>0</sub>	Für k <sub>1</sub>	Für k <sub>2</sub>	ln L	∂lnL/∂ μ	∂lnL/∂ σ	ln L <sub>max</sub> (Summe)	-798,68	
3														∂ln L/∂μ (Summe)	0,0000	
4														∂ln L/∂σ (Summe)	0,0000	a
5	1	1	749	151	0,00			-453,2	201,0	293,2	-408,4	18,8	0,0	Mittelwert μ <sub>1</sub> (Startwert)	0,1653	1

Abbildung 63: Zelle \$H5

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	<b>Bestimmung des Bekanntheitsgrades (Fragebögen A bis E)</b>																		
2	Da- ten- Nr.	Häufig- keits- merk- mal a	Be- kannt	Nicht be- kannt	ln a	Hilfs- grö- ße 1	Hilfs- grö- ße 2	Für k <sub>0</sub>	Für k <sub>1</sub>	Für k <sub>2</sub>	ln L	∂lnL/∂ μ	∂lnL/∂ σ	ln L <sub>max</sub> (Summe)	-798,68				
3														∂ln L/∂μ (Summe)	0,0000				
4														∂ln L/∂σ (Summe)	0,0000	a	p <sub>1</sub>	p <sub>99%</sub>	p <sub>0</sub>
5	1	1	749	151	0,00			-453,2	201,0	293,2	-408,4	18,8	0,0	Mittelwert μ <sub>1</sub> (Startwert)	0,1653	1	83%	76%	81%

Abbildung 64: Zelle \$I5

66 <http://www.openoffice.org>

$$=(PI()/WURZEL(3)*(\$E5-\$O\$7)/\$O\$8)^2*EXP(PI()/WURZEL(3)*(\$E5-\$O\$7)/\$O\$8)/(1+EXP(PI()/WURZEL(3)*(\$E5-\$O\$7)/\$O\$8))^2*(\$C5+\$D5)$$

F	G	H	I	J	K	L	M	N	O	P	Q	R	S
<b>ntheitsgrades (Fragebögen A bis E)</b>													
								ln L <sub>max</sub> (Summe)	-798,68				
Hilfsgrö- ße 1	Hilfsgrö- ße 2	Für k <sub>0</sub>	Für k <sub>1</sub>	Für k <sub>2</sub>	ln L	∂lnL/∂ μ	∂lnL/∂ σ	∂ln L/∂μ (Summe)	0,0000				
		-453,2	201,0	293,2	-408,4	18,8	0,0	∂ln L/∂σ (Summe)	0,0000	a	p <sub>1</sub>	p <sub>95%</sub>	p <sub>0</sub>
								Mittelwert μ <sub>1</sub> (Startwert)	0,1653	1	83%	76%	81%

Abbildung 65: Zelle \$J5

K	L	M	N	O	P	Q	R	S	T
<b>n A bis E)</b>									
			ln L <sub>max</sub> (Summe)	-798,68					
		∂lnL/∂ μ	∂lnL/∂ σ	∂ln L/∂μ (Summe)	0,0000				
k <sub>2</sub>	ln L			∂ln L/∂σ (Summe)	0,0000	a	p <sub>1</sub>	p <sub>95%</sub>	p <sub>0</sub>
3,2				=WENN(\$D5>0; \$C5*LN(EXP(PI()/WURZEL(3)*(\$E5-\$O\$7)/\$O\$8)/(EXP(PI()/WURZEL(3)*(\$E5-\$O\$7)/\$O\$8)+1))+ \$D5*LN(1-EXP(PI()/WURZEL(3)*(\$E5-\$O\$7)/\$O\$8)/(EXP(PI()/WURZEL(3)*(\$E5-\$O\$7)/\$O\$8)+1)); \$C5*LN(EXP(PI()/WURZEL(3)*(\$E5-\$O\$7)/\$O\$8)/(EXP(PI()/WURZEL(3)*(\$E5-\$O\$7)/\$O\$8)+1)))					
7,7									
3,1									

Abbildung 66: Zelle \$K5

$$= \$C5-EXP(PI()/WURZEL(3)*(\$E5-\$O\$7)/\$O\$8)/(EXP(PI()/WURZEL(3)*(\$E5-\$O\$7)/\$O\$8)+1)*(\$C5+\$D5)$$

L	M	N	O	P	Q	R	S	T	U
<b>is E)</b>									
			ln L <sub>max</sub> (Summe)	-798,68					
	∂lnL/∂ μ	∂lnL/∂ σ	∂ln L/∂μ (Summe)	0,0000					
			∂ln L/∂σ (Summe)	0,0000	a	p <sub>1</sub>	p <sub>95%</sub>	p <sub>0</sub>	
8,4	18,8	0,0	Mittelwert μ <sub>1</sub> (Startwert)	0,1653	1	83%	76%	81%	
			Steilheit σ (Startwert)	0,0000	0	87%	88%	88%	

Abbildung 67: Zelle \$L5

$$= (\$C5-EXP(PI()/WURZEL(3)*(\$E5-\$O\$7)/\$O\$8)/(EXP(PI()/WURZEL(3)*(\$E5-\$O\$7)/\$O\$8)+1)*(\$C5+\$D5))*\$E5$$

M	N	O	P	Q	R	S	T	U
			ln L <sub>max</sub> (Summe)	-798,68				
	∂lnL/∂ μ	∂lnL/∂ σ	∂ln L/∂μ (Summe)	0,0000				
			∂ln L/∂σ (Summe)	0,0000	a	p <sub>1</sub>	p <sub>95%</sub>	p <sub>0</sub>
3,8	0,0		Mittelwert μ <sub>1</sub> (Startwert)	0,1653	1	83%	76%	81%

Abbildung 68: Zelle \$M5

=\$K\$14		
M	N	O
	$\ln L_{\max}$ (Summe)	-798,68
	$\partial \ln L / \partial \mu$ (Summe)	0,0000

Abbildung 69: Zelle \$O\$2

=\$L\$14		
M	N	O
	$\ln L_{\max}$ (Summe)	-798,68
$\partial \ln L / \partial \sigma$	$\partial \ln L / \partial \mu$ (Summe)	0,0000

Abbildung 70: Zelle \$O\$3

=\$M\$14		
M	N	O
	$\ln L_{\max}$ (Summe)	-798,68
$\partial \ln L / \partial \sigma$	$\partial \ln L / \partial \mu$ (Summe)	0,0000
$\sigma$	$\partial \ln L / \partial \sigma$ (Summe)	0,0000

Abbildung 71: Zelle \$O\$4

=\$E\$13-0,5*\$E\$5*\$C\$5/(\$C\$5+\$D\$5)-0,5*\$F\$14		
M	N	O
	$\ln L_{\max}$ (Summe)	-798,68
$\partial \ln L / \partial \sigma$	$\partial \ln L / \partial \mu$ (Summe)	0,0000
$\sigma$	$\partial \ln L / \partial \sigma$ (Summe)	0,0000
.8	0,0	Mittelwert $\mu_1$ (Startwert)
		0,1653
	Steilheit $\sigma$ (Startwert)	0,0000

Abbildung 72: Zelle \$O\$5

=WURZEL(ABS(\$E\$13^2-\$E\$5^2*\$C\$5/(\$C\$5+\$D\$5)-\$G\$14-\$O\$5^2))						
	M	N	O	P	Q	R
		ln L <sub>max</sub> (Summe)	-798,68			
$\frac{\partial \ln L}{\partial \mu}$		$\frac{\partial \ln L}{\partial \mu}$ (Summe)	0,0000			
$\frac{\partial \ln L}{\partial \sigma}$		$\frac{\partial \ln L}{\partial \sigma}$ (Summe)	0,0000	a	p <sub>1</sub>	p <sub>95%</sub>
1,8	0,0	Mittelwert $\mu_1$ (Startwert)	0,1653	1	83%	76%
1,2	-17,4	Steilheit $\sigma_1$ (Startwert)	0,3978	2	87%	89%

Abbildung 73: Zelle \$O\$6

=\$C\$14+\$D\$14			
	M	N	O
		ln L <sub>max</sub> (Summe)	-798,68
$\frac{\partial \ln L}{\partial \mu}$		$\frac{\partial \ln L}{\partial \mu}$ (Summe)	0,0000
$\frac{\partial \ln L}{\partial \sigma}$		$\frac{\partial \ln L}{\partial \sigma}$ (Summe)	0,0000
3,8	0,0	Mittelwert $\mu_1$ (Startwert)	0,1653
5,2	-17,4	Steilheit $\sigma_1$ (Startwert)	0,3978
3,5	-10,5	$\mu_0$ (max. Mutmaßl.)	-1,0996
3,9	5,4	$\sigma_0$ (max. Mutmaßl.)	1,3671
3,8	6,1	Anzahl N d. Einzelgr.	3000

Abbildung 74: Zelle \$O\$9

$\frac{\partial \ln L}{\partial k_0}$	0,100	100%
$\frac{\partial \ln L}{\partial k_0}$	=-\$H14/\$O\$9	
$\frac{\partial \ln L}{\partial k_0}$	0,2737	100%

Abbildung 75: Zelle \$O\$12

$k_1$	$=PI()*\$114/\$0\$9/WURZEL(3)$
-------	--------------------------------

Abbildung 76: Zelle \$O\$13

$k_2$	$=\$J14/\$0\$9$
-------	-----------------

Abbildung 77: Zelle \$O\$14

Devianz D	$=(-2)*\$0\$2$
-----------	----------------

Abbildung 78: Zelle \$O\$16

$\chi^2(1-V; N-2)$	$=CHIINV(1-\$0\$10;\$0\$9-2)$
--------------------	-------------------------------

Abbildung 79: Zelle \$O\$17

Modell gut lt. D-Test?	$=WENN(\$0\$16<\$0\$17;"ja";"nein")$
------------------------	--------------------------------------

Abbildung 80: Zelle \$O\$18

$=\$C5/(\$C5+\$D5)$					
	M	N	O	P	Q
$\partial \ln L / \partial \mu$		$\partial \ln L / \partial \mu$ (Summe)	0,0000		
$\partial \ln L / \partial \sigma$		$\partial \ln L / \partial \sigma$ (Summe)	0,0000	a	$P_1$
0,8	0,0	Mittelwert $\mu_1$ (Startwert)	0,1653	1	83%

Abbildung 81: Zelle \$Q3

R	S	T	U	V	W	X
$P_{95\%}$	$P_0$					
$=NORMVERT((LN(\$B5)-\$0\$7)/\$0\$8-WURZEL(\$0\$11/\$0\$9/\$0\$12*(1+(\$0\$12*(LN(\$B5)-\$0\$7)/\$0\$8+\$0\$13)^2/(\$0\$12*\$0\$14-\$0\$13^2))),0;1;1)$						

Abbildung 82: Zelle \$R3

= =EXP(PI()/WURZEL(3)\*(LN(\$B5)-\$O\$7)/\$O\$8)/(EXP(PI()/WURZEL(3)\*(LN(\$B5)-\$O\$7)/\$O\$8)+1)

S	T	U	V	W	X
p <sub>0</sub>					
81%					

Abbildung 83: Zelle \$S3

=(\$E6-\$E5)\*(\$C6/(\$C6+\$D6)+\$C5/(\$C5+\$D5))

	F	G	H	I	J	K
<b>intheitsgrades (Fragebögen A bis</b>						
	Hilfs- grö- ße 1	Hilfs- grö- ße 2	Für k <sub>0</sub>	Für k <sub>1</sub>	Für k <sub>2</sub>	ln L
	,00		-453,2	201,0	293,2	-408,4
	69	1,2	0,4	-144,0	104,1	247,7

Abbildung 84: Zelle \$F6

=(\$E6-\$E5)\*(\$E6\*\$C6/(\$C6+\$D6)+\$E5\*\$C5/(\$C5+\$D5))

	F	G	H	I	J	K	L
<b>intheitsgrades (Fragebögen A bis E)</b>							
	Hilfs- grö- ße 1	Hilfs- grö- ße 2	Für k <sub>0</sub>	Für k <sub>1</sub>	Für k <sub>2</sub>	ln L	∂lnL/∂ μ
	,00		-453,2	201,0	293,2	-408,4	18,8
	69	1,2	0,4	-144,0	104,1	247,7	-223,6

Abbildung 85: Zelle \$G6

f(x) Σ = =SUMME(C5:C13)

	C	D	E	F	G	H	I	J	K	L	M
1	299	7	1,39	0,5	0,7	-34,6	34,7	114,4	-34,2	3,9	5,4
5	179	1	1,61	0,4	0,7	-15,4	16,8	60,5	-8,5	3,8	6,1
3	178	2	1,79	0,4	0,6	-12,2	14,3	54,8	-11,5	1,8	3,2
7	150	0	1,95	0,3	0,6	-8,4	10,3	41,6	-2,6	2,6	5,0
3	150	0	2,08	0,3	0,5	-7,1	9,1	38,2	-2,2	2,2	4,5
3	209	1	2,20	0,2	0,5	-8,5	11,3	49,4	-7,0	1,6	3,5
	2737	263		4,1	4,6	-741,1	452,7	1048,8	-798,7	0,0	0,0

Abbildung 86: Zelle \$C\$14

In die Zellen  $\$D\$14$  und  $\$F\$14$  bis  $\$M\$14$  sind die entsprechenden Summenformeln (also  $=\text{SUMME}(D5:D13)$  usw.) einzutragen (ohne Abbildung).

Die Formeln in den Zellen unterhalb von  $\$E5$ ,  $\$F6$ ,  $\$G6$ ,  $\$H5$  bis  $\$M5$  und  $\$Q5$  bis  $\$S5$  erhält man am einfachsten durch Kopieren der genannten Zellen und anschließendes Einfügen. Beispiel: In  $\$E6$  steht dann  $=\text{LN}(\$B6)$ , in  $\$E7$  steht dann  $=\text{LN}(\$B7)$  usw.

Weiterhin ist das Makro ModNewton zum Finden der maximalen Mutmaßlichkeit auszuführen<sup>67</sup>. Es wendet das modifizierte Newton-Verfahren mit zwei Variablen an. Dieses Makro befindet sich nicht standardmäßig im Lieferumfang des Tabellenkalkulationsprogramms OpenOffice.org Calc. Es wurde vom Autor der vorliegenden Arbeit erstellt<sup>68</sup> und muss implementiert werden. Es folgt der Quellcode (Programmiersprache: OpenOffice.org Basic).

```
Sub ModNewton
Dim weiter As Boolean
Dim Infgr As Double
Dim nds As Double
Dim w3 As Double
Dim pi As Double
Dim genau As Double
Dim banf As Double
Dim sanf As Double
Dim b As Double
Dim s As Double
Dim o As Double
Dim p As Double
Dim f As Double
Dim g As Double
Dim fb As Double
Dim fs As Double
Dim gb As Double
Dim gs As Double
Dim fbs As Double
Dim fbb As Double
Dim fss As Double
Dim gbs As Double
Dim gbb As Double
Dim gss As Double
Dim fsb As Double
Dim gsb As Double
Dim hml As Double
Dim hpo As Double
```

<sup>67</sup> Falls Microsoft Excel als Tabellenkalkulationsprogramm verwendet wird, ist ein Makro mit der beschriebenen Funktion nicht erforderlich: Es kann hier das „Add In“ namens „Solver“ verwendet werden; Einstellungen: „Zielzelle“:  $\$O\$3$ ; „Wert“: 0; „Zu verändernde Zellen“:  $\$O\$7:\$O\$8$ ; „Randbedingungen“:  $\$O\$4=0$ .

<sup>68</sup> Mathematischer Hintergrund: Kapitel 4.4.5.

```
Dim mmf As Double
Dim mmg As Double
Dim i As Integer
Dim zeile As Integer
Dim treff As Integer
Dim verf As Integer
Dim zaehl As Integer
Dim zaehlmax As Integer
Dim oSheet As Object
Dim oCell As Object

oSheet = ThisComponent.CurrentController.ActiveSheet
zaehl = 0
zaehlmax = 20000
nds = oSheet.getCellByPosition(14,14).getValue()
banf = oSheet.getCellByPosition(14,4).getValue()
sanf = oSheet.getCellByPosition(14,5).getValue()
b = banf
s = sanf
w3 = Sqr(3.0)
pi = 3.141592654
genau = 0.00001
f = 0.0
g = 0.0
fb = 0.0
fs = 0.0
gb = 0.0
gs = 0.0
fbs = 0.0
fbb = 0.0
fss = 0.0
gbs = 0.0
gbb = 0.0
gss = 0.0
weiter = TRUE

While (weiter = TRUE)

f = 0.0
g = 0.0
fb = 0.0
fs = 0.0
gb = 0.0
gs = 0.0
fbs = 0.0
fbb = 0.0
fss = 0.0
gbs = 0.0
gbb = 0.0
gss = 0.0

For i = 1 To nds

zeile = i + 4 - 1
treff = oSheet.getCellByPosition(2,zeile).getValue()
```



```

verf = oSheet.getCellByPosition(3,zeile).getValue()
Infgr = oSheet.getCellByPosition(4,zeile).getValue()
o = Exp(pi/w3*(Infgr-b)/s)
p = o/(o + 1.0)
hml = treff + verf
hpo = hml*p/(1.0 + o)
f = f + treff - p*hml
g = g + (treff - p*hml)*Infgr
fb = fb + pi/w3/s*hpo
gb = gb + Infgr*pi/w3/s*hpo
fs = fs + hpo*Log(o)/s
gs = gs + Infgr*hpo*Log(o)/s
gbb = gbb + Infgr*(-1.0)*pi*pi/3.0/s/s*hpo/(1.0 + o)*(1.0 - o)
fbb = fbb + (-1.0)*pi*pi/3.0/s/s*hpo/(1.0 + o)*(1.0 - o)
fss = fss + (-1.0)*hpo*Log(o)*(2 + Log(o) - 2*Log(o)*p)/s/s
gss = gss + Infgr*(-1.0)*hpo*Log(o)*(2.0 + Log(o) - 2.0*Log(o)*p)/s/s
fbs = fbs + (-1.0)*pi/w3*hpo*(1.0 + Log(o) - 2.0*Log(o)*p)/s/s
gbs = gbs + Infgr*(-1.0)*pi/w3*hpo*(1.0 + Log(o) - 2.0*Log(o)*p)/s/s

```

Next i

```

zaehl = zaehl + 1
If abs(f)+abs(g) <= genau OR zaehl > zaehlmax Then
weiter = FALSE
Else
fsb = fbs
gsb = gbs
If zaehl MOD 2 > 0.1 Then
mmf = f*((fbb*fb + fbs*fs)*fb + (fsb*fb + fss*fs)*fs)/(fb*fb + fs*fs)/(fb*fb + fs*fs)
b = b - 1.0/(1.0 - mmf)*f*fb/(fb*fb + fs*fs)
s = s - 1.0/(1 - mmf)*f*fs/(fb*fb + fs*fs)
Else
mmg = g*((gbb*gb + gbs*gs)*gb + (gsb*gb + gss*gs)*gs)/(gb*gb + gs*gs)/(gb*gb +
gs*gs)
b = b - 1.0/(1.0 - mmg)*g*gb/(gb*gb + gs*gs)
s = s - 1.0/(1.0 - mmg)*g*gs/(gb*gb + gs*gs)
End If
End If

```

Wend

```

oCell = oSheet.getCellByPosition(14,6)
oCell.setValue(b)
oCell = oSheet.getCellByPosition(14,7)
oCell.setValue(s)
End Sub

```

## Verwendete Abkürzungen und Symbole

a	1. Häufigkeitsmerkmal 2. (Modell-)Parameter
$a_1$	(Modell-)Parameter
$a_2$	(Modell-)Parameter
$a_{90 95}$	Häufigkeitsmerkmal der zuverlässigen Bekanntheit
A	1. Adjektiv 2. (Modell-)Parameter 3. Zahl der agglutinierenden Konstruktionen
Adv	Adverb
AdvP	Adverbphrase
AP	Adjektivphrase
ART	Artikel
A/J	Index der Agglutination
b	(Modell-)Parameter
B	(Modell-)Parameter
Bewege- $\alpha$	Bewegungstransformation
c	(Modell-)Parameter
C	1. Konjunktion (= Kopf einer durch eine Konjunktion eingeleiteten Flexionsphrase CP) 2. Modell-)Parameter
CP	durch eine Konjunktion eingeleitete Flexionsphrase
d	Mittelwert der Häufigkeiten (bestimmter Wörter, z. B. mit Häufigkeitsmerkmal a) im Dietze-Korpus
D	1. Bestimmtheitsmaß 2. Devianz
E	Matrix (im Gauß-Newton-Verfahren); $E = K L$
f'	erste Ableitung von f
f''	zweite Ableitung von f
$f_1$	1. Freiheitsgrad 1 2. relative Häufigkeit des empirisch ermittelten häufigsten Phonems



j	Esperanto-Lernzeit
J	1. Jacobi-Matrix (im Gauß-Newton-Verfahren) 2. Zahl der Morph-Junktoren
k	Index (für Nullstellen)
L	1. Matrix (im Gauß-Newton-Verfahren); $L = J^T R$ 2. Lexemlänge (in Silbenzahl)
$L_{th}$	mit einem Modell berechnete Lexemlänge (in Silbenzahl)
$k_0$	Parameter im Modell der logistischen Regression (für Vertrauensgrad-Berechnung)
$k_1$	Parameter im Modell der logistischen Regression (für Vertrauensgrad-Berechnung)
$k_2$	Parameter im Modell der logistischen Regression (für Vertrauensgrad-Berechnung)
K	Matrix (im Gauß-Newton-Verfahren); $K = (J^T J)^{-1}$
L	1. Länge eines Lexems, ausgedrückt durch die Anzahl seiner Silben 2. Mutmaßlichkeit („likelihood“)
$L_0$	maximale Mutmaßlichkeit („maximum likelihood“)
m	Anzahl der Fragebogen
$m_k$	Hilfsgröße (beim modifizierten Newton-Verfahren)
M	Zahl der Morpheme
M/W	Isoliertheitsgrad
n	1. Anzahl 2. Stichprobenumfang
N	1. Anzahl 2. Stichprobenumfang 3. Kumulierte Wortzahl 4. Nomen, allg. auch lexikalische Ebene der Nominalphrase (kann auch durch ein Pronomen lexikalisch realisiert sein)
$N_R$	Anzahl der Rangwerte (im Kruskal-Wallis-Test)
$N'$	Parameter im Modell der logistischen Regression (für Vertrauensgrad-Berechnung)
NEG	Negation
NP	Nominalphrase

O	Zahl der nur durch ihre Stellung signalisierten syntaktischen Konnexionen
O/N	Index der Isolation
p	Bekanntheitsgrad
p'	Bekanntheitswahrscheinlichkeit im Modell der logistischen Regression
p <sub>0</sub>	Bekanntheitsgrad ohne Vertrauensgrad im Modell der logistischen Regression
p <sub>0, k</sub>	Bekanntheitsgrad in Anlehnung an ein Modell aus Köhler, Rapp (2007)
p <sub>1</sub>	Anteil der bekannten Wörter
p <sub>29, v</sub>	Bekanntheitsgrad bei Vertrauensgrad V nach der 29/29-Methode
p <sub>v</sub>	Bekanntheitsgrad bei Vertrauensgrad V im Modell der logistischen Regression
p <sub>95%</sub>	Bekanntheitsgrad bei Vertrauensgrad 95% im Modell der logistischen Regression
P	1. Präposition 2. Polysemie (Anzahl) 3. Zahl der Präfixe
PP	Präpositionalphrase
P/M	Index der Präfixation
q	1. (Modell-)Parameter 2. Quantil der Standard-Normalverteilung
R	1. Rangsumme (im Kruskal-Wallis-Test) 2. Residuenmatrix (im Gauß-Newton-Verfahren)
r	Residuum
r <sup>2</sup>	Abweichungsquadratsumme
s	Selbsteinschätzung des eigenen Esperanto-Anwendungsvermögens
S	1. Startsymbol (konventionelle Phrasenstrukturgrammatik) 2. Parameter-Vektor (im Gauß-Newton-Verfahren; Startwerte) 3. Zahl der Suffixe
S/M	Index der Suffixation
Spez	Spezifizierer

V	1. Vertrauensgrad (Vertrauensniveau, Konfidenzniveau) 2. Parameter-Vektor (im Gauß-Newton-Verfahren; verbesserte Werte) 3. Verb
VP	Verbalphrase
W	Zahl der Wörter
W/M	Index des Synthetismus
x	1. Zahl der Gekannt-Aussagen in der 29/29-Methode 2. Rang 3. Worthäufigkeit
$x_k$	genäherte Nullstelle k im modifizierten Newton-Verfahren
$x_{k+1}$	verbesserte genäherte Nullstelle k+1 im modifizierten Newton-Verfahren
X	Kopf einer Phrase; ist außer bei Flexionsphrase IP lexikalisch realisiert; X steht für A, Adv, C, I, N, P oder V
y	1. Dichotomie-Merkmal ( $y = 1$ oder $y = 0$ ) 2. Wortvertrautheit 3. Häufigkeit
$\gamma$	Parameter zur Vertrauensgrad-Berechnung im Modell der logistischen Regression
$\Gamma$	Gamma-Funktion
$\mu$	Mittelwert-Parameter
$\mu_0$	Mittelwert-Parameter bei maximaler Mutmaßlichkeit
$\mu_1$	Startwert für die Bestimmung von $\mu_0$
$\sigma$	Streuungs-Parameter
$\sigma_0$	Streuungs-Parameter bei maximaler Mutmaßlichkeit
$\sigma_1$	Startwert für die Bestimmung von $\sigma_0$
$\Phi$	kumulative Standard-Normalverteilung
$\chi^2$	Chi-Quadrat
$\langle \rangle$	Mittelwert
->	Ersetzung
	1. Übergang (z. B. von NP nach N') ohne Links- oder Rechtsverzweigung 2. syntaktische Konnexion

/	1. Übergang mit Linksverzweigung 2. syntaktische Konnexion
\	1. Übergang mit Rechtsverzweigung 2. syntaktische Konnexion
/_\	Auslassung der Ebenen X' und X; steht zwischen XP und der Oberflächenstruktur
'	„Bar“: Ebene zwischen Phrase XP und Wortkategorie X, z. B. bezeichnet N' die Ebene zwischen NP und N
∅	Leere Menge
[Prät.]	Präteritum

## Abbildungsverzeichnis

Abbildung 1: Herkömmliches Phrasenstrukturdiagramm eines Esperanto-Satzes .....	28
Abbildung 2: Der Esperanto-Satz von Abbildung 1 in X-Bar-Konvention.....	29
Abbildung 3: Gesamtgrundstruktur einer Phrase.....	30
Abbildung 4: Nominalphrase (Beispiel).....	30
Abbildung 5: Anzahl der Wörter gegen den Rang in der Offizialisierungsfolge; schraffiert: tatsächliche Zahlen; unshraffiert: nach Modellfunktion (aus: Kück 2008: 19).....	33
Abbildung 6: Relative Phonemhäufigkeit des Esperanto gegen den Rang sowie Anpassung mit dem Yule-Modell.....	49
Abbildung 7: Syntaktische Konnexionen in einem Beispielsatz.....	54
Abbildung 8: Häufigkeit der 20 häufigsten Morphemstrukturen des Esperanto (vgl. Kapitel 2.3.10, Tabelle 4) gegen den Rang sowie Modellierung nach Gleichung 9.....	59
Abbildung 9: Flexionsphrase.....	60
Abbildung 10: Flexionsphrase mit rechts eingebetteter Subjekt-Nominalphrase .....	61
Abbildung 11: Links eingebettete Objekt-Nominalphrase in einer Verbalphrase .....	62
Abbildung 12: Rechts eingebettete Adjektivphrase bei Korrelativwort.....	64
Abbildung 13: Rechts eingebettete komplexe Adjektivphrase.....	65
Abbildung 14: Rechts eingebettete Präpositionalphrase.....	66
Abbildung 15: Rechts vom Bezugssubstantiv eingebettete Präpositionalphrase .....	67
Abbildung 16: Links vom Bezugssubstantiv eingebettete Präpositionalphrase.	68
Abbildung 17: Adverbphrase als rechts eingebettetes Adjunkt in einer Verbalphrase.....	69
Abbildung 18: Adverbphrase als links eingebettetes Adjunkt in einer Verbalphrase.....	70
Abbildung 19: Negation ne als links eingebettetes Adjunkt in einer Verbalphrase .....	71



Abbildung 20: Negation ne als rechts eingebettetes Adjunkt in einer Verbalphrase (bei Kontrast).....	72
Abbildung 21: Rechtseinbettung von Relativsätzen im Esperanto.....	73
Abbildung 22: Linkseinbettung von Relativsätzen im Japanischen.....	74
Abbildung 23: Kumulierte Wortzahl N gegen Rang i in der Offizialisierungsfolge; gestrichelte farblose Säulen: N (aus Tabelle 5 ermittelt, dabei nur „Fundamento“ [i = 1] und „Beifügungen“ berücksichtigt); graue Säulen: N (modelliert nach Gleichung 10).....	77
Abbildung 24: Länge von Lexemen als Funktion der Polysemie im Esperanto.	78
Abbildung 25: Mittelwert d der Häufigkeiten im Dietze-Korpus gegen Häufigkeitsmerkmal a.....	80
Abbildung 26: Annahmen im Berens-Modell; in der Vorlage (Berens 1989) steht „defect size“ statt „Häufigkeitsmerkmal“ und „probability of detection“ statt „Bekanntheitsgrad“.....	88
Abbildung 27: Fragebogen A (Quellcode: Anhang E).....	102
Abbildung 28: Fragebogen B.....	104
Abbildung 29: Fragebogen C.....	104
Abbildung 30: Fragebogen D.....	105
Abbildung 31: Fragebogen E.....	105
Abbildung 32: Bekanntheitsgrad p gegen Häufigkeitsmerkmal a; Grundlage: Fragebogen E (vgl. Abbildung 27); gestrichelte farblose Säulen: Bekanntheitsgrad nach der Vorgehensweise von Kapitel 4.4.2 (Schritt 3; also ohne Vertrauensgrad); graue Säulen: Bekanntheitsgrad bei 95% Vertrauensgrad.....	108
Abbildung 33: Bekanntheitsgrad p gegen Häufigkeitsmerkmal a; Grundlage: Fragebogen A bis E; gestrichelte farblose Säulen: Bekanntheitsgrad nach der Vorgehensweise von Kapitel 4.4.2 (Schritt 3; also ohne Vertrauensgrad); graue Säulen: Bekanntheitsgrad bei 95% Vertrauensgrad.....	112
Abbildung 34: Bekanntheitsgrad p gegen Häufigkeitsmerkmal a; bei Muttersprache aus nicht-indogermanischer Sprachfamilie; gestrichelte farblose Säulen: Bekanntheitsgrad nach der Vorgehensweise von Kapitel 4.4.2 (Schritt 3; also ohne Vertrauensgrad); graue Säulen: Bekanntheitsgrad bei	

	95%Vertrauensgrad.....	113
Abbildung 35:	Bekanntheitsgrad p gegen Häufigkeitsmerkmal a; bei Muttersprache aus indogermanischer Sprachfamilie; gestrichelte farblose Säulen: Bekanntheitsgrad nach der Vorgehensweise von Kapitel 4.4.2 (Schritt 3; also ohne Vertrauensgrad); graue Säulen: Bekanntheitsgrad bei 95%Vertrauensgrad.....	114
Abbildung 36:	Bekanntheitsgrad p gegen Häufigkeitsmerkmal a; bei Muttersprache Chinesisch oder Japanisch; gestrichelte farblose Säulen: Bekanntheitsgrad nach der Vorgehensweise von Kapitel 4.4.2 (Schritt 3; also ohne Vertrauensgrad); graue Säulen: Bekanntheitsgrad bei 95%Vertrauensgrad.....	115
Abbildung 37:	Bekanntheitsgrad p gegen Häufigkeitsmerkmal a; bei Muttersprache Japanisch; gestrichelte farblose Säulen: Bekanntheitsgrad nach der Vorgehensweise von Kapitel 4.4.2 (Schritt 3; also ohne Vertrauensgrad); graue Säulen: Bekanntheitsgrad bei 95%Vertrauensgrad.....	116
Abbildung 38:	Bekanntheitsgrad p gegen Häufigkeitsmerkmal a; bei Muttersprache aus uralischer Sprachfamilie; gestrichelte farblose Säulen: Bekanntheitsgrad nach der Vorgehensweise von Kapitel 4.4.2 (Schritt 3; also ohne Vertrauensgrad); graue Säulen: Bekanntheitsgrad bei 95%Vertrauensgrad.....	117
Abbildung 39:	Bekanntheitsgrad p gegen Häufigkeitsmerkmal a; bei Muttersprache aus baltischem oder slawischem Sprachzweig; gestrichelte farblose Säulen: Bekanntheitsgrad nach der Vorgehensweise von Kapitel 4.4.2 (Schritt 3; also ohne Vertrauensgrad), graue Säulen: Bekanntheitsgrad bei 95%Vertrauensgrad.....	118
Abbildung 40:	Bekanntheitsgrad p gegen Häufigkeitsmerkmal a; bei Muttersprache aus germanischem oder romanischem Sprachzweig; gestrichelte farblose Säulen: Bekanntheitsgrad nach der Vorgehensweise von Kapitel 4.4.2 (Schritt 3; also ohne Vertrauensgrad), graue Säulen: Bekanntheitsgrad bei 95%Vertrauensgrad.....	119
Abbildung 41:	Bekanntheitsgrad p gegen Häufigkeitsmerkmal a; bei	

Muttersprache aus romanischem Sprachzweig; gestrichelte farblose Säulen: Bekanntheitsgrad nach der Vorgehensweise von Kapitel 4.4.2 (Schritt 3; also ohne Vertrauensgrad), graue Säulen: Bekanntheitsgrad bei 95%Vertrauensgrad.....	120
Abbildung 42: Selbsteinschätzung des eigenen Esperanto-Anwendungsvermögens s gegen Esperanto-Lernzeit j; durchgezogene Linie: Vertrauensgrad 95%; gestrichelte Linie: ohne Vertrauensgrad.....	123
Abbildung 43: Vergleich der mit Gleichung 15 (gestrichelte farblose Säulen) und Gleichung 99 (graue Säulen) gegebenen Modelle.....	125
Abbildung 44: Länge von Lexemen als Funktion der Polysemie im Esperanto .....	140
Abbildung 45: Auswertung von Fragebogen A.....	160
Abbildung 46: Auswertung von Fragebogen B.....	161
Abbildung 47: Auswertung von Fragebogen C.....	162
Abbildung 48: Auswertung von Fragebogen D.....	163
Abbildung 49: Auswertung von Fragebogen E.....	164
Abbildung 50: Auswertung der Fragebogen A bis E.....	165
Abbildung 51: Auswertung, Muttersprache aus nicht-indoeuropäischer Sprachfamilie.....	166
Abbildung 52: Auswertung, Muttersprache aus indoeuropäischer Sprachfamilie .....	167
Abbildung 53: Auswertung, Muttersprache Chinesisch oder Japanisch.....	168
Abbildung 54: Auswertung, Muttersprache Japanisch.....	169
Abbildung 55: Auswertung, Muttersprache aus uralischer Sprachfamilie.....	170
Abbildung 56: Auswertung, Muttersprache aus baltischem oder slawischem Sprachzweig.....	171
Abbildung 57: Auswertung, Muttersprache aus germanischem oder romanischem Sprachzweig.....	172
Abbildung 58: Auswertung, Muttersprache aus romanischem Sprachzweig...	173
Abbildung 59: Selbsteinschätzung des eigenen Esperanto-Anwendungsvermögens s und Esperanto-Lernzeit j (Teil 1 von 2) .....	174
Abbildung 60: Selbsteinschätzung des eigenen Esperanto-	

Anwendungsvermögens s und Esperanto-Lernzeit j (Teil 2 von 2)	175
Abbildung 61: Modellvergleich.....	176
Abbildung 62: Zelle \$E5.....	177
Abbildung 63: Zelle \$H5.....	177
Abbildung 64: Zelle \$I5.....	177
Abbildung 65: Zelle \$J5.....	178
Abbildung 66: Zelle \$K5.....	178
Abbildung 67: Zelle \$L5.....	178
Abbildung 68: Zelle \$M5.....	178
Abbildung 69: Zelle \$O\$2.....	179
Abbildung 70: Zelle \$O\$3.....	179
Abbildung 71: Zelle \$O\$4.....	179
Abbildung 72: Zelle \$O\$5.....	179
Abbildung 73: Zelle \$O\$6.....	180
Abbildung 74: Zelle \$O\$9.....	180
Abbildung 75: Zelle \$O\$12.....	180
Abbildung 76: Zelle \$O\$13.....	181
Abbildung 77: Zelle \$O\$14.....	181
Abbildung 78: Zelle \$O\$16.....	181
Abbildung 79: Zelle \$O\$17.....	181
Abbildung 80: Zelle \$O\$18.....	181
Abbildung 81: Zelle \$Q3.....	181
Abbildung 82: Zelle \$R3.....	181
Abbildung 83: Zelle \$S3.....	182
Abbildung 84: Zelle \$F6.....	182
Abbildung 85: Zelle \$G6.....	182
Abbildung 86: Zelle \$C\$14.....	182

## Tabellenverzeichnis

Tabelle 1: Phoneminventar des Esperanto (erarbeitet unter Rückgriff auf Kalocsay, Waringhien 1980 <sup>4</sup> : 39ff., Wells 1978: 15ff. und Wennergren 2006: 21ff. mit Artikulationsort-Bezeichnungen in Anlehnung an die von International Phonetic Association 2008 zusammengestellten).....	15
Tabelle 2: Phonem-Häufigkeitsrangliste des Esperanto (erarbeitet unter Rückgriff auf die Ergebnisse von Haszpra 2001).....	17
Tabelle 3: Korrelativwörter des Esperanto; das ihnen allen gemeinsame Korrelativitäts-Morphem wird durch das Morph /i/ realisiert.....	25
Tabelle 4: Die 20 häufigsten Morphemstrukturen im Esperanto; Stern (*): lexikalisches Morphem, das kein autosemantisches Affix ist (Blahuš 2008: 24).....	26
Tabelle 5: „Fundamento“- und offizielle Esperanto-Wörter; Anzahl (Übersicht)	32
Tabelle 6: Dietze-Korpus: statistische Merkmale (erarbeitet unter Rückgriff auf die Angaben von Dietze 1989: 4ff.).....	44
Tabelle 7: Anwendung verschiedener Modelle zur Beschreibung der relativen Phonemhäufigkeit im Esperanto.....	48
Tabelle 8: Ausgewählte morphologische Indizes von „Klasse D“-Sprachen und Esperanto (die Werte der Indizes [mit Ausnahme der für Esperanto genannten] entstammen Altmann, Lehfeldt 1973: 40).....	57
Tabelle 9: Zusammenfassung der Ergebnisse zur Einbettung von Phrasen und Relativsätzen.....	76
Tabelle 10: Häufigkeiten der BRO-Wörter im Dietze-Korpus.....	79
Tabelle 11: Analogie zwischen Sprachwissenschaft und Werkstoffprüftechnik.	82
Tabelle 12: Parameter $\gamma$ in Abhängigkeit von n bei 95% Vertrauensgrad (bestimmt nach Gleichung 32; Hinweis: Dezimalpunkt statt Dezimalkomma verwendet).....	94
Tabelle 13: Häufigkeitsmerkmal mit zuverlässiger Bekanntheit, ermittelt aus den Fragebogen A bis E.....	109
Tabelle 14: Hilfstabelle für Kruskal-Wallis-Test auf Ähnlichkeit.....	111
Tabelle 15: Zusammenfassung von Gesamt- und Einzelauswertungen.....	121

Tabelle 16: Tabellenkalkulationsblatt zur Anpassung des Yule-Modells an die empirisch erhaltene relative Phonemhäufigkeit mit Hilfe des Gauß-Newton-Verfahrens (monophonematische Wertung von Affrikaten)	134
Tabelle 17: Tabellenkalkulationsblatt zur Anpassung des Yule-Modells an die empirisch erhaltene relative Phonemhäufigkeit mit Hilfe des Gauß-Newton-Verfahrens (biphonematische Wertung von Affrikaten)...	135
Tabelle 18: Tabellenkalkulationsblatt zur Anpassung des Zipfschen Gesetzes an die Häufigkeitsverteilung von Morphemstrukturen des Esperanto	136
Tabelle 19: Länge und Polysemie von 100 Lexemen (Teil 1 von 3).....	137
Tabelle 20: Länge und Polysemie von 100 Lexemen (Teil 2 von 3).....	138
Tabelle 21: Länge und Polysemie von 100 Lexemen (Teil 3 von 3).....	139
Tabelle 22: Länge und Polysemie von 100 Lexemen; Auswertung; die Anpassung erfolgte mittels OpenOffice Calc 2.3, Zusatzprogramm („Extension“) EuroSolver Optimierer 1.6 (Option: „Das nichtlineare Verfahren“).....	140
Tabelle 23: Fragebogen A; Inhalt (Teil 1 von 3).....	145
Tabelle 24: Fragebogen A; Inhalt (Teil 2 von 3).....	146
Tabelle 25: Fragebogen A; Inhalt (Teil 3 von 3).....	147
Tabelle 26: Fragebogen B; Inhalt (Teil 1 von 3).....	148
Tabelle 27: Fragebogen B; Inhalt (Teil 2 von 3).....	149
Tabelle 28: Fragebogen B; Inhalt (Teil 3 von 3).....	150
Tabelle 29: Fragebogen C; Inhalt (Teil 1 von 3).....	151
Tabelle 30: Fragebogen C; Inhalt (Teil 2 von 3).....	152
Tabelle 31: Fragebogen C; Inhalt (Teil 3 von 3).....	153
Tabelle 32: Fragebogen D; Inhalt (Teil 1 von 3).....	154
Tabelle 33: Fragebogen D; Inhalt (Teil 2 von 3).....	155
Tabelle 34: Fragebogen D; Inhalt (Teil 3 von 3).....	156
Tabelle 35: Fragebogen E; Inhalt (Teil 1 von 3).....	157
Tabelle 36: Fragebogen E; Inhalt (Teil 2 von 3).....	158
Tabelle 37: Fragebogen E; Inhalt (Teil 3 von 3).....	159

## Literaturverzeichnis

Air Force Research Laboratory (Hrsg.) (2002): *Handbook for Damage Tolerant Design*. Dayton, Air Force Research Lab, <http://www.siresearch.info/projects/dtdh/pdfs.php> 05.11.2007.

Akademio de Esperanto (Hrsg.) (1958): *Sesa Oficiala Aldono al la Universala Vortaro. Oficiala Bulteno de la Akademio de Esperanto N-ro 1*. [ohne Ort].

Akademio de Esperanto (Hrsg.) (2007<sup>2</sup>): *Oficiala Bulteno de la Akademio de Esperanto – N-ro 10 – Aktoj de la Akademio II 1968-1974*. Rotterdam, Akademio de Esperanto.

Altmann, Gabriel (1993): Phoneme counts. In: *Glottometrika* 14, 55-70.

Altmann, Gabriel; Beőthy, Erzsėbeth; Best, Karl-Heinz (1982): Die Bedeutungskomplexitat der Worters und das Menzerathsche Gesetz. In: *Zs. fur Phonetik, Sprachwissenschaft und Kommunikationsforschung* 35, 537-543.

Altmann, Gabriel; Lehfeldt, Werner (1973): *Allgemeine Sprachtypologie. Prinzipien und Meverfahren*. Munchen, Fink.

Annis, Charles (2001): *StatisticalEngineering*, <http://www.statisticalengineering.com/glm2.htm> 05.11.2007.

Atanasov, Atanas D. (1983): *La lingva esenco de Esperanto*. Rotterdam, Universala Esperanto-Asocio.

Baayen, R. Harald (2001): *Word frequency distributions*. Dordrecht (u. a.), Kluwer.

Bailey, William; Rollet de l'Isle, Maurice (1935): *Sesa Oficiala Aldono al la Universala Vortaro. Oficialaj sciigoj de la Akademio*, Nr. 9. Paris, Centra Presejo Esperantista.

Bak, Giuan (1991): *Hanguowa esuperanthoui hyongthae taejo yongu*. A Study of Morphological Contrast of the Korean and Esperanto. Studo pri Morfologia Kontrasto de la Korea kaj Esperanto [Seoul: Konkuk University], Seoul, Seula Esperanto-Kulturcentro.

Bastien, Louis; Rollet de l'Isle, Maurice (1934): *Kvina Oficiala Aldono al Universala Vortaro. Oficialaj sciigoj de la Akademio* Nr. 4. Paris, Centra Presejo Esperantista.

Bausani, Alessandro (1970): *Geheim- und Universalsprachen. Entwicklung und Typologie*. Stuttgart, Kohlhammer.

Berens, Alan P. (1989): *NDE reliability data analysis. Metals Handbook*, Vol 17: *Nondestructive Evaluation and Quality Control*. Metals Park (Ohio), ASM International, 689-701.

Berens, Alan P.; Hovey, Peter W. (1981): *Evaluation of NDE Reliability Characterization*. Vol I. Dayton (Ohio), Dayton University Ohio Research Institute.

Berlin, B.; Kay, P. (1969): *Basic color terms: their universality and evolution*. Berkeley, University of California Press.

Björck, Åke (1996): *Numerical methods for least squares problems*. Philadelphia, SIAM.

Blaas, Leo (1951): Statistiko de 50 000 tekstvortoj [Statistik von 50 000 Textwörtern]. In: *Esperantologio*, Vol I, 2 (06/1950), 107-132, und 3 (03/1951), 160-200.

Blahuš, Marek (2008): *A Spell Checker for Esperanto*. Bachelor Thesis, Faculty of Informatics, Masaryk University, Brno.

Blanke, Detlev (1981): Plansprache und Nationalsprache. Einige Probleme der Wortbildung des Esperanto und des Deutschen in konfrontativer Darstellung. Berlin, *Linguistische Studien 85*, Akademie der Wissenschaften der DDR (Zentralinstitut für Sprachwissenschaft).

Blanke, Detlev (1986<sup>2</sup>): *Esperanto und Wissenschaft*. Berlin, Esperanto-Verband im Kulturbund der DDR.

Bovet, Pierre; Ith, Henriette (1956): zitiert nach Akademio de Esperanto (Hrsg.) (2007<sup>2</sup>: 25).

Bußmann, Hadumod (2002<sup>3</sup>): *Lexikon der Sprachwissenschaft*. Stuttgart, Kröner.



- Carroll, John B. (1967): On sampling from a lognormal model of word frequency distribution. In: Kučera, Henry; Francis, W. Nelson (Hrsg.): *Computational Analysis of Present-Day American English*, Brown University Press, Providence.
- Cart, Théophile (1909): Unua Oficiala Aldono al Universala Vortaro. In: *Esperantista Dokumentaro*, 13. Paris, Esperantista Centra Oficejo, 7-31.
- Cart, Théophile (1919): Dua Oficiala Aldono al Universala Vortaro. In: *Esperantista Dokumentaro*, 30. Paris, Esperantista Centra Oficejo, 3-12.
- Cheng, Russel C. H.; Iles, T. C. (1983): Confidence Bands for Cumulative Distribution Functions of Continuous Random Variables. In: *Technometrics*, 1, 25, 77-86.
- Cheng, Russel C. H.; Iles, T. C. (1988): One-Sided Confidence Bands for Cumulative Distribution Functions. In: *Technometrics*, 2, 30, 155-159.
- Chomsky, Noam (1970): Remarks on nominalization. In: Jacobs, R.; Rosenbaum, P. (Hrsg.): *Reading in English Transformational Grammar*, 184-221. Waltham: Ginn.
- Corsetti, Renato; Dasgupta, Probal; Mattos, Geraldo; Moon, Brian (2007): Pri apartaj teknikaj bezonoj rilate al niaj alfabeto kaj ortografio. In: *Akademio de Esperanto – Oficialaj Informoj - N-ro 6 2007*, [http://www.akademio-de-esperanto.org/oficialaj\\_informoj/oficialaj\\_informoj\\_6\\_2007.html?s=h](http://www.akademio-de-esperanto.org/oficialaj_informoj/oficialaj_informoj_6_2007.html?s=h) 05.11.2007.
- Cox, David Roxbee (1970): *The analysis of binary data*. London, Methuen (*Methuen's monographs on applied probability and statistics*).
- Deutsches Institut für Normung (Hrsg.) (1982): *DIN 55350-24 1982-11 Begriffe der Qualitätssicherung und Statistik – Begriffe der Statistik – Schließende Statistik*. Berlin, Beuth Verlag.
- Deutsches Institut für Normung (Hrsg.) (1985): *DIN 53804-4:1985-03 Statistische Auswertungen, Attributmerkmale*. Berlin, Beuth Verlag.
- Dietze, Joachim (1989): *Frequenzwörterbuch Esperanto-Deutsch. Die meistgebrauchten Wurzeln der Esperanto-Literatursprache*. Halle (Saale), Wissenschaftspublizistik der Martin-Luther-Universität Halle-Wittenberg.

Dominte, Constantin (2001): Fonemstatistikaj rimarkoj pri Esperanto. In: Fiedler, Sabine; Haitao, Liu (Hrsg.): *Studien zur Interlinguistik. Festschrift für Detlev Blanke zum 60. Geburtstag*. Dobřichovice (Prag), KAVA-PECH.

Duc Goninaz, Michel (1987a): Kiel priskribi Esperanton? Problemoj metodologiaj kaj terminologiaj. In: *Serta Gratulatoria in honorem Juam Régulo. II: Esperantismo*. La Laguna, Universidad de la Laguna 1987, 141-149.

Duc Goninaz, Michel (1987b): Problèmes sémantiques et sociolinguistiques dans l'usage de l'espéranto. In: Duc Goninaz, M. (Hrsg.): *Studioj pri la Internacia Lingvo*. Gent, AIMAV 1987, 139-143.

Đurčo, Peter (2001): Bekanntheit, Häufigkeit und lexikographische Erfassung von Sprichwörtern. Zu parömiologischen Minima für DaF. In: Häcki Buhofer, Annelies; Burger, Harald; Gautier, Laurent (Hrsg.): *Phraseologiae Amor: Aspekte europäischer Phraseologie*. Hohengehren, Schneider Verlag, 99-106.

Fiedler, Sabine (1999): *Plansprache und Phraseologie: Empirische Untersuchungen zu reproduziertem Sprachmaterial im Esperanto*. Frankfurt am Main, Peter Lang.

Fiedler, Sabine (2002): *Esperanta frazeologio*. Rotterdam, Universala Esperanto-Asocio.

Georgiou, George A. (2006): *Probability of Detection (PoD) curves. Derivation, applications and limitations (Research Report 454)*. London, Jacobi Consulting Limited, <http://www.hse.gov.uk/RESEARCH/rrpdf/rr454.pdf> 05.11.2007.

Gledhill, Christopher (2000<sup>2</sup>): *The Grammar of Esperanto. A corpus-based description*. München, Lincom Europa.

Good, Irving John (1969): Statistics of language. In: Meethoun, A. R., Hudson, R. A. (Hrsg.): *Encyclopedia of information, linguistics and control*, 567-581. Oxford, Pergamon.

Gottwald, Wolfgang (2000): *Statistik für Anwender*. Weinheim, Berlin, New York et al., Wiley-VCH.

Greenberg, Joseph H. (1960): A quantitative approach to the morphological typology of language. *International Journal of American Linguistics*, 26, 178-194.

Greenberg, Joseph H. (1963): *Universals of language*. Cambridge, Mass., MIT Press.

Greenberg, Joseph H. (1974): Language typology: a historical and analytical overview. Den Haag, Mouton.

Grosjean-Maupin, Émile (1910): *Dictionnaire complet Espéranto-français*. Paris, Hachette.

Grosjean-Maupin, Émile (1922): Tria Oficiala Aldono al Universala Vortaro. In: *Oficiala Gazeto Esperantista*, Vol 8, 5-6 (82-83). Paris, Esperantista Centra Oficejo, 153-164.

Grosjean-Maupin, Émile (1929): Kvara Aldono al U. V. – Oficiala Listo. In: *Oficiala Bulteno de la Esperantista Akademio*. I (06/1929), 4-8 und II (02/1930), 25-29.

Hagège, Claude (1985): *L'homme de paroles*. Paris, Fayard.

Hallsteinsdóttir, Erla; Šajánková, Monika; Quasthoff, Uwe (2006): Phraseologisches Optimum für Deutsch als Fremdsprache. Ein Vorschlag auf der Basis von Frequenz- und Geläufigkeitsuntersuchungen. In: *Linguistik online* 27, 2/06, [http://www.linguistik-online.de/27\\_06/hallsteinsdottir\\_et\\_al.pdf](http://www.linguistik-online.de/27_06/hallsteinsdottir_et_al.pdf) 05.11.2007.

Harry, Ralph (1967): Relativaj oftecoj de lingvaj elementoj en Esperanto. In: *Scienca Revuo* 19/2, 49-50.

Haszpra, Ottó (1998): *Eszperantó az ábécétől a felsőfokig*. Budapest, HEA.

Haszpra, Ottó (2001): Liter-ofteco en Esperantaj tekstoj. In: Fiedler, Sabine; Haitao, Liu (Hrsg.): *Studien zur Interlinguistik. Festschrift für Detlev Blanke zum 60. Geburtstag*. Dobřichovice (Prag), KAVA-PECH.

Heil, Anett (1999): *Grammatische Reduktion in Franko-Kreolsprachen und Plansprachen*. (Rostocker Romanistische Arbeiten, Band 2). Frankfurt am Main, Peter Lang.

Hempel, Tino (1996): *Darstellung von Iterationsverfahren zur Nullstellenbestimmung von Funktionen und visualisierende Computerexperimente*. Hausarbeit zur Ersten Staatsprüfung für das Lehramt an Gymnasien im Fach

Numerische Mathematik, Chemnitz, Technische Universität Chemnitz-Zwickau, <http://www.tinohempel.de/info/mathe/iter/haus.pdf> 13.02.2007.

Herdan, Gustav (1960): *Type-Token Mathematics*. Den Haag, Mouton.

Hoffmann, Christiane (2005): Morphologisch orientierte Typologie. In: Köhler, Reinhard; Altmann, Gabriel; Piotrowski, Raimund G.: *Quantitative Linguistik. Ein Internationales Handbuch*, Berlin, Walter de Gruyter.

International Phonetic Association (2008): *Reproduction of The International Phonetic Alphabet (Revised to 2005)*. <http://www2.arts.gla.ac.uk/IPA/pulmonic.html> 07.08.2008.

Jackendoff, Ray (1977): *X-bar-Syntax: A Study of Phrase Structure, Linguistic Inquiry Monograph 2*. Cambridge, Mass., MIT Press.

Jansen, Wim (2007): *Woordvolgorde in het Esperanto*. Utrecht, LOT.

Janton, Pierre (1993<sup>2</sup>): *Einführung in die Esperantologie*. Hildesheim (u. a.), Olms.

Jung, Theo (1926): Statistik. La ofteco de la literoj en la Esperanto-alfabeto. In: *Heroldo de Esperanto* 48 (376).

Juska-Bacher, Britta: Phraseologische Befragungen und ihre statistische Auswertung. In: *Linguistik online* 27, 2/06, [http://www.linguistik-online.de/27\\_06/hallsteinsdottir\\_et\\_al.pdf](http://www.linguistik-online.de/27_06/hallsteinsdottir_et_al.pdf) 01.02.2007.

Kalocsay, Kálmán (1938<sup>2</sup>): *Plena gramatiko de Esperanto: kun detala komentario pri vorto kaj frazo*. Budapest, Literatura Mondo.

Kalocsay, Kálmán; Waringhien, Gaston (1980<sup>4</sup>): *Plena analiza gramatiko de Esperanto*. Rotterdam, UEA.

Kempgen, Sebastian; Lehfeldt, Werner (2004): Quantitative morphologische Typologie. In: Booij, Geert: *Morphologie. Ein Internationales Handbuch zur Flexion und Wortbildung*. Berlin, Walter de Gruyter.

Köhler, Reinhard (1990): Linguistische Analyseebenen. Hierarchisierung und Erklärung im Modell der sprachlichen Selbstregulation. In: *Glottometrika* 11, 1-18.

Köhler, Reinhard (1999). Der Zusammenhang zwischen Lexemlänge und Polysemie im Maori. In: Ondrejovič, S., Genzor, J. (Hrsg.): *Pange lingua. Zbornik na počest' Viktora Krupu*, 27-33. Bratislava, Veda.

Köhler, Reinhard; Rapp, Reinhard (2007): Word Familiarity and Frequency. A Psycholinguistic Application of Synergetic Linguistics. In: *Glottometrics* 15. Lüdenscheid, RAM-Verlag.

Krause, Erich-Dieter (1999): *Großes Wörterbuch Esperanto-Deutsch*. Hamburg, Buske.

Kruskal, William H.; Wallis, W. Allen (1952): Use of ranks in one-criterion variance analysis. In: *Journal of the American Statistical Association*, 47, 583-621.

Kück, Andreas (2007): Interessaj dokumentoj el Austrujo [Interessante Dokumente aus Österreich]. Beitrag zum Netzforum *Akademia diskuto*, <http://groups.yahoo.com/group/AdE-diskuto/message/6836> 19.03.2007.

Kück, Andreas (2008): *Oficialaj lingvoelementoj de Esperanto (OLEO)*. Rechtenfleth, Selbstverlag Andreas Kück.

Kühlmeier, Manfred (2001): *Statistische Auswertungsmethoden für Ingenieure mit Praxisbeispielen*. Berlin (u. a.), Springer.

Kuno, S. (1973): *The structure of the Japanese language*. Cambridge, Mass., MIT Press.

Langenscheidt-Verlag (Hrsg.) (1959): *Langenscheidts Lilliput-Wörterbuch Nr. 118 Esperanto-Deutsch*. Berlin, Langenscheidt.

Lyons, John (1995<sup>8</sup>): *Einführung in die moderne Linguistik*. München, C. H. Beck.

Manders, Wilhelmus Johannes Arnoldus (1947): *Vijf kunsttalen. Vergelijkend onderzoek naar de waarde van het Volapük, Esperanto, Ido, Occidental en Novial*. Purmerend, Muusses.

Manders, Wilhelmus Johannes Arnoldus (1950): *Interlingvistiko kaj esperantologio*. Purmerend, Muusses.

Meier, Klaus (1999): *Bericht zur Zertifizierung der Legierungsbestandteile (Ca, Sn) und der Verunreinigungen (Al, Ag, Bi, Cu) in einer Blei-Akku-Legierung (PbCaSn) – Zertifiziertes Referenzmaterial BAM-102*. Berlin, Bundesanstalt für Materialforschung und -prüfung, [http://www.bam.de/pdf/service/referenzmaterialien/zertifikate/nichteisen/bam\\_102\\_report.pdf](http://www.bam.de/pdf/service/referenzmaterialien/zertifikate/nichteisen/bam_102_report.pdf) 15.02.2007.

Ministère de l'éducation nationale (1954): *Le français fondamental (1er degré: 1300 mots – 2ème degré: 1000 mots – 2ème degré)*. Paris, IPN.

Narayan, Sundaresan; Balasubrahmanyam, Vriddhachalam K. (1992a): Information theoretic models in statistical linguistics - Part I: A model for word frequencies. In: *Current Science* 63, 261-269.

Narayan, Sundaresan; Balasubrahmanyam, Vriddhachalam K. (1992b): Information theoretic models in statistical linguistics - Part II: Word frequencies and hierarchical structure in language - statistical tests. In: *Current Science* 63, 297-306.

Narayan, Sundaresan; Balasubrahmanyam, Vriddhachalam K. (2000): Information theory and algorithmic complexity: Applications to linguistic discourses and DNA sequences as complex systems. In: *Journal of Quantitative Linguistics* 7, 129-183.

Natrella, Mary G. (1963): *Experimental statistics*. National Bureau of Standards Handbook 91. Washington, D. C., U. S. Government Print Office.

Nüssel, Frank (2000): *The Esperanto Language*. New York, Legas.

Piron, Claude (1986): *Esperanto: Europeese of Aziatische taal?*. Antwerpen, Flandra Esperanto-Ligo.

Rublev, S. (1927): Provo de kvanta element-analizo en la lingvo Esperanto. In: *Sennacieca Revuo*, Nr. 10/12. Paris, Sennacieca Asocio Tutmonda, 155-163.

Sadler, Victor (1959): Relativaj oftecoj de kelkaj lingvaj elementoj en Esperanto. In: *Scienca Revuo* 2-3, 67-72.

Sakaguchi, Alicja (1998): *Interlinguistik. Gegenstand, Ziele, Aufgaben, Methoden*. Frankfurt am Main, Peter Lang.

- Setälä, Vilho (1949): Lingva esprimkapablo kiel funkcio de la vorttrezoro [Sprachliche Ausdrucksfähigkeit als Funktion des Wortschatzes]. In: *Esperantologio* I, 1, 5-32.
- Sigurd, Bengt (1968). Rank-frequency distribution for phonemes. In: *Phonetica* 18, 1-15.
- Sikosek, Ziko Marcus (2003<sup>2</sup>): *Esperanto sen mitoj*. Antwerpen, Flandra Esperanto-Ligo.
- Stamatiadis, Anakreon (1986): Aglutina karaktero de E[speranto]. In: *Enciklopedio de Esperanto*, 14-15.
- Stancliff, Fenton (1933): Literofteco en Esperanto. In: *Heroldo de Esperanto* 49 (754).
- Strauss, Udo; Altmann, Gabriel (2008): *Change in language. From Laws in Quantitative Linguistics*, [http://www-alt.uni-trier.de/uni/fb2/ldv/lqj\\_wiki/index.php/Change\\_in\\_language](http://www-alt.uni-trier.de/uni/fb2/ldv/lqj_wiki/index.php/Change_in_language) 21.09.2008.
- Strauss, Udo; Altmann, Gabriel; Best, Karl-Heinz (2007): *Phoneme frequency. From Laws in Quantitative Linguistics*, [http://www-alt.uni-trier.de/uni/fb2/ldv/lqj\\_wiki/index.php/Phoneme\\_frequency](http://www-alt.uni-trier.de/uni/fb2/ldv/lqj_wiki/index.php/Phoneme_frequency) 02.08.2008.
- Symoens, Edward (1989): *Bibliografio de universitataj kaj altlernejaj diplomverkoj, disertacioj kaj tezoj pri Esperanto kaj interlingvistiko*. Rotterdam, Universala Esperanto Asocio.
- Symoens, Edward (1995): *Bibliografio de Disertacioj pri Esperanto kaj Interlingvistiko. Suplemento*. Rotterdam, Universala Esperanto-Asocio.
- Tiede, Manfred (1995): *Statistische Logit-Analyse: eine Orientierungshilfe für die Verwendung des binären Logit-Modells*. Bochum, Fakultät für Sozialwissenschaft der Ruhr-Universität Bochum.
- Tambovtsev, Yuri; Martindale, Colin (2007): Phoneme Frequencies Follow a Yule Distribution (The Form of the Phonemic Distribution in World Languages). In: *SKASE Journal of Theoretical Linguistics*, Bd. 4, Nr. 2.

Tonkin, Humphrey (Hrsg.) (1997): *Esperanto, Interlinguistics, and Planned Language*. Lanham, University Press of America/Rotterdam-Hartford (Center for Research and Documentation on World Language Problems).

Tonkin, Humphrey; Fettes, Mark (1996): *Esperanto Studies: An Overview*. *Esperanto Document 43a*. Rotterdam, Universala Esperanto-Asocio.

Tuldava, Juhan (1988): Opyt kvantitativnogo analiza sistemy fonem estonskogo jazyka. In: *Acta et Commentationes Universitatis Tartuensis* 838, 120-133.

U. S. Department of Defense (Hrsg.) (1999): MIL-HDBK-1823 *Nondestructive evaluation system reliability assessment*. Washington, D. C., U. S. Department of Defense.

Urban, Dieter (1993): *Logit-Analyse : statistische Verfahren zur Analyse von Modellen mit qualitativen Response-Variablen*. Stuttgart (u. a.), Fischer.

Walther, F. J. (Hrsg.) (1970): Esperanto, eine Sprache höchster Effizienz. In: *Dokumente – Informationen*. Nürnberg, Deutsches Esperanto-Institut.

Waringhien, Gaston (2002): *La Nova Plena Ilustrita Vortaro de Esperanto*. Paris, Sennacieca Asocio Tutmonda.

Waringhien, Gaston (2005): *La Nova Plena Ilustrita Vortaro de Esperanto*. Paris, Sennacieca Asocio Tutmonda.

Wells, John C. (1978): *Lingvistikaj aspektoj de Esperanto*. Rotterdam, Universala Esperanto-Asocio.

Wells, John C. (1989<sup>2</sup>): *Lingvistikaj aspektoj de Esperanto*. Rotterdam, Universala Esperanto-Asocio.

Wendt, Heinz F. (1977): *Sprachen*. Frankfurt am Main, Fischer-Taschenbuch-Verlag.

Wennergren, Bertilo (2006): *Plena Manlibro de Esperanta Gramatiko*. El Cerrito, Esperanto-Ligo por Norda Ameriko.

Wennergren, Bertilo (2007a): Nua Oficiala Aldono al la Universala Vortaro. In: *Akademio de Esperanto – Oficialaj Informoj – N-ro 8*, [http://www.akademio-de-esperanto.org/oficialaj\\_informoj/oficialaj\\_informoj\\_8\\_2007.html?s=h](http://www.akademio-de-esperanto.org/oficialaj_informoj/oficialaj_informoj_8_2007.html?s=h) 10.03.2007.



Wennergren, Bertilo (2007b): *Akademia Vortaro*. [http://www.akademio-de-esperanto.org/akademia\\_vortaro/index.html](http://www.akademio-de-esperanto.org/akademia_vortaro/index.html) 11.10.2007.

Wood, Richard E. (1982): *Current Work in the Linguistics of Esperanto. Esperanto Documents 28 A*. Rotterdam, Universala Esperanto-Asocio.

Yule, George Udny (1924): A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F. R. S. In: *Philosophical Transactions of the Royal Society of London Biological Sciences* 213, 21-87.

Zamenhof, Leyzer Ludwik (1887): *Unua Libro*. Warschau, Ch. Kelter.

Zamenhof, Leyzer Ludwik (1925): *Lingvaj respondoj (plena kolekto)*. Paris, Esperantista Centra Librejo.

Zamenhof, Leyzer Ludwik (1991<sup>10</sup>): *Fundamento de Esperanto*. Pisa, Edistudio.

Zamenhof, Leyzer Ludwik (1992<sup>18</sup>): *Fundamenta Krestomatio de la lingvo Esperanto*. Rotterdam, Universala Esperanto-Asocio.

Zipf, George Kingsley (1929): Relative frequency as a determinant of phonetic change. In: *Harvard Studies in Classical Phonology* 40, 1-95.

Zipf, George Kingsley (1935): *The Psycho-Biology of Language. An Introduction to Dynamic Philology*. Boston: Houghton-Mifflin. (21968) Cambridge: M.I.T. Press.

Zipf, George Kingsley (1949): *Human behavior and the principle of least effort*. Cambridge, Addison-Wesley.

## Stichwortverzeichnis

29/29-Methode.....	86
Affrikaten.....	50
Agglutination.....	55
Ähnlichkeitsanalyse.....	109
Akademio de Esperanto.....	13
Antonymie.....	38
autosemantische Affixe.....	19
Bedeutungserweiterung.....	41
Bedeutungsverengung.....	41
Bekanntheitsgrad.....	84
Berens-Methode.....	87
BRO.....	43
Bulonja Deklaracio.....	13
Devianz-Test.....	124
Dietze-Korpus.....	43
Ersatzschreibweise.....	14
Esperanto.....	13
Farben.....	39
Flexion.....	55
Fundamento de Esperanto.....	13
Gauß-Newton-Verfahren.....	49
Grapheme.....	14
Greenberg-Indizes.....	53
Häufigkeitsmerkmal.....	83
Häufigkeitsmerkmal mit zuverlässiger Bekanntheit.....	108
Hauptthesen.....	11
Homonymie.....	35
Hyperonymie.....	39
Hyponymie.....	39
Interlinguistik.....	33

Internationalismus.....	31
Isolation.....	53
klassifikatorischer Schlüssel.....	53
Kollokationen.....	42
Konnotationen.....	41
konzeptuelles Feld.....	39
Korrelativwörter.....	23
Kruskal-Wallis-Test.....	109
Lexeme.....	30
Lexikon des Esperanto.....	30
Lingva Komitato.....	13
logistische Regression.....	87
maximum likelihood estimation.....	90
Meronymie.....	38
Modell-Tauglichkeit.....	124
modifiziertes Newton-Verfahren.....	98
Morpheme.....	19
Morphemstrukturen.....	59
Morphologie.....	19
Neologismen.....	37
Offizielle Basis-Wortwurzelsammlung.....	43
Oficialaj Aldonoj.....	13
Paronymie.....	36
Phoneme.....	14
Phonemhäufigkeit.....	47
Phrasenstrukturdiagramm.....	27
Phraseologismen.....	45
Piotrowski-Gesetz.....	76
Plansprachen (aposteriorisch, apriorisch).....	33
Polysemie.....	34
polysynthetische Merkmale.....	58
Präfixation.....	56
Prototyp.....	40
Rechts- und Linkseinbettung.....	60

Schätzung der maximalen Mutmaßlichkeit.....	90
Semantik.....	33
Sprachfamilien.....	112
Sprachtypologie.....	19
Suffixation.....	56
Synonymie.....	36
syntaktische Konnexionen.....	53
Syntax.....	27
Synthetismus.....	53
Vertrauensgrad.....	84
Welthilfssprache.....	13
Werkstoffprüftechnik.....	82
Wortbildung.....	21
Wortfeld.....	39
Worthäufigkeit.....	43
X-Bar-Theorie.....	29
Yule-Modell.....	18
Ziel.....	8
Zipf-Modell.....	18
Zusammenfassung.....	126