

# INHALT

Saburo OKAMURA, Willi LANGE, Joachim SCHARLOTH:

Vorwort .....1

Susumu ZAIMA:

Frequenzbasierte Grundwortschätze: Probleme und Perspektiven .....3

Noah BUBENHOFER:

Lehrwerke und Referenzwortschätze

Der Nutzen frequenzbasierter Grundwortschätze .....13

Saburo OKAMURA, Willi LANGE, Joachim SCHARLOTH:

Methoden der Bestimmung des Kernwortschatzes Deutsch .....29

Katsumi IWASAKI:

Ansätze zur Erstellung eines Grundwortschatzes für DaF in Japan .....45

Angelika WERNER, Shiro KATTO:

Wortschatzlernen aus der didaktischen Perspektive .....67

## 目 次

まえがき

..... 岡村三郎, Willi Lange, Joachim Scharloth 1

頻度に基づく「基本語彙」リスト作成の問題点と展望

..... 在間 進 3

教科書と基準語彙 頻度に基づく基本語彙集の価値

..... Noah Bubenhofer 13

ドイツ語の中心的語彙を決定するための方法

..... 岡村三郎, Willi Lange, Joachim Scharloth 29

日本の DaF におけるドイツ語基礎語彙へのアプローチ

..... 岩崎克己 45

語彙の学習 – ドイツ語教授法の観点から –

..... Angelika Werner, 甲藤史郎 67

## VORWORT

Der vorliegende Band enthält die überarbeiteten Versionen der Beiträge, die auf dem Symposium „*Grundwortschatz Deutsch: lexikografische und fremdsprachendidaktische Perspektiven*“ gehalten wurden. Dieses Symposium fand am 16. Oktober 2011 im Rahmen der Herbsttagung der JGG in Kanazawa statt.

Die wissenschaftliche Beschäftigung mit Wortschatzfragen ist aus einer Vielzahl von Gründen wichtig. Hier seien nur zwei genannt: Es gibt erstens eine deutliche Korrelation zwischen dem Wortschatzwissen und den fremdsprachlichen Leistungen von Lernern in den Bereichen Lesen, Schreiben, Sprechen und Grammatikbeherrschung. Zweitens ist der Wortschatzerwerb eine langfristige und herausfordernde Aufgabe. Der Wortschatzerwerb ist inkrementell, sowohl in die Breite als auch in die Tiefe. Wortschatzwissen hat viele Facetten: Das reicht vom phonetischen über grammatisches und semantisches bis hin zum pragmatischen und stilistischen Wissen. Zusätzlich lässt sich unterscheiden zwischen passivem und aktivem Wissen. Wenn man die zugrundeliegenden Strukturen des Zielwortschatzes besser kennt, die zugrundeliegenden Lernprozesse versteht und den Lernweg strukturieren kann, dann ist den Lernenden viel geholfen.

Die Vortragenden des Symposiums gingen gemeinsam davon aus, dass der Sprachunterricht für (fortgeschrittene) Anfänger eine der Hauptaufgaben vieler japanischer Germanistinnen und Germanisten ist. Dabei stellt sich auch immer wieder die Frage, welchen Wortschatz die Studierenden in welcher Reihenfolge mit welchen Methoden erwerben sollen, um möglichst

schnell und möglichst effektiv zu authentischer Interaktion fähig zu werden. Diese Frage stellt sich in Japan umso dringlicher, da die Stunden für den Deutschunterricht im letzten Jahrzehnt oft deutlich reduziert wurden.

Es gab auf dem Symposium fünf Beiträge. Dieser Sammelband enthält die leicht überarbeiteten oder erweiterten Versionen: *Zaima* stellt anhand von konkreten Beispielen die Probleme und Perspektiven frequenzbasierter Grundwortschätze dar und diskutiert ihre Signifikanz im Kontext des japanischen Unterrichts Deutsch als Fremdsprache. *Bubenhof*er setzt eine breitere korpusanalytische Untersuchung zu japanischen Deutschlehrwerken in Beziehung zu verschiedenen Referenzwortschätzen, insbesondere zum DeReWo des Instituts für deutsche Sprache in Mannheim. *Okamura / Lange / Scharloth* stellen ihren datengeleiteten Ansatz zur Bestimmung des Kernwortschatzes Deutsch vor und diskutieren erste Ergebnisse. *Iwasaki* stellt das für japanische Lernende aufgebaute DJPD (Deutsch-Japanisches Pararellkorpus für Deutschlernende) vor und setzt das DJPD in Beziehung zu „Profile deutsch“ und den „Grund- und Aufbauwortschatz“ von Tschirner (2008). *Werner / Katto* betrachten den Wortschatzerwerb aus didaktischer Perspektive und stellen ein neues Format von Wortschatzlernkarten vor.

Wir hoffen, dass die vorliegenden Beiträge vielfältige Diskussionen um die Bedeutung des Wortschatzes bzw. des Grundwortschatzes für den Deutschunterricht in Japan auslösen.

im Mai 2012

Die Herausgeber

# 頻度に基づく「基本語彙」リスト作成の問題点と展望<sup>1</sup>

## 在間 進

### 1 はじめに

語学教育の体系化を目指す CEFR の提唱, 大規模コーパスを活用した使用頻度調査などに触発されて「基本語彙」リスト作成の研究が盛んになってきている。「基本語彙」リスト作成の目的は, その成果の語学学習, 教材作成などへの応用である。より正確に言うならば, 語学学習の効率化である。

本稿の目的は, 「語彙学習」という観点のみならず, 「初級文法」学習という観点からも「基本語彙」リスト作成が喫緊の課題であるとの認識のもとで, 頻度に基づく「基本語彙」リスト作成に関する一般的な問題点と展望を述べることである。

### 2 「基本語彙」リスト作成の問題点

まず, 「基本語彙」リスト作成における問題点を挙げる。

#### 2.1 単位としての「語彙項目」の認定

その一つは, 「語彙項目」の認定である。

##### 2.1.1 派生語

「基本語彙」の「語彙項目」は, 語形と意味の組み合わせである。語形と意味の組み合わせという単位の認定において, まず問題になるのは, 派生形の取り扱いである。語形が部分的に変化する派生形の場合, 元の中心的な意味が保持されるが, 周辺的な意味が変化する。したがって, 派生のどの場合までを同一「語彙項目」のバリエーションとみなし, どの場合から別の独立した「語彙項目」とみなすかということが問題になる。

たとえば *gelöst* という語形には, (1) のような, 動詞 *lösen* の完了形および受動形を作る用法があるが,

- (1) a. Er hat ein Problem gelöst.
- b. Das Problem ist gelöst worden.

---

<sup>1</sup> 本研究は科研費(21520430)の助成を受けたものである。

他方、(2) a のような付加語的用法もある。そして、それには、b のような、否定の接頭辞の付いた形容詞が反義語として存在する。

- (2) a. ein gelöstes Problem  
b. ein ungelöstes Problem

(1) の用例は、動詞 *lösen* の派生形として捉えることに問題はないであろうが、(2) a の取り扱いは必ずしも自明ではない。(1) との関係を重視するならば、*lösen* の派生的用法と捉えることになるが（この場合、単独の語彙項目にならない）、他方、(2) b との関連性を重視するならば、自立的な形容詞と捉えることになる（この場合、単独の語彙項目になる）。

しかし、この問題は、何が「正しい」のか、ア priori に論じうるものではなく、有用性という実用的な観点、実践的な観点から取り扱うべきものであろう。

### 2.1.2 多義語

語形と意味の組み合わせという単位の認定において、次に問題になるのは、「多義性」の取り扱いである。「基本語彙」の「語彙項目」は、語形と「意味」の組み合わせというよりは、語形と「語義」の組み合わせである。多義語の場合、一つの語形に対して複数の語義が存在する。

語形が同一でも語義が異なれば、語義が重視され、それぞれの組み合わせが独立した別々の語彙項目になる。

また多くの場合、想定しうる語義の間には、学習上の重要度に相違がある。すなわち学習上重要な語義もあれば、それほど重要でない語義もある。そして、一般的に、重要な語義のみが「基本語彙」の対象になる。

したがって、多義的な語彙の場合、それを多義とするのかしないのか（すなわち別々の語彙項目とするのかどうか）、また多義とする場合、どのような語義に区分するのか（すなわちどのような語彙項目をたてるのか）を決めなければならない。さらに言えば、「基本語彙」の対象にする語義は、それらのすべてなのかあるいは一部なのかを決めなければならない。

しかし、この問題は、何が「正しい」のか、ア priori に論じうるものではなく、有用性という実用的な観点、実践的な観点から取り扱うべきものであろう。

なお、語義を確定しえたとしても、当該の語義区分が情報としてコーパスに組み込まれていない限り、プログラムを用いた、語義を軸とした大規模かつ本格的な使用頻度調査は不可能である。ではこのような、語義区分の情報も組み込んだコーパスがいつ構築されうるのかと問われるならば、そもそもそのようなコーパスが構築されうるのかと問わざるをえない。現実的な、唯一可能な「解決策」は、実用的な観点からの「対症療法」として、問題になる語彙項目の各

語義の使用割合をサンプル的に算定し、それに基づいて使用数を推定することであろう。

## 2.2 「基本語彙」の選定

「基本語彙」リスト作成における問題点のもう一つは、「基本語彙」の選定である。

「基本語彙」を選定するためには、以下の二つの作業が必要となる。

(a) 語彙を序列化する

(b) 一定数の語彙を「基本語彙」として選定する

学習上の重要性に基づく語彙の序列化は、従来、教師の経験などに基づいていたが、最近では、コーパスを用いた使用頻度調査の結果に基づく試みが盛んになっている。

すでに何種類もの、頻度に基づいた語彙リストが入手可能であるが、個々の語彙の序列（すなわち重要度）は、リスト間で異なっており、したがって、一定数で区切った場合、含まれる語彙はリストごとに異なる。

たとえば Institut für deutsche Sprache (Mannheim)<sup>2</sup> の作成した使用語彙頻度リスト derewo (2007-v-30000g-2007-12-31-0.1) の上位 4000 語（母体コーパスは Deutsches Referenzkorpus DeReKo 約 11 億語）と Jones, R./Tschirner, E. (2006)<sup>3</sup> に掲載された 4000 語リスト（以下、Herder/BYU と呼ぶ）を比較した場合、多くの語彙の序列は異なり、また、一方に含まれているが他方に含まれていない語彙が約 25% 前後ある。

「基本語彙」リストにおける語彙項目の序列およびそれがリストに含まれるか否かは、どのようなコーパスとどのような序列基準を用いるかによって左右される。したがって、より代表性のあるコーパスの構築や客観的かつ検証可能な基準を求めようとするのは極めて自然な発想であろう。

しかし、ドイツ語を真の意味で代表するコーパスの構築は、実際上不可能である。たとえどのようなコーパスを構築したとしても、それがドイツ語を代表するものであることを証明することはできないからである。また、どのように序列基準を工夫したとしても、「基本」語彙というものがアプリアリに存在しているわけでないのであるから、「基本」語彙と「非基本」語彙の区別は、どこまでも人為的な作業となる。

---

<sup>2</sup> <http://www.ids-mannheim.de>.

<sup>3</sup> Jones, R./Tschirner, E. (2006): A Frequency Dictionary of German, Core vocabulary for learners. London: Routledge.

### 3 語学学習上の有用性

以上のような問題点があるのにも拘わらず、「基本語彙」リスト作成を試みるのは、語彙に学習上の重要度の差が認められ、かつ人間の学習能力に限界があるため、学習上の順序を決めることが語学学習上有用であるからである。

「基本語彙」を語学学習上の有用性という観点から研究する場合、考えるべき重要な点が二つある。

#### 3.1 目的の明確化と評価の導入

その一つは、「基本語彙」リスト作成の目的の明確化と評価の導入である。

##### 3.1.1 意義

有用性を論じる際、まず、目的が明確に定められていなければならない。目的が明確になって始めて、有用かどうか、どの程度有用かを知ることができ、それに基づく改善が可能になる。

なお、冒頭で、「基本語彙」リスト作成の目的は、語学学習の効率化であると述べたが、目的という視点を明確に研究の枠組みに取り入れることによって、「語学学習」の「何」（たとえば読解力、会話力あるいは文法学習などの「どれ」）を効率化するののかということがより明確に意識化され、分析が精緻化されるであろう。

また、評価という作業を「基本語彙」リスト作成の一過程として導入し、データ収集、選定基準などの検証を行うことは、「基本語彙」リスト作成が一つの本格的な「研究」になりうるための不可欠な作業の一つである。

##### 3.1.2 私の目的

上で derewo と Herder/BYU の 2 つの語彙頻度リストについて触れた。前者はドイツ語母語話者のドイツ語使用のデータであり、また、後者は外国語学習の視点を取り入れてはいるが（副題が *core vocabulary for learners*）、日本人のドイツ語学習に特化したものではない。

現在、私は、「基本語彙」リスト作成の目的として、(3) のような 2 点を想定している。

- (3) a. ドイツ人によってよく使用されるドイツ語語彙学習の効率化
- b. 日本における ドイツ語初級文法学習の効率化

前者は、ドイツ語学習上級者を想定したもので、一般的な課題と言えるが、後者は、必ずしも一般的なものではないため、具体的に少し説明する。



たとえば、複数形の無語尾式にウムラウトするものがある。初級文法教材では、具体的な事例として Vater, Garten, Apfel などが候補となるが、Herder/BYU における順位およびネット上のヒット数などでは、以下のような順位になる。

(4) Herder/BYU での順位

Vater : 216 > Garten : 959 > Apfel : 3837

(5) yahoo.de (Seiten auf Deutsch) のヒット数 (2012年1月13日現在)

【単数形】 Garten : 15.600.000 > Vater : 8.980.000 > Apfel : 2.550.000

【複数形】 Gärten : 2.140.000 > Äpfel : 1.400.000 > Väter : 1.380.000

しかし、教師としての主観的判断になるが、特に複数形の練習問題 (文例) を考えると (たとえば Wie viele ...?), 無語尾式ウムラウトの事例として Apfel/Äpfel が最も適している。

また、受動文を教える際の具体的な動詞として loben が多く使用される。しかし、動詞が受動形として用いられる比率を調査した須田佳寿子 (2008)<sup>4</sup>によれば、loben の、受動形として用いられる比率は極めて低く、受動形として用いられる比率が最も高い動詞は wählen である。両者の受動形として用いられる比率は (6) のようになる。

(6) 受動形での使用比率

loben 9% (すなわち 100 例中 9 例が受動形)

wählen 59% (すなわち 120 例中 71 例が受動形)

また、ネット上のヒット数 (gelobt wurde と gewählt wurde という結合形式で調査) も、loben の受動形の方が圧倒的に少ない。

(7) yahoo.de (Seiten auf Deutsch) のヒット数 (2012年1月13日現在)

gelobt wurde 20.900

gewählt wurde 172.000

しかし、教師としての経験的判断になるが、受動形の人称変化および能動文の主語の、von 前置詞句への転換のことを考えると、受動文を教える際、loben の使用が最も適している。

---

<sup>4</sup> 須田佳寿子 (2008) : 『ドイツ語動作受動文の、コーパスに基づく使用頻度分析』, 東京外国語大学大学院地域文化研究科修士論文。

要するに、ドイツ語母語話者の中で使用頻度の高い語彙がドイツ語の初級文法学習に適しているとは必ずしも言えないわけである。ではどのようにして教育的観点を組み込んだ「基本語彙」を作成するのかということになる。しかし、現在、残念ながら、全体的な構想とその成果を語れる段階に未だない。そのため基本作業として、— 英語教育での成果を取り入れながら —、ドイツおよび日本で作成されている教材などの使用語彙、使用用例を収集し、分析を行っているところである。

### 3.2 有用性の評価方法

「基本語彙」を語学学習上の有用性という観点から研究する場合、考えるべき重要な点のうち1つは、有用性の評価方法である。

#### 3.2.1 「差分」調査<sup>5</sup>

読解対策としての「基本語彙」の場合、有用性を検証する一つの方法として、「差分」調査が考えられる。これは、テキストの読解における有用性、すなわちテキストでの使用語彙がどの程度当該の語彙リストでカバーされているのか（すなわちテキストでの使用語彙の何%が当該リスト内にあるのか）を調査するものである。

一つのサンプル例として、2011年度のドイツ語技能検定試験（略称：独検）3級の試験問題（秋期）の使用語彙（計306語<sup>6</sup>）を調査対象とし、上述の *derewo* と Herder/BYU の語彙リストの上位1000語、2000語、3000語、4000語のカバー率を調査した。なお、カバー率は、一般的に、テキスト全体の延べ語数に対するある語または語の集団の比率として用いられるが、ここでは、テキストに使用されている語（いわゆる変化形を一つにまとめた「レマ（見出し語）」）の総数を分母とし、それぞれの語彙リストによってカバーされている語の総数を分子とする比率を指す。<sup>7</sup>その結果、それぞれ語彙リストのカバー率は、表1のようになった。なお、（ ）内の数字は、カバーされている語数である。

	1000 語	2000 語	3000 語	4000 語
Herder/BYU	約 62% (189)	約 73% (223)	約 75% (231)	約 80% (245)
<i>derewo</i>	約 56% (172)	約 66% (203)	約 71% (217)	約 77% (235)

表1

<sup>5</sup> この調査には、山田善久氏（岐阜経済大学）作成のソフトウェア Teddy-Z を使用。

<sup>6</sup> 固有名詞、前置詞と定冠詞の融合形などは除いた数字である。

<sup>7</sup> この点については、本稿末尾の【追記】も参照。

この表に限って言えば、両者の差は、1000語で約6%、2000語で最大になり約7%、その後4%、2%と差が縮まっている。<sup>8</sup>前者が外国語学習者を対象にした結果であるとするならば、このことは、この種のテキストの読解（すなわち日本人受験者を想定した、日本人作成のテキストの読解）に対して、一たとえば語彙学習上最も効率的と考えられる2000語、3000語レベルで<sup>9</sup>—、有用性のより高い「基本語彙」リストの作成が、「日本におけるドイツ語学習」という観点をより強く押し出すことによって可能であることを示唆していると言えよう。

### 3.2.2 「基本語彙」の評価方法

3.1.2. で、「基本語彙」リスト作成の目的として、「ドイツ人によってよく使用されるドイツ語語彙学習の効率化」と「ドイツ語初級文法学習の効率化」を想定していることを述べた。前者の評価に関しては、上で、述べたような「差分」調査などが考えられるが、後者の評価に関しては、残念ながら、具体的な方法が未だない。現在考えられる方法は、作成した「基本語彙」リストを文法学習、語彙学習などの実際の現場で使用しながら、それらの有用性を評価するという「対症療法」的なものであろう。しかし、その際も、単に経験によるのではなく、明確な方法論を確立する必要がある。今後の大きな課題としたい。

## 4 用例の選定

「基本語彙」リストを作成する上で残された問題は、まだまだ多くある。その中で一番重要な課題は、「基本語彙」の用例の選定である。「基本語彙」の用例は、当然のことながら、実際によく使用されるものであることが学習効率の上からも望ましい。以下、使用頻度の高い用例、すなわち当該の語彙項目を主要構成素とする、使用頻度の高い語句結合を調査する二つの方法について試行した結果を述べる。

### 4.1. n-グラム分析

使用頻度の高い語句結合を調査する一つの方法はn-グラム分析の応用である。n-グラム分析とは、検索語と共起語が隣接するパターンを分析し、次にくる単語を予測する計算モデルのことである。

日本語の語連結のパターンに関して Google が 255 億語からなる言語データを元に以下の調査結果を示している。

<sup>8</sup> 500語の場合のカバー率は約50%と約47%で、1000語の場合より差が縮まっている。

<sup>9</sup> この調査では、2000語以降それぞれのリストでカバー率が急激に下がっている。このことは、語彙学習上、2000語までが最も効率的であることを示していると言えよう。

(8) Google 日本語単語 n-グラムデータ一例<sup>10</sup>

1-gram	2-gram	3-gram	4-gram	5-gram	6-gram	7-gram	頻度
グーグル	で	検索	し	て	おく	べき	27
グーグル	で	検索	し	て	い	た	39
グーグル	で	検索	し	て	い	て	45
グーグル	で	検索	し	て	い	たら	143
グーグル	で	検索	し	て	いる	と	433

ドイツ語の場合、現在、この種の本格的な調査は未だ実施されていない。<sup>11</sup>したがって、この着想に準じる形で、ドイツ語における「語連結のパターン」を抽出するため、試行的に、*dass* 文の事例を収集し分析を行った。*dass* 文を分析対象としたのは、ドイツ語の場合、「副文では、動詞が文末に置かれ、他の語句は、動詞との関係が強いものほど文末に置かれる」および「副文では、副詞などの添加成分の共起が比較的少ない」との作業仮説に基づく。特に「語連結のパターン」が有意義な課題になる動詞の場合、*dass* 文が「語連結のパターン」の抽出に最も適していると考えた。

分析対象として、動詞 *verbringen* を選び、Institut für deutsche Sprache のコーパス DeReKo から *&verbringen*<sup>12</sup> の現れる *dass* 文を Cosmas II によって収集し（事例数約 12 万）、*verbringen* がどの語とどの頻度で結合して現れるかを調査したところ、結果は、以下の通りである。<sup>13</sup>

(9) *&verbringen* の結合事例

	5 語	4 語	3 語	2 語	1 語		事例数
a.	_____	_____	_____	<i>zu</i>	<i>Hause</i>	<i>&amp;verbringen</i>	621
b.	_____	_____	<i>Jahre</i>	<i>seines</i>	<i>Lebens</i>	<i>&amp;verbringen</i>	263
c.	_____	<i>seines</i>	<i>Lebens</i>	<i>hinter</i>	<i>Gittern</i>	<i>&amp;verbringen</i>	88
d.	<i>sie</i>	<i>den</i>	<i>Rest</i>	<i>ihres</i>	<i>Lebens</i>	<i>&amp;verbringen</i>	69

この分析の最大の問題は、上記のように、語数が増えるごとに事例数が急激に下がることである。有意義と認めうるような数値の事例を得るためには、少なくとも数百億語を越える規模のコーパスが必要であろうが、現時点では、そのようなドイツ語コーパスは存在しない。また、そのような規模のコーパスが構

<sup>10</sup> <http://googlejapan.blogspot.com/27/11/n-gram.html>.

<sup>11</sup> 2009 年に Google のブック検索で収集されたコーパスの N-Gram が Google から無償配布されているが (<http://books.google.com/ngrams/datasets>)、ドイツ語に関してまだまだ利用可能の状況ではない。

<sup>12</sup> *&verbringen* の *&* は、現在形、過去形などの変化形をすべて含むことを表す。

<sup>13</sup> Zaima, S./et al. (2011): *Korpusbasierte Analyse der syntakto-semantischen Konstituentenverbindungen des Verbs verbringen*, *Japanische Gesellschaft für Germanistik (Hrsg.): Mapping zwischen Syntax, Prosodie und Informationsstruktur*. Iudicium. 103-120.

築されたとしても、それによって語学学習上有意義な成果が得られるという保証も未だない。

#### 4.2 構成素の結合頻度分析

もう一つの方法は、結合する構成素の頻度分析を行うことである。上述の動詞 *verbringen* は、必須成分として 2 構成素を取る事例が一番多いが、それらの組み合わせパターンの割合は、(10) (11) (12) のようになる (分析全事例数 1968 ; Zaima/et al. (2010))。

(10) a. 「何かをどこかで」 (約 58% ; 1151 事例)

dass Dreiviertel der Bewohner das Wochenende zu Hause verbringen

b. 「何かを何かの仕方で」 (約 29% ; 569 事例)

dass sie ihr Leben gemeinsam verbringen wollten.

(11) a. 「どこかで何かを」 (約 7% ; 142 事例 )

dass gut jeder dritte Deutsche hier seinen Urlaub verbringt

b. 「何かの仕方で何かを」 (約 5% ; 104 事例)

dass sie mit Mördern den Abend verbringen

(12) 「何かの仕方でどこかで」 (約 0.1% ; 2 事例)

dass sie meist jahrzehntelang unschuldig in der Todeszelle verbracht hatten

動詞 *verbringen* の場合、「何かをどこかで過ごす」という結び付きが最も多く使用されているのであるから、少なくともこのパターンの事例を「基本語彙」の用例として載せるということが学習効率上最も望ましいことになる。この方法は、構成素分析という、機械では不可能な作業が入るため、相当の労力が必要であるが、不可能ではない。

以上、「基本語彙」の用例として望ましいものを調査する方法を 2 つ示した。「基本語彙」リスト作成という目標は、極めて実用的なことであるが、有用性のより高いものにするためには、ドイツ語そのものの本格的な研究が不可欠である。

## 5 終わりに (私の想い)

今回の「基本語彙」をめぐる考察から得られる知見は、3.2.1 で取り扱った差分調査からも推測できるように、言語運用というものが基本的に数千の語彙によって担われているということである。他方、これまでのドイツ語研究の経験によれば、文法現象のほとんどすべてが基本的な数千の語彙の中に網羅されていると言える。

重要な文法現象のほとんどすべてが数千の語彙に網羅されており、かつこれらの数千の語彙が言語運用の主要部分を占めているとするならば、これら数千の語彙の分析は、実質上、ドイツ語全体の分析と等しい価値を持つと言っても過言ではない。

数十万ならいざ知らず、数千の語彙の分析は、大規模なコーパスの構築とともに、様々な検索ソフトが開発されつつある現在、決して不可能な目標ではない。もしそうであるならば、上記のような作業仮説（「ドイツ語全体の分析」＝「ドイツ語の数千の語彙の分析」）のもと、「ドイツ語全体を研究するんだ」という気概を持ち、数千の語彙の分析を本格的に構想することがあってもよいのではないであろうか？むしろそうあるべきではないであろうか？

### 【追記】

3.2.1 でテキストの「カバー率」を取り上げた。「カバー率」(text coverage) は、一般に、ある語または語の集合がテキスト全体の延べ語数に対して占める割合を示す。たとえば、「頻度上位の 2000 語の、日常語におけるカバー率は約 90% である」などと主張されるのは、この種の研究の成果である。しかし、その際、このような、頻度上位 2000 語が具体的にどの語なのかが示されることは（あまり）なかった。したがって、「カバー率」についていくら語られようとも、それは、語彙学習者にとって、どの語を学べばよいのかが分からないのであるから、何の意味もないと言っても過言ではない。そこで、近年、具体的な「基本語彙」リストの作成が盛んに計画され、提唱されるようになった。

ドイツ語テキストの内容把握の指標を求めるという観点から見た場合、「カバー率」の調査方法には少なくとも二つの可能性があると考えられる。その一つは、上述のように、テキスト全体の延べ語数を対象とするものである。もう一つは、当該テキストで使用されている語の総数（いわゆる変化形を一つにまとめた「レマ」の総数）を対象とするものである。

テキストの読解能力には、語彙数以外の、他の多くの要因も絡むのであり、読解能力とカバー率（学習語彙数）の因果関係はそれほど簡単でない（『英語の語彙知識に基づくリーディングの Can-do List の構築』（相澤一美著、科学研究費補助金研究成果報告書））。

したがって、語数という観点から読解に効果的な「基本語彙」リストのあり方を考察する場合、まずは、基礎作業として、テキストで使用されている「レマ」の総数を対象とする調査が行われるべきであると考えられる。特に、語学試験では多く語彙力そのものも出題対象になりうるのであるから、「レマ」の総数を対象にする調査の方がより実践的と言えよう。3.2.1 での「差分」調査において、私が「レマ」の総数を対象にしたのは、以上のような理由による。

# LEHRWERKE UND REFERENZWORTSCHÄTZE

## Der Nutzen frequenzbasierter Grundwortschätze

Noah BUBENHOFER

### 1 Fragestellung

Bei der Erstellung von Lehrwerken für den Fremdsprachenunterricht stellt sich immer die Frage nach der Auswahl des Vokabulars. Fürs Deutsche existiert eine Art De-facto-Standard, nämlich „Profile deutsch“<sup>1</sup>, das für die Stufen A1 bis B2 den kanonischen Wortschatz definiert.

Der vorliegende Aufsatz fragt nach dem Nutzen von Frequenzwortschätzen: Spiegeln die japanischen Deutsch-Lehrwerke die häufigsten Wörter wieder? Die Frage zielt dabei in zwei Richtungen:

Welchen Häufigkeitsklassen entspricht das tatsächlich in den Lehrwerken eingeführte Vokabular?

Welche Anteile des häufigsten Vokabulars der Sprache finden sich in den Lehrbüchern wieder?

Implizit wird damit auch untersucht, wie der Referenzwortschatz von „Profile deutsch“, der wohl bis zu einem gewissen Grad der Hintergrund für die Wortschatzauswahl in den Lehrwerken bildet, mit einem rein frequenzbasiert bestimmten Wortschatz übereinstimmt. Es geht bei diesem Vergleich nicht um eine Kritik der Lehrwerke, sondern darum, zu prüfen, ob eine frequenzbasierte Bestimmung des Wortschatzes mit anderen Prinzipien, welche implizit bei der Erstellung der Lehrbücher zur Anwendung kamen, konvergiert.

---

<sup>1</sup> Glaboniat, Manuela (2005): Profile deutsch: Gemeinsamer europäischer Referenzrahmen; Lernzielbestimmungen, Kannbeschreibungen, Kommunikative Mittel; Niveau A1-A2, B1-B2, C1-C2, Berlin u.a.: Langenscheidt.

## 2 Vorgehen

Die Basis für die vorliegende Untersuchung bilden fünf japanische Lehrwerke, die digitalisiert vorliegen. Die Erarbeitung dieser Daten gehen auf ein Vorgängerprojekt zurück, bei dem die Bücher gescannt und mit einer OCR-Software in digitalen Text konvertiert wurden<sup>2</sup>. Es handelt sich um folgende Bücher:

- Ogino, Kurahei/Raab, Andrea (2009): Ein Sommer in Deutschland. 4. Auflage, Tokyo: Asahi. (Im Folgenden: „Sommer“.)
- Itayama, Mayumi/Shioji, Ursula/Motokawa, Yuko/Yoshimitsu, Takako (2007): Farbkasten Deutsch neu 1. 26. Auflage, Tokyo: Sanshusha. (Im Folgenden: „Farbkasten“.)
- Sekiguchi, Ichiro (2008): Hallo München. Neu. Tokyo: Hakusuisha. (Im Folgenden: „München“.)
- Seino, Tomoaki (2008): Meine Deutschstunde. 4. Auflage, Tokyo: Asahi. (Im Folgenden: „Deutschstunde“.)
- Riessland, Andreas/Waragai, Ikumi/Kimura, Goro Christoph/Hirataka, Fumiya/Raindl, Marko/Ohta, Tatsuya (2009): Modelle neu 1. 6. Auflage, Tokyo: Sanshusha. (Im Folgenden: „Modelle“.)

Die Texte wurden mit Metadaten versehen, japanische Textteile ausgesondert und die Lektionen und weitere Untergliederungen (Lesetext, Übung, Grammatik, Wortschatz etc.) markiert. Zudem wurden die Texte unter Verwendung des TreeTaggers<sup>3</sup>maschinell mit Wortarten und Lemmata annotiert. Dabei sollte man sich dessen bewusst sein, dass an mehreren Stellen der Korpusaufbereitung Fehler entstehen können: Bereits beim OCR-Prozess können Erkennungsfehler entstehen sowie später bei der Annotation und der maschinellen Wortartenerkennung und Lemmatisierung.

---

<sup>2</sup> Bubenhofer, Noah/u.a.: Welcher Wortschatz? Korpuslinguistische Untersuchungen zur Wortschatzselektion japanischer Deutschlehrbücher für Anfänger. In: *Doitsugo Kyoiku – Deutschunterricht in Japan* 16, S. 43–60.

<sup>3</sup> Schmid, Helmut (1994): Probabilistic Part-of-Speech Tagging Using Decision Trees.



Ein frequenzbasierter Wortschatz muss sich an der Analyse großer Textmengen, also Korpora, orientieren. Für das Deutsche liegt es nahe, mit dem Deutschen Referenzkorpus DeReKo zu arbeiten, das momentan mit etwa 5 Mia. Textwörtern das größte öffentlich verfügbare linguistische Korpus der deutschen Sprache ist<sup>4</sup>. Die Erstellung einer nach Frequenz geordneten Liste der Wörter daraus ist nicht trivial, da sich Fragen stellen wie danach, was überhaupt als „Wort“ gezählt, wie trennbare Verben, Eigen- und Gattungsnamen behandelt oder welches Frequenzmaß verwendet werden soll. Mit „DeReWo“<sup>5</sup> steht jedoch eine solche Frequenzliste auf der Basis des DeReKo zur Verfügung.<sup>6</sup> Die Liste wird in zwei Versionen angeboten: der Wortformen- und der Grundformenliste. Für die vorliegende Untersuchung wurde letztere verwendet.<sup>7</sup> Sie basiert auf einer DeReKo-Version, die etwa 3 Mia. Textwörter umfasst und klassifiziert 40.000 Grundformen nach Häufigkeitsklassen.<sup>8</sup> Die Häufigkeitsklasse (HK) setzt

---

<sup>4</sup> Institut für Deutsche Sprache: Das Deutsche Referenzkorpus DeReKo.

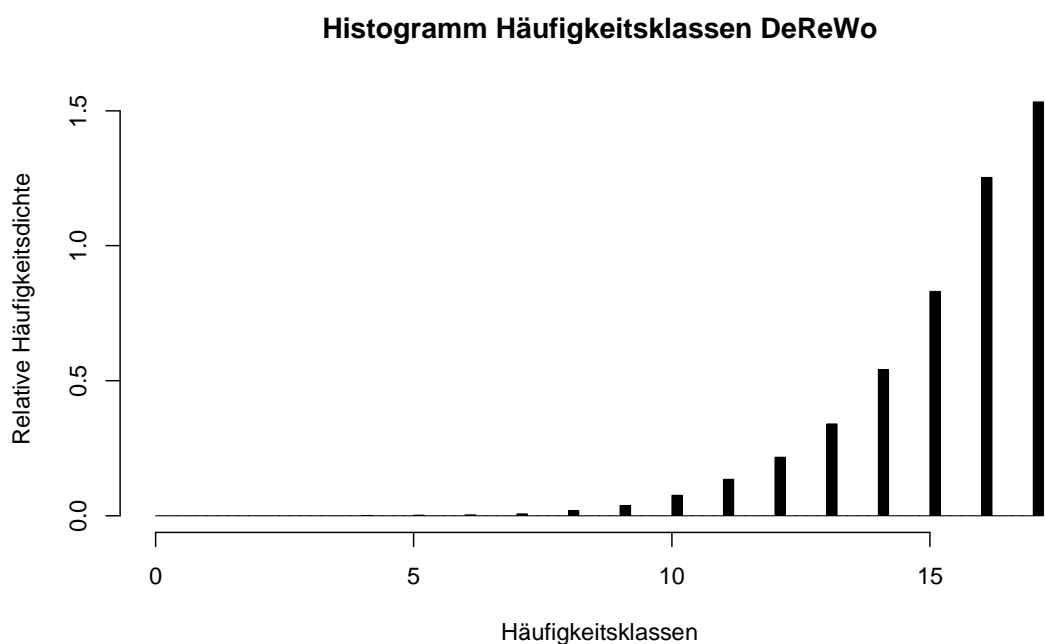
<sup>5</sup> Programmbereich Korpuslinguistik: „Korpusbasierte Grundformenliste DeReWo, v-40000g-2009-12-31-0.1, mit Benutzerdokumentation“, <http://www.ids-mannheim.de/kl/derewo/> (08. Mai 2012).

<sup>6</sup> Es gibt neben der Liste aus dem Jahr 2009 eine weitere aus dem Jahr 2011 (Programmbereich Korpuslinguistik: „Korpusbasierte Wortgrundformenliste DeReWo, v-ww-bll-250000g-2011-12-31-0.1, mit Benutzerdokumentation“, <http://www.ids-mannheim.de/kl/derewo/>.) die umfangreicher ist, sich jedoch gemäß Angaben in der Benutzerdokumentation als Beta-Version versteht. Für die vorliegende Untersuchung ist die Wortmenge sekundär, da sowieso nur die häufigsten Wörter berücksichtigt werden, weshalb sich die Verwendung der älteren Liste von 2009 anbietet.

<sup>7</sup> Vgl. die Benutzerdokumentation Programmbereich Korpuslinguistik: „Korpusbasierte Grundformenliste DeReWo, v-40000g-2009-12-31-0.1, mit Benutzerdokumentation“. zu den Details der Aufbereitung.

<sup>8</sup> Problematisch ist der Umgang mit Eigen- und Gattungsnamen: Eigennamen wurden aus den Frequenzlisten mit einem halbautomatischen Verfahren entfernt und die Frequenzen von Wortformen, welche sowohl Eigen- als auch Gattungsnamen sein können, angepasst, indem der Anteil der Eigennamen-Verwendung herausgerechnet wurde (vgl. „Korpusbasierte Grundformenliste DeReWo, v-40000g-2009-12-31-0.1, mit Benutzerdokumentation“, <http://www.ids-mannheim.de/kl/derewo/>). Allerdings sind Unstimmigkeiten in der Liste zu beobachten: So wurden gemäß Dokumentation „über 400 Bezeichner für die Bewohner von Städten und Regionen“ (Benutzerdokumentation: 9) herausgefiltert, wobei das nur für die maskulinen Formen zu stimmen scheint, da beispielsweise *Mannheimerin* enthalten ist.

die Frequenz des Wortes in Bezug zum häufigsten Wort im Korpus. Ein Wort  $x$  mit HK  $n$  bedeutet: Das häufigste Wort kommt  $2^n$  mal häufiger vor als das Wort  $x$ . Bezogen auf die Grundformenliste von DeReWo bedeutet das: Der bestimmte Artikel *der/die/das* hat HK 0, die Grundform *Tag* weist HK 7 auf; der Artikel *der/die/das* kommt demnach  $2^7$  (=128) mal häufiger vor als *Tag*.



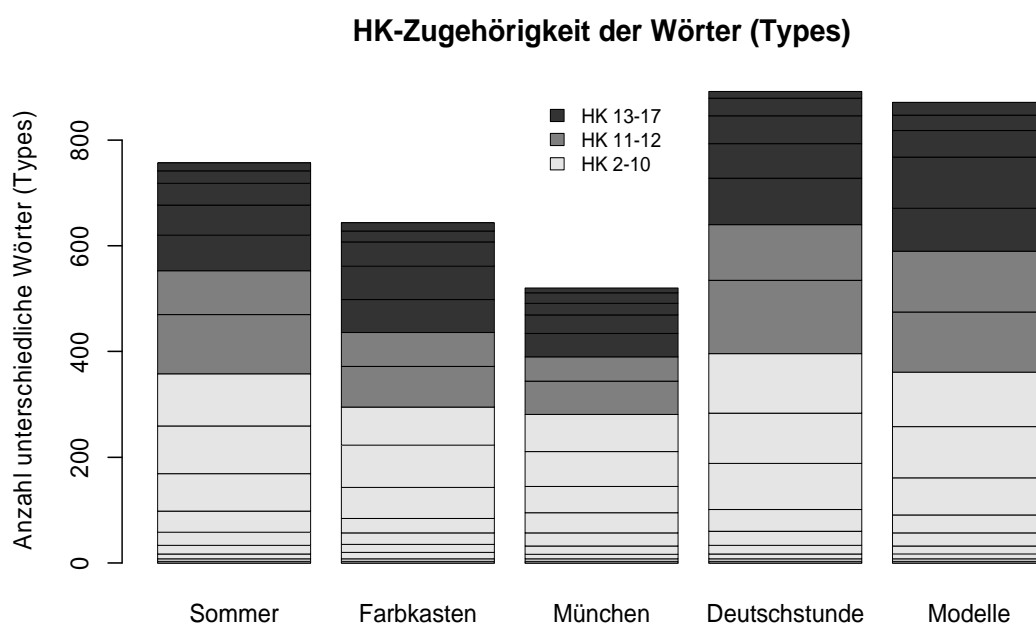
*Abb.1: Histogramm der Häufigkeitsklassen der DeReWo-Liste*

Wie Abbildung 1 zeigt, gehören nur wenige Wörter den niedrigen Häufigkeitsklassen an (sind also sehr häufig) und es gibt sehr viele seltene Wörter – das Histogramm visualisiert also die Krux des Fremdsprachenlernens: Von den wirklich häufigen Wörtern gibt es gar nicht so viele, von den seltenen hingegen schon. Ein Blick in die DeReWo-Liste zeigt diesen Sachverhalt noch dramatischer: Mit dem Wort an Stelle 100 der häufigsten Wörter (*Mensch*) bewegen wir uns schon im Bereich der Häufigkeitsklasse 7. Häufigkeitsklasse 9 reicht bis zum 589. Wort der Liste (*Theater*) und um alles bis und mit Häufigkeitsklasse 10 abzudecken, müssen 1196 Wörter,

bis Häufigkeitsklasse 12 bereits 4018 Wörter gelernt werden. Diese Liste ist nun die Grundlage, um die in den untersuchten Lehrwerken verwendeten Wörter zu analysieren.

### 3 Abdeckung der Häufigkeitsklassen in den Lehrbüchern

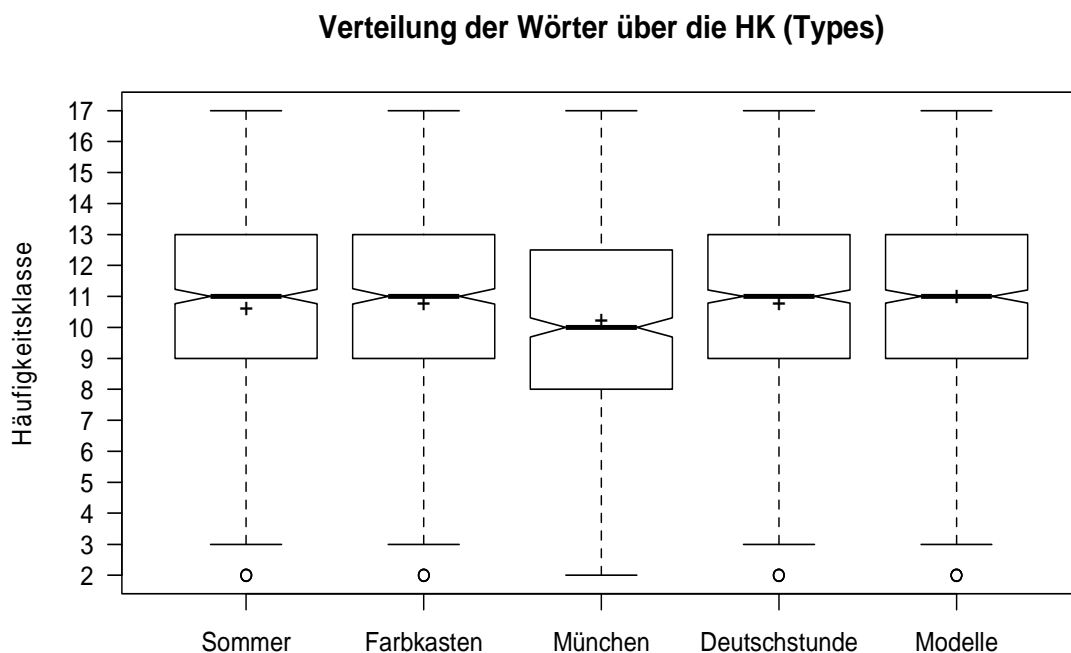
Als erstes interessiert die Frage, welchen Häufigkeitsklassen die in den Lehrbüchern vorkommenden Wörter entsprechen. Abbildung 2 zeigt pro



*Abb.2: Welchen Häufigkeitsklassen gehören die Wörter in den Lehrbüchern an?*

Lehrbuch die Anzahl der unterschiedlichen Wörter (Types) in den einzelnen Häufigkeitsklassen. Die Grafik zeigt einerseits, dass die Anzahl der Wörter der Grundformenliste je nach Lehrbuch sehr unterschiedlich ist: „Deutschstunde“ und „Modelle“ enthalten mit je knapp 900 unterschiedlichen Wörtern den größten Umfang, „München“ mit 520 den geringsten. An den Grauschattierungen sind die unterschiedlichen Bereiche der Häufig-

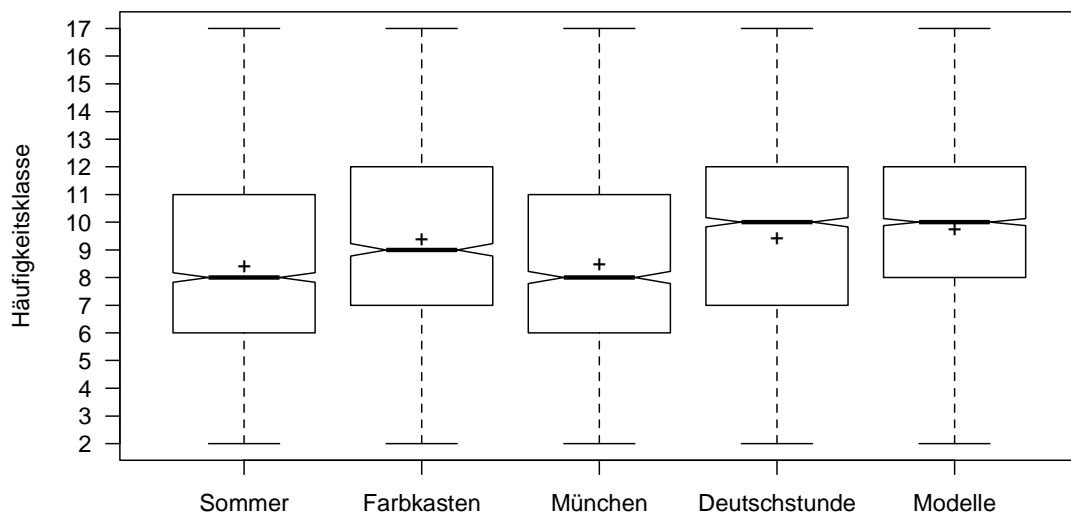
keitsklassen ablesbar: Der hellste Bereich gibt die Klassen 2-10 an.<sup>9</sup> „Deutschstunde“ weist mit 396 die höchste Zahl an Wörtern auf, die dem Bereich HK 2-10 zugeordnet werden können, dicht gefolgt von „Modelle“ (361) und „Sommer“ (358). „Farbkasten“ (295) und „München“ (281) enthalten am wenigsten Wörter dieser Häufigkeitsklassen.



*Abb.3: Mediane (dicker waagrechtlicher Balken), Mittelwerte (+) und Begrenzung des oberen und unteren Quartils (Box) der Häufigkeitsklassen in den Lehrbüchern; 50% der Daten liegen innerhalb der Box; unterschiedliche Wörter (Types)*

<sup>9</sup> Zu HK 0 gehört nur der bestimmte Artikel, der bei der vorliegenden Untersuchung generell ignoriert wurde. In HK 1 gibt es keine Wörter in DeReWo.

**Verteilung der Wörter über die HK (Types nach Lektionen)**



*Abb.4: Mediane, Mittelwerte und Begrenzung des oberen und unteren Quartils der Häufigkeitsklassen in den Lehrbüchern; 50% der Daten liegen innerhalb der Box; unterschiedliche Wörter (Types) pro Lektion*

Abbildung 3 gibt genauer darüber Auskunft, in welchen Häufigkeitsklassen-Bereichen sich die Wörter in den einzelnen Lehrbüchern bewegen. Nur „München“ unterscheidet sich diesbezüglich signifikant von den anderen Büchern. Dort liegt der Median bei HK 10, bei den anderen bei 11. Somit weist „München“ also insgesamt etwas mehr höherfrequentes Vokabular auf als die anderen Bücher. Die Hälfte des Vokabulars bewegt sich dort in den HK 8 bis 12, während es bei den anderen Büchern bei 9 bis 13 liegt. Alle Bücher verwenden aber Wörter aus allen Häufigkeitsklassen von 0 bis 17.

Während in Abbildung 3 jedes unterschiedliche Wort mit seiner Häufigkeitsklasse nur einmal gezählt wurde, zeigt Abbildung 4 die Verteilung auf die Häufigkeitsklassen, wenn die unterschiedlichen Wörter pro Lektion separat gezählt werden. Wenn frequente Wörter also in mehreren Lektionen vorkommen, führt dies im Vergleich zur Berechnung in Abbildung 3 zu

einem tieferen Median der verwendeten Häufigkeitsklassen. Diese Berechnung ist insofern realistischer, als dass die mehrfache Verwendung frequenter Wörter „belohnt“ wird und die bloß einmalige Verwendung eines niedrigfrequenten Wortes nicht so sehr „bestraft“ wird. Denn es kann ja durchaus sinnvoll sein, auch seltene Wörter in einem Lehrbuch zu verwenden, sie aber nicht als Teil des primären Lernwortschatzes zu sehen.

Die Werte in Abbildung 4 unterscheiden sich dann auch leicht von den vorhergehenden. „Sommer“ und „München“ weisen mit HK 8 die niedrigsten Mediane auf, „Deutschstunde“ und „Modelle“ mit 10 die höchsten.

Weiter interessiert bezüglich des verwendeten Vokabulars der Verlauf über die Lektionen hinweg. So wäre z.B. zu erwarten, dass sich die Lehrbücher bemühen, in den ersten Lektionen eher einfaches, also frequentes Vokabular und in den späteren Lektionen auch niedrigerfrequentes Vokabular zu verwenden.

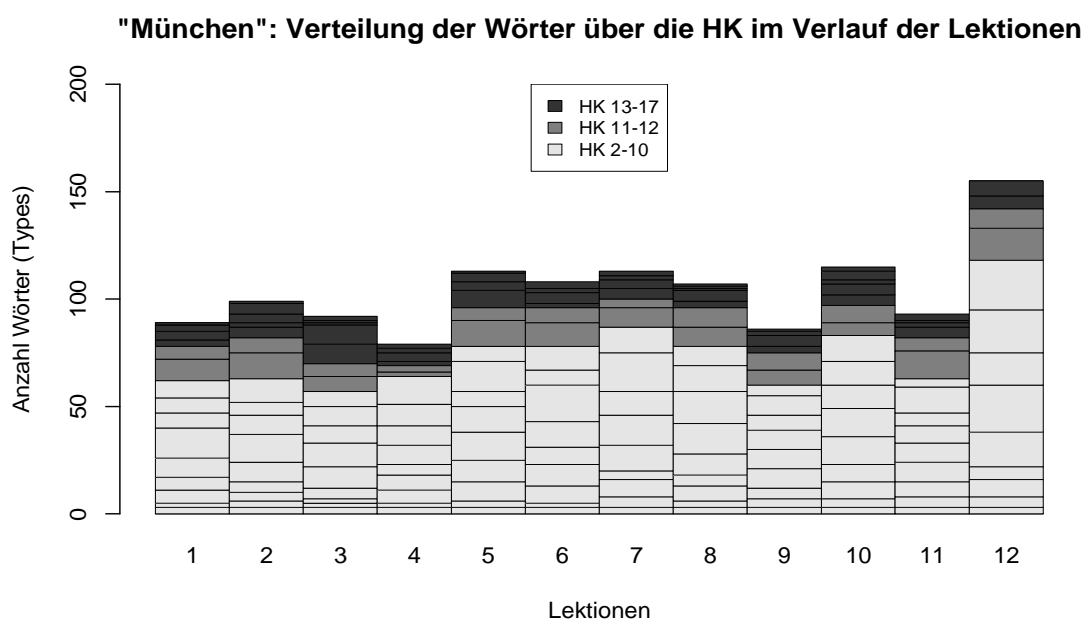
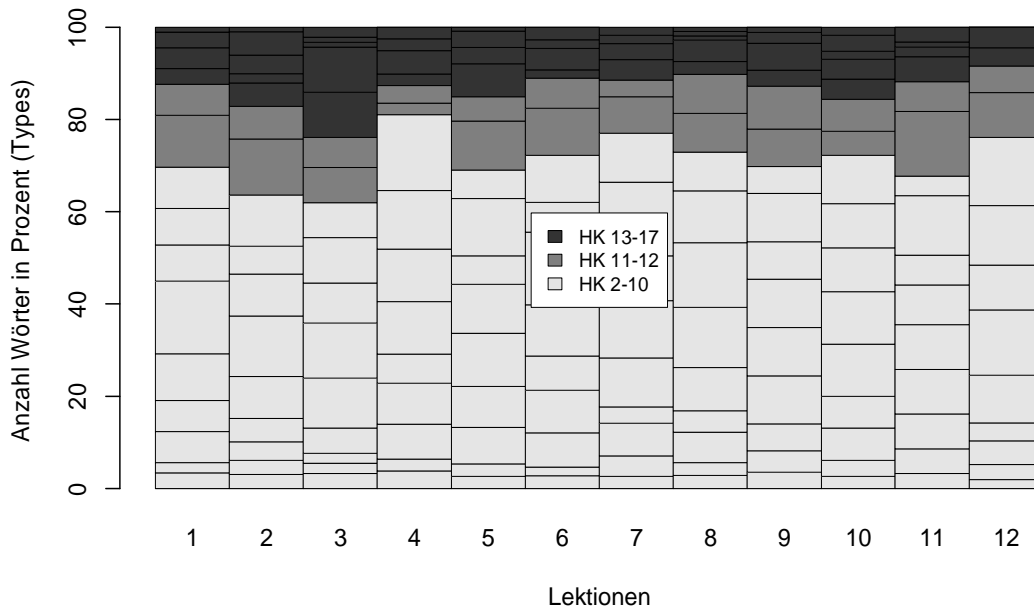


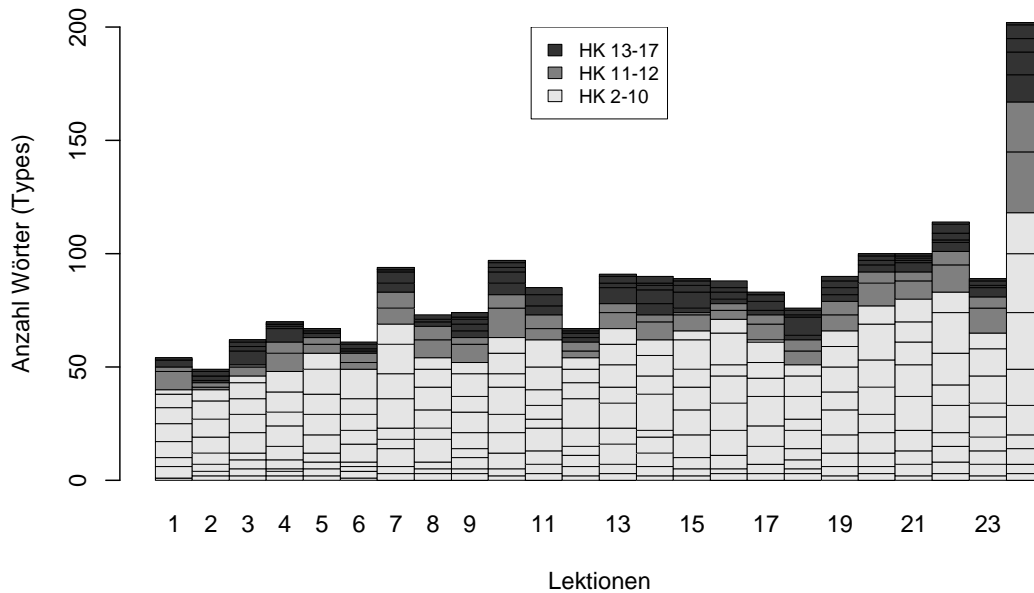
Abb.5: Entwicklung der Häufigkeitsklassen über die Lektionen; „München“ (absolute Frequenzen)

**"München": Verteilung der Wörter über die HK im Verlauf der Lektionen**



*Abb.6: Entwicklung der Häufigkeitsklassen über die Lektionen; „München“ (relative Frequenzen)*

**"Sommer": Verteilung der Wörter über die HK im Verlauf der Lektionen**



*Abb.7: Entwicklung der Häufigkeitsklassen über die Lektionen; „Sommer“ (absolute Frequenzen)*

### "Sommer": Verteilung der Wörter über die HK im Verlauf der Lektionen

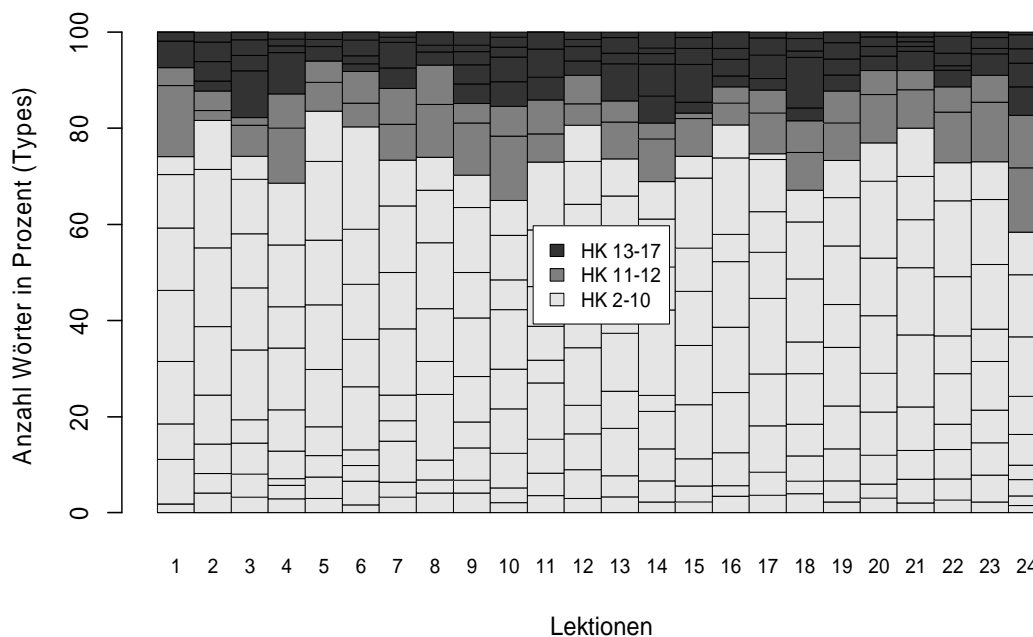


Abb.8: Entwicklung der Häufigkeitsklassen über die Lektionen; „Sommer“ (relative Frequenzen)

Solche Tendenzen sind aber kaum zu erkennen, wie exemplarisch die Daten für „München“ und „Sommer“ zeigen (Abb.5 bis Abb.8).<sup>10</sup>

#### 4 Differenzen zum Frequenzwortschatz und zu „Profile deutsch“

Zuletzt soll nun noch ein genauerer Blick auf das Vokabular die Differenzen zwischen der Wortverwendung in den Lehrbüchern, in DeReWo und in „Profile deutsch“ geworfen werden.

<sup>10</sup> Der starke Anstieg der unterschiedlichen Wörter in der letzten Lektion geht auf Vokabellisten am Ende der Bücher zurück.



Sommer	Farbkasten	München	Deutschstunde	Modelle
Endung	Anglistik	Brotzeit	Bücherregal	Berghof
Gaststube	Badeanzug	Metzger-	Dauerwelle	Bergsteigen
Geheim-	Blumengeschäft	meister	Deutschkurs	Blumengeschäft
nummer	Bücherregal	Portier	Ingenieurin	Briefumschlag
Gästezimmer	Deutschlehrer	Straßenbahn-	Kriminal-	Deutschlehrer
Lauch	Nikolaustag	haltestelle	geschichte	Endung
Nachspeise	Orakel	Verkehrsnetz	Rutsch	Hochzeitsfeier
Nickerchen	Plural	Verkehrsstau	Studenten-	Kopierer
Reiseroute	Shampoo	Weißbier	wohnheim	Mumm
Sprach-	Silvesterabend	erkälten	Vollkornbrot	Nomen
unterricht	Theater-	kämmen	betrinken	Notizblock
Toast	wissenschaft		erkälten	Otter
Weinkenner	Verb		nächst	Pilotin
appetitlich	Wohngegend		unsympathisch	Sehen
eichen	Woldecke		zwitschern	Stecker
mitfliegen	buchstabieren			Sweatshirt
schwindelig	nächst			Verb
				Weihnachtskarte
				Weinflasche
				Wohnen
				abgefahren
				anmalen
				irgendwohin
				unsympathisch

Tabelle 1: Beispiele für Wörter aus den Lehrbüchern mit Häufigkeitsklasse 17

Zunächst interessiert die Frage, welcher Art das Vokabular in den Lehrbüchern ist, das die höchste Häufigkeitsklasse (17) aufweist. In Tabelle 1 sind Beispiele für solche Wörter aufgeführt. Darunter fallen Wörter, die spezifisch für die Lebenswelt des Sprachunterrichts sind wie z.B. *Verb*, *Deutschkurs*, *Studentenwohnheim*, *Deutschlehrer*, *Nomen* oder *Kopierer*. Weiter fallen Komposita auf, deren Komponenten durchaus frequent sind und wo die Lehrbuchautoren wahrscheinlich davon ausgehen, dass sie transparent sind, da sie kompositionell verstehbar sind: *Brotzeit* (12/7)<sup>11</sup>, *Straßenbahnhaltestelle* (8/9/13), *Bücherregal* (9/14), *Blumengeschäft* (12/9) oder *Weinflasche* (11/12). Zudem ist immer denkbar, dass situativ Vokabular eingeführt wird, das für eine Geschichte wichtig ist oder zum

<sup>11</sup> In Klammern sind die Häufigkeitsklassen für die Glieder des Kompositums angegeben.

Lokalkolorit gehört, es aber nicht zwingend Ziel ist, das Wort zu lernen. Darunter könnten Wörter fallen wie *Weißbier*, *Orakel*, *Metzgermeister*, *Dauerwelle* etc.

Interessant ist nun, ob Wörter, die eine hohe Häufigkeitsklasse aufweisen, also eher selten sind, in „Profile deutsch“ genannt sind. In *Tabelle 2* sind Wörter aus „Farbkasten“ aufgeführt, deren HK-Wert 14 bis 17 ist, die aber in „Profile deutsch“ (A1 bis B2) vorkommen. Die Wörter wurden thematisch gruppiert, um einen besseren Überblick zu ermöglichen.

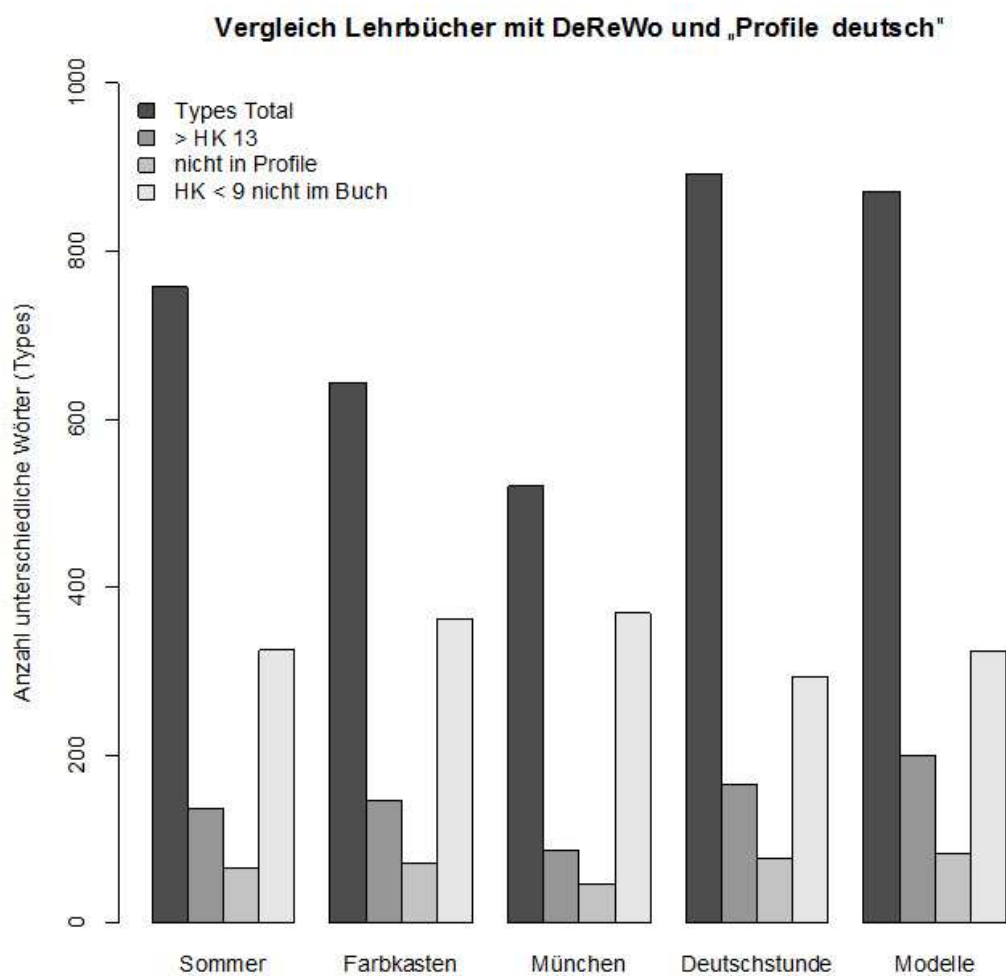
Auch hier fallen, wie bereits oben, Komposita auf, die mehr oder weniger kompositionell transparent sind, sowie Wörter aus der Lebenswelt der Studierenden. Auffallender ist jedoch, dass es sich um einen stark kommunikativ ausgerichteten Wortschatz handelt, der in der DeReWo-Liste mit hohen Häufigkeitsklassen ausgezeichnet ist.

<b>Freizeit</b>	<b>Nahrungsmittel</b>	<b>Unileben</b>	<b>Wohnung</b>	<b>Diverses</b>
Badeanzug	Banane	Grammatik	Schrank	gebären
Bikini	Birne	Hausaufgabe	Sofa	gefangen
duschen	Bonbon	Stundenplan	Regal	gehalten
Radtour	Brötchen	Vorlesung	Vase	drinnen
Picknick	Saft	Vokabel	Pullover	eilig
angeln	Salat	buchstabieren	Wäsche	Altersheim
Woldecke	Schinken	Betriebs-	Kissen	prima
losfahren	Schnaps	wirtschaft	Kommode	woher
mitkommen	Tomate		Häuschen	
Pfingsten	Weißwein		Lampe	
Kellner	Süßigkeit		putzen	
Portemonnaie	Mineralwasser		aufgeräumt	
spazieren	Pizza		aufräumen	
ausschlafen	Orangensaft		<b>Kochen</b>	
fernsehen	Plätzchen		gekocht	
Fußballplatz	Blumen-		Geschirr	
Geige	geschäft		Tasse	
Postkarte	Kiosk			
Beamtin	Verkäuferin			
Briefmarke				
Kugelschreiber				

*Tabelle 2: Wörter in „Farbkasten“ mit HK > 13, die aber in „Profile“ genannt sind (Auswahl, thematisch gruppiert)*

Dies liegt an der Zusammensetzung des DeReKo, dem Korpus, das der Frequenzliste DeReWo zugrunde liegt. Denn das Referenzkorpus enthält zu einem sehr großen Teil Zeitungstexte.<sup>12</sup> L2-Lehrbücher wollen normalerweise jedoch nicht nur die typische Zeitungssprache lehren, sondern legen mehr oder weniger Gewicht auf das Erlernen von kommunikativen Routinen wie sie in der Alltagskommunikation notwendig sind.

Zum Abschluss gibt Abbildung 9 Auskunft über die Übereinstimmung von DeReWo und „Profile deutsch“ mit den verwendeten Lehrbüchern.



*Abb.9: Anzahl unterschiedliche Wörter in den Lehrbüchern und Anteile davon die gemäß DeReWo selten (HK > 13) oder nicht in „Profile deutsch“ sind, sowie im Vergleich Anzahl Wörter, die gemäß DeReWo häufig sind (HK < 9), jedoch nicht im Lehrbuch erscheinen*

<sup>12</sup> Vgl. <http://www.ids-mannheim.de/kl/projekte/korpora/archiv.html> (29. Februar 2012).

Der jeweils erste Balken gibt die Anzahl unterschiedlicher Wörter im Buch wieder. Der zweite Balken zeigt, wie viele Wörter des Lehrbuchvokabulars gemäß DeReWo eher selten sind ( $HK > 13$ ). Hier ist ersichtlich, dass (in absoluten Werten) „Modelle“ am meisten, „München“ am wenigsten solche Wörter enthalten.

Der dritte Balken zeigt, wie gut das Vokabular des Lehrbuchs der Referenz „Profile deutsch“ angepasst ist, denn damit wird die Anzahl der Wörter ausgedrückt, die nicht in „Profile“ genannt sind.<sup>13</sup> Auch hier ist dieser Anteil bei „München“ am kleinsten, wobei der Wert bei allen Büchern niedrig ist.

Der vierte Balken gibt die Anzahl der Wörter an, die gemäß DeReWo zu den häufigen Wörtern gehören ( $HK < 9$ ), aber nicht in den Lehrbüchern erscheinen. Würde man sich bei der Wortschatzzusammenstellung also stärker an DeReWo orientieren, so müssten die seltenen Wörter (zweiter Balken) entfernt und die häufigen (vierter Balken) hinzugefügt werden. Allerdings wurden mit  $HK > 13$  für die seltenen und  $HK < 9$  für die häufigen Wörter relativ willkürliche Schnitte gewählt; sie orientierten sich an den Analysen zu den in den Lehrbüchern vertretenen Häufigkeitsklassen (vgl. *Abb.3*). Genauere Analysen müssten aber zeigen, welche Häufigkeitsklassen den Niveaus A1 bis B2 entsprechen sollten.

## 5 Fazit

Die Studie untersuchte die Frage, wie stark sich japanische Deutschlehrbücher an einem Frequenzwortschatz orientieren. Die zugrunde gelegte Frequenzliste ist DeReWo, was auf dem Deutschen Referenzkorpus DeReKo des Instituts für Deutsche Sprache beruht. Die Analysen zeigten, dass die

---

<sup>13</sup> Dieser Wert müsste noch um der Anzahl Wörter aus „Profile deutsch“, die nicht im Lehrbuch erscheinen, ergänzt werden. Da der Fokus aber nicht auf „Profile“ liegt, wurde der Übersichtlichkeit halber darauf verzichtet.

fünf untersuchten japanischen Deutschlehrbücher relativ ähnliche Profile aufweisen, was das verwendete Vokabular in Hinsicht auf Häufigkeitsklassen betrifft: Die Hälfte des Vokabulars bewegt sich zwischen den Häufigkeitsklassen 9 bis 13 bzw. 6 bis 12, wenn Wörter, die mehrmals im Buch vorkommen, stärker gewichtet werden. Nach der DeReWo-Liste gehören gut 4000 Wörter zu den Häufigkeitsklassen 0 bis 12, was weit mehr ist als L2-Lehrwerke der Stufen A1 bis B2 an Vokabular einführen können. Deshalb ist es nicht verwunderlich, dass in jedem Lehrbuch zwischen 300 und 350 Vokabeln fehlen, die gemäß DeReWo besonders häufig sind ( $HK < 9$ ).

Ein weiteres Problem stellt sich darin, welche Frequenzliste als Referenz verwendet werden soll. Die DeReWo-Liste beruht auf einem der größten verfügbaren deutschsprachigen Korpora, trotzdem zeigte sich bei der Analyse, dass die Liste stark von Zeitungstexten geprägt ist und deshalb Wortschatz aus dem kommunikativen Alltag der Studierenden teilweise fehlt.

Die benannten Probleme sind jedoch kein grundsätzlicher Einwand gegen die Orientierung an Frequenzlisten, um den Lernwortschatz zu bestimmen. Es wird sich jedoch lohnen, das zugrundeliegende Korpus sorgfältig zusammenzustellen und bezüglich Textsorten heterogener zu gestalten. Dann könnte ein frequenzbasierter Lernwortschatz eine transparentere Alternative zu De-fakto-Standards wie „Profile deutsch“ sein, der zudem nicht nur den Bereich A1 bis B2, sondern auch die fortgeschrittenen Stufen C1 und C2 abdecken könnte.



# METHODEN DER BESTIMMUNG DES KERNWORTSCHATZES DEUTSCH

Saburo OKAMURA / Willi LANGE / Joachim SCHARLOTH

## 1 Einleitung

Die Bestimmung des Kernwortschatzes ist für die Disziplin Deutsch als Fremdsprache von zentraler Bedeutung. Sie bietet nicht nur Lernenden eine wichtige Orientierung, welche Wörter zentral für eine gelingende Kommunikation sind. Kernwortschätze sind auch eine wichtige Grundlage für die Erstellung von Lehrbüchern und Curricula.

Es gibt eine Vielzahl von Grund- und Aufbauwortschätzen zur deutschen Sprache. Fast alle haben jedoch das Problem, dass sie nicht auf transparenter, empirischer Basis zusammengestellt wurden. In den meisten Fällen handelt es sich um Wortlisten, die auf individuellen Kenntnissen und Erfahrungen von Lehrenden oder Wissenschaftlern beruhen. Seltener fußen sie auch auf Recherchen in kleineren Textsammlungen (Korpora).

Im Projekt der Autoren „korpusgeleiteter Grund- und Aufbauwortschatz Deutsch“<sup>1</sup> soll zunächst ein umfangreiches eigenes Korpus erstellt werden. Im Anschluss sollen operationalisierbare Selektionskriterien ermittelt und ein Kernwortschatz des Deutschen berechnet werden. Für die Kernwortschatzlemmata sollen dann im Korpus datengestützte Beispielsätze herausgefiltert werden. Dies soll jedoch nicht willkürlich geschehen. Grundlage für die Auswahl sind vielmehr Kollokationsanalysen und Häufigkeitsbe-

---

<sup>1</sup> Dieses von 2011 bis 2014 dauernde Projekt wird durch Fördermittel des japanischen Ministeriums für Forschung und Kultur unterstützt (コーパス駆動型研究に基づく学習用ドイツ語語彙 科研費 23320120). Als Berater fungieren Stefan Gries (University of California, Santa Barbara) und Noah Bubenhofer (Institut für Deutsche Sprache Mannheim).

rechnungen für alle Elemente im Satz. Schließlich sollen die Ergebnisse so aufbereitet werden, dass sie für Lernende und Lehrende von Deutsch als Fremdsprache gut nutzbar sind. Schließlich werden die Ergebnisse des Projektes den Lernenden zur Verfügung gestellt. Sowohl Informationen im Internet als auch Publikationen in Buchform sind vorgesehen.

## **2 Zentraler Wortschatz des Deutschen**

Die Forschung zu zentralen Wortschätzen ist durch eine Vielzahl von Ansätzen geprägt, deren unterschiedliche Erkenntnisinteressen sich auch in der Terminologie widerspiegeln.<sup>2</sup> Für uns ist im Folgenden der Begriff „Zentraler Wortschatz“ der Oberbegriff für zwei verschiedene Typen begrenzender Wortschatzbeschreibung. Während „Kernwortschatz“ eine zweckfreie Beschreibung bezeichnet, wird „Grundwortschatz“ für alle Formen der Beschreibung verwendet, die eine sprachdidaktische Zielsetzung haben. Dabei ist zunächst unerheblich, ob die Zielsetzung muttersprachlich oder fremdsprachlich ist.

In der Forschung lassen sich drei Hauptansätze zur Bestimmung von zentralen Wortschätzen erkennen: der frequenzorientierte Ansatz, der kommunikativ-pragmatische Ansatz und der lexikografische Ansatz. Diese Stränge sollen hier kurz dargestellt und anschließend noch zusätzlich erläutert werden.

Der *frequenzorientierte Ansatz* ist ein linguistisches Paradigma, über Häufigkeitszählungen in Sprachdatensammlungen zu einer Wortschatz-begrenzung zu kommen. Frühe Beispiele wären Pfeffer (1970) und

---

<sup>2</sup> Siehe Haderlein, Veronika (2008): Das Konzept zentraler Wortschätze. Bestandsaufnahme, theoretisch-methodische Weiterführung und praktische Untersuchung. Diss. München. S. 43ff. für eine ausführliche Diskussion der Terminologieprobleme.



Rosengren (1972). Aktuelle Beispiele sind Jones/Tschirner (2006) und Tschirner (2008)<sup>3</sup>.

Der *kommunikativ-pragmatische Ansatz* ist ein auf sozialwissenschaftlichen und sprechakttheoretischen Grundkategorien beruhendes Paradigma. Über die Definition der relevanten (alltäglichen) Situationen und sozialen Rollen kommt man zu Sprechakten. Diesen Sprechakten werden dann die sprachlichen Mittel und somit auch der Wortschatz zugeordnet. Wichtige Beispiele sind hier die „Kontaktschwelle Deutsch“, das „Zertifikat Deutsch als Fremdsprache“ und „Profile deutsch“<sup>4</sup>.

Der *lexikographische Ansatz* ist ein linguistisches Paradigma, auf der Basis von vorhandenen Wörterbüchern und oder Wortschatzsammlungen einen Kern bzw. ein Zentrum zu destillieren. Neuere Repräsentanten sind Schnörch (2002) und Haderlein (2008)<sup>5</sup>.

Ein zentraler Punkt, in dem sich die drei Ansätze unterscheiden, ist die Datenbasis, die für die Wortschatzidentifizierung herangezogen wird: Während sich der frequenzorientierte Ansatz normalerweise auf primäre Daten (Texte, Korpora) stützt, nutzt der lexikografische Ansatz sekundäre Daten (vorhandene Wörterbücher oder Wortschatzsammlungen). Der kommunikativ-pragmatische Ansatz ist durch eine allenfalls nachrangige, selektive oder sporadische Nutzung empirischer Daten gekennzeichnet. Zusammenfassend lassen sich die Ansätze in der folgenden Übersicht darstellen:

---

<sup>3</sup> Pfeffer, Allan, J. (1970): Grunddeutsch. Basic (Spoken) German Dictionary. Englewood Cliffs; Rosengren, Inger (1970-1977): Ein Frequenzwörterbuch der deutschen Zeitungssprache. Die Welt. Süddeutsche Zeitung. 2 Bde. Lund; Jones, Randall L./Tschirner, Erwin (2006): A Frequency Dictionary of German. Core vocabulary for learners. London; Tschirner, Erwin (2008): Grund- und Aufbauwortschatz Deutsch als Fremdsprache nach Themen. Berlin.

<sup>4</sup> Baldegger, Markus/u.a. (1980): Kontaktschwelle Deutsch als Fremdsprache. Berlin; Deutscher Volkshochschulverband/Goethe-Institut (1985): Das Zertifikat Deutsch als Fremdsprache. 3. Aufl. Bonn/Frankfurt; Glaboniat, Manuela/u.a. (2005): Profile deutsch. Berlin.

<sup>5</sup> Schnörch, Ulrich (2002): Der zentrale Wortschatz des Deutschen. Strategien zu seiner Ermittlung, Analyse und lexikografischen Aufarbeitung. Tübingen; Haderlein (2008).

Ansatz	Frequenz-orientierter Ansatz	Kommunikativ-pragmatischer Ansatz	Lexikografischer Ansatz
Beispiele für Vertreter	Pfeffer (1970) Rosengren (1972-1977), Jones/Tschirner (2006), Tschirner (2008)	Baldegger/ u.a. (1980) Volkshochschulverband/ Goethe-Institut (1985) Glaboniat/ u.a. (2005)	Schnörch (2002) Haderlein (2008)
Empirische Basis	Eigene Datenerhebungen oder Zeitungstexte	keine	Wörterbücher, Wortschatzsammlungen
Hauptkriterium der Selektion	Häufigkeit	Kommunikative Relevanz	Schnittmengenbildung oder Bündel aus statistischen Maßen + Introspektion

*Tabelle 1: Übersicht über die Ansätze zur Bestimmung des zentralen Wortschatzes*

Versucht man, ein Entwicklungsbild der skizzierten Ansätze für den Bereich Deutsch als Fremdsprache der letzten 50 Jahren zu zeichnen, so lässt sich Folgendes sagen: Die zu Beginn recht starke Frequenzorientiertheit wird zunehmend abgelöst und dann völlig dominiert von der Orientierung am kommunikativ-pragmatischen Paradigma. Mit der Übernahme von korpuslinguistischen Verfahren im Bereich Deutsch als Fremdsprache – besonders seit 2005 – scheint es wieder eine stärkere Tendenz zu empirischen Verfahren zu geben. Der lexikografische Ansatz ist konstant als Nebenströmung vorhanden.

In unserem Argumentationszusammenhang sei hier noch auf einige weitere wichtige Punkte hingewiesen:

a) Frequenzorientierter Ansatz:

- Besonders die frühen Arbeiten mussten mit technischen Begrenzungen kämpfen. Entsprechend klein waren die Korpora.
- Das Hauptaugenmerk richtete sich auf die absolute Frequenz von Wörtern im jeweiligen Korpus. Eine Untersuchung der Verteilung in

Subkorpora war ebenso wenig möglich wie die automatische Analyse der Umgebung von Wörtern (z.B. Kookurrenzanalysen).

b) Kommunikativ-pragmatischer Ansatz:

- Der kommunikativ-pragmatische Ansatz setzte sich historisch ganz bewusst und explizit vom frequenzorientierten Ansatz ab. Die Hauptkritikpunkte waren: Nackte Wortschatzlisten haben keinen oder nur einen sehr eingeschränkten Wert. Hohe Frequenz und kommunikative Wichtigkeit fallen nicht notwendigerweise zusammen. Die frequenzorientierten Ansätze berücksichtigen nicht die potentiellen Grundwortschatznutzer, die Kommunikationssituationen und die relevanten Handlungsmuster.
- Innerhalb der Praxis Deutsch als Fremdsprache hat der kommunikativ-pragmatische Ansatz besonderes Gewicht. Internationale Deutschprüfungen wie das Zertifikat Deutsch (früher „Zertifikat Deutsch als Fremdsprache“, seit neuestem „B1-Prüfung“) oder auch Start 1 (A1-Prüfung) und Start 2 (A2-Prüfung) sind dem kommunikativ-pragmatischen Ansatz verpflichtet<sup>6</sup>. Mit der Ausrichtung der zentralen Prüfungen Deutsch als Fremdsprache am kommunikativ-pragmatischen Ansatz ergibt sich dann, dass auch die Lehrwerke diesem Ansatz stark verpflichtet sind.
- Am kommunikativ-pragmatischen Ansatz lässt sich kritisieren: die mangelnde theoretische Fundierung und empirische Absicherung der Basiskategorien (Situationen, Themen, Sprechakte), die Intransparenz der Wortschatzselektionskriterien und die Intransparenz der

---

<sup>6</sup> Die deutsche Ausarbeitung des europäischen Referenzrahmens – Glaboniat, Manuela/u.a. (2005): Profile deutsch. Berlin – ist eine explizite Überarbeitung von Baldegger, Markus/u.a. (1980): Kontaktschwelle Deutsch als Fremdsprache. Berlin. Im Prinzip werden die sprachlichen Mittel der “Kontaktschwelle Deutsch” aus ihrem ursprünglichen Rahmen (starke Mündlichkeit und face-to-face Kommunikation) herausgelöst und in den neuen Rahmen des europäischen Referenzrahmens mit den breiten Kann-Bestimmungen eingebaut.

Wortschatzzuordnung (zu den einzelnen Themen bzw. zu den einzelnen Lernstufen).

c) Lexikografischer Ansatz:

- Schnörch (2002) versucht, den deutschen Kernwortschatz über die Schnittmengenbildung von sieben vorhandenen Grundwortschätzen zu beschreiben. Schnörchs Forschungsinteresse liegt aber mehr in der Etablierung einer Grundwortschatzlexikografie, die sich besonders der lexikografischen Aufbereitung und Strukturierung widmet. Seine intendierte zentrale Grundeinheit ist konsequenterweise nicht das Lexem sondern das Semem.
- Haderlein (2008) widmet sich sehr ausführlich der Diskussion von möglichen Selektionskriterien für Lexeme. Leider ist die Operationalisierung der Kriterien (z.B. „Temporale Stabilität“ mit Hilfe von zweisprachigen Wörterbüchern) nicht überzeugend. Auch die praktische Durchführung der Zentralwortschatzselektion bleibt auf den Buchstaben 0 des Duden Universalwörterbuchs beschränkt.

Nach einem kurzen Blick auf die Grundwortschatzarbeit in Japan wenden wir uns der Darstellung der Verfahren zu, die in unserem Wortschatzprojekt verwendet werden.

### **3 Grundwortschatzarbeit in Japan**

Die Grundwortschatzarbeit in Japan geht in erster Linie von der didaktischen Fragestellung aus, wie viele Wörter die Lernenden im Deutschunterricht lernen können und sollen, um sich einigermaßen im Deutschen zurechtfinden zu können. Grundwortschatzlisten bzw. -sammlungen findet man in großen Mengen.

Es gibt einige ernsthafte Versuche, den Grundwortschatz zu bestimmen, wie z.B. Nishina/ u.a. (1971)<sup>7</sup> und Asakura/ u.a. (1994)<sup>8</sup>. Nishina et al. versuchten, Wortfrequenzen aus den Analysen etlicher deutscher Lehrbücher festzustellen und darauf aufbauend eine Grundwortschatzliste zu erstellen. Dies scheiterte jedoch letztendlich an der geringen Quantität der Quellenbücher.

Asakura/ u.a. (1994) versuchten, aus vorhandenen deutschen Wortschatzlisten wie z.B. dem „Grund- und Aufbauwortschatz Deutsch“<sup>9</sup>, verschiedenen deutsch-japanischen Wörterbüchern und auch ausgewählten Büchern eine Grundwortschatzliste zu erstellen. Asakura/ u.a. (1994) könnte man somit eher dem lexikografischen Ansatz zuordnen. Aber letztlich ausschlaggebend bei der Auswahl der Einträge waren die Erfahrungswerte der Wissenschaftler bzw. Deutschlehrer.

„Wichtige“ Wörter als solche werden in den meisten gängigen deutsch-japanischen Wörterbüchern markiert (oft rot oder fett). Aber in der Regel erfährt man nicht, wie diese Wörter ausgewählt wurden. Dass die Auswahl der „wichtigen“ bzw. „wichtigsten“ Wörter nicht transparent und nachvollziehbar vorgenommen worden ist, ist die gemeinsame Schwäche.

Während verschiedene großangelegte deutsche Korpora inzwischen vorhanden sind und man davon Gebrauch machen könnte, ist das Access-Wörterbuch von Zaima (2010)<sup>10</sup> zurzeit das einzige auf dem Markt, das sich rühmt, auf Häufigkeitslisten basiert zu sein<sup>11</sup>. Das Access-Wörterbuch ist somit der einzige japanische Vertreter eines frequenzorientierten Ansatz-

---

<sup>7</sup> Nishina, Takemitsu/u.a. (1971): Wichtige Wörter für japanische Studenten. In: Ausschuss für Deutschunterricht und Sprachwissenschaft der Gesellschaft für Germanistik Japans (Hrsg.): Grundprobleme des Deutschunterrichts in Japan. S. 233-534.

<sup>8</sup> Asakura, Takumi/u.a. (1994): 基礎語彙策定の試み(Grundwortschatz). In: Berichte des Japanischen Deutschlehrerverbandes 46, S. 5-12.

<sup>9</sup> Oehler, Heinz (1991): Grund- und Aufbauwortschatz Deutsch. Stuttgart.

<sup>10</sup> Zaima, Susumu (Hg.) (2010): アクセス独和辞典. 3. Aufl., Tokyo.

<sup>11</sup> Zaima (2010: Vorwort, o. S.). Die Literaturliste auf S. 2117 nennt die Namen der Korpora: DEREKO (IDS), dewac (Lexical Computing Ltd.), DWDS (BBAW) und Wortschatz Portal (Universität Leipzig).

zes. Aber wie man tatsächlich die „wichtigsten“ Wörter aus diesen Listen ausgewählt hat, obwohl man annehmen könnte, dass es zwischen diesen Korpora durchaus Differenzen geben müsste, erfährt man leider nicht.

#### **4 Das Forschungsprojekt „Korpusgeleiteter Grund- und Aufbauwortschatz Deutsch“**

Im Forschungsprojekt „Korpusgeleiteter Grund- und Aufbauwortschatz Deutsch“ verfolgen wir einen frequenzorientierten Ansatz. So lässt sich die Entscheidung, warum ein Wort zum Kernwortschatz des Deutschen gehört und deshalb in einen Grundwortschatz für Lernende aufgenommen werden sollte, methodisch transparent und damit nachvollziehbar machen. Wir vertreten dabei ähnlich wie Tschirner (2008) einen radikal frequenzorientierten Ansatz. Wir ziehen Lemmafrequenzen nicht nur in Zweifelsfällen als Entscheidungshilfe heran sondern machen sie überhaupt zur Grundlage der Berechnung des Kernwortschatzes. Unser Vorgehen möchten wir daher als *datengeleitet* (im Gegensatz zu datenbasiert) bezeichnen.<sup>12</sup>

Anders als für Tschirner bedeutet Frequenzorientierung für uns jedoch nicht ausschließlich, eine Rangfolge der relativen Frequenzen von Lemmata in einem Korpus als Kriterium für die Aufnahme in den Kernwortschatz zu wählen. Wir differenzieren den Frequenzaspekt vielmehr in mehrere Dimensionen:

1. *Relative Frequenz*: Zum zentralen Wortschatz zählen wir all jene Lemmata, die eine hohe relative Frequenz aufweisen oder – anders ausgedrückt – eine niedrige Häufigkeitsklasse haben.
2. *Temporale Stabilität*: Ein weiteres Kriterium für die Zugehörigkeit zum zentralen Wortschatz ist das über einen längeren Zeitraum hin gleichmäßig häufige Auftreten eines Lemmas oder die Zunahme der Häufigkeit seines Auftretens über einen längeren Zeitraum.

---

<sup>12</sup> Vgl. Tognini-Bonelli, Elena (2001): *Corpus linguistics at work*. Amsterdam: Benjamins (Studies in corpus linguistics 6).

3. *Themenübergreifende Stabilität*: Ein weiteres Kriterium für die Zugehörigkeit zum zentralen Wortschatz ist seine ressort- bzw. sachbereichsspezifische Stabilität, d.h., dass ein Lemma in unterschiedlichen Teilkorpora gleichmäßig häufig auftritt.
4. *Produktivität*: Ein weiteres Kriterium ist die Wortbildungspotenz lexikalischer Morpheme. Wenn ein lexikalisches Morphem durch Komposition oder Derivation (a) Teil vieler Lemmata ist und (b) diese Lemmata zudem häufig vorkommen, dann ist dies ein Kriterium dafür, das zum lexikalischen Morphem gehörige Lemma zum Kernwortschatz zu zählen.

Es ist offensichtlich, dass bei unserer datengeleiteten Vorgehensweise die *Wahl des Korpus* von entscheidender Bedeutung ist. Anders als Tschirner, der die Ausgewogenheit und Repräsentativität seines Herder/BYU-Korpus für zentral hält<sup>13</sup>, sind wir der Ansicht, dass ein ausgewogenes und repräsentatives Korpus, das die Vielzahl der unterschiedlichen kommunikativen Verwendungsweisen des Deutschen in ihrem jeweiligen Verhältnis abbildet, nicht existieren kann. Zudem halten wir den Umfang des Herder/BYU-Korpus mit ca. 4,2 Mio. Wortformen für völlig unzureichend, um belastbare Aussagen über Stabilität und Produktivität zu machen. Auch im Hinblick auf die zuverlässige Ermittlung relevanter schweizerischer und österreichischer Varianten des Standarddeutschen sind erhebliche Zweifel angebracht, ob der Umfang der länderspezifischen Teilkorpora von 420.000 (Schweiz) und 840.000 Wortformen (Österreich) ausreicht.

Aus diesen Gründen haben wir uns dafür entschieden, zum einen mit sehr großen Korpora, d.h. Korpora mit mehr als 100 Millionen Wörtern, zu arbeiten und zum anderen die Vorstellung von Repräsentativität und Ausgewogenheit zugunsten der Abbildung prototypischer Kommunikations-

---

<sup>13</sup> Vgl. Tschirner, Erwin (2005): Korpora, Häufigkeitslisten, Wortschatzerwerb. In: Heine, Antje/Hennig, Mathilde/Tschirner, Erwin (Hrsg.), *Deutsch als Fremdsprache – Konturen und Perspektiven eines Fachs*. München: Iudicium. S. 133-149.

konstellationen aufzugeben: Unser Korpus soll einerseits massenkommunikative Texte enthalten, d.h. mehrfachadressierte und konzeptionell schriftliche Texte, andererseits aber auch Texte, die persönlich adressiert und konzeptionell mündlich sind. Für die massenkommunikativen Texte greifen wir auf Zeitungen zurück, die persönlich adressierten Texte entnehmen wir Online-Diskussionsforen.

Im Folgenden möchten wir anhand zweier exemplarischer Analysen demonstrieren, warum die Orientierung an prototypischen Kommunikationskonstellationen sinnvoll und die Differenzierung in mehrere Frequenzdimensionen bei der Berechnung des zentralen Wortschatzes notwendig sind.

## **5 Exemplarische Analyse I: Häufigkeitsklassen in prototypischen Kommunikationskonstellationen**

Um zu illustrieren, wie unterschiedlich einzelne Lemmata in den Texten der oben skizzierten Kommunikationskonstellationen verteilt sein können, haben wir eine exemplarische Analyse anhand zweier kleinerer Korpora vorgenommen: Dem Unterforum „Allgemeine Themen“ des Online-Diskussionsforums seniorentreff.de (6.357.610 Wortformen aus den Jahren 2005-2010), sowie den Jahrgängen 2009 und 2010 der Wochenzeitung DIE ZEIT (9.706.498 Wortformen). Die Korpora wurden tokenisiert und mit dem TreeTagger<sup>14</sup> lemmatisiert.

Bei unserer Analyse diene die Häufigkeitsklasse als statistische Kenngröße. Die Häufigkeitsklasse (HK) ist ein statistisches Maß für die Gebrauchshäufigkeit eines Wortes in einem Korpus. Bei der Berechnung dient

---

<sup>14</sup> Vgl. Schmid, Helmut (1994): Probabilistic part-of-speech tagging using decision trees. Stuttgart: [Working paper] Universität Stuttgart, Institut für maschinelle Sprachverarbeitung; Schiller, Anne/Teufel, Simone/Thielen, Christine (1995): Guidelines für das Tagging deutscher Textcorpora mit STTS. Stuttgart: [Working Paper] Universität Stuttgart, Institut für maschinelle Sprachverarbeitung, Tübingen: Seminar für Sprachwissenschaft.



die am häufigsten auftretende Wortform (im Deutschen *der*) bzw. das frequenteste Lemma (im Deutschen der bestimmte Artikel *der/die/das*) als Referenzpunkt. Die Häufigkeitsklasse berechnet sich aus dem ganzzahligen Anteil des Logarithmus zur Basis 2 des Quotienten der Frequenz des häufigsten Wortes durch die Frequenz des betreffenden Wortes<sup>15</sup>. Je kleiner die Häufigkeitsklasse ist, desto frequenter ist also das Wort. Je kleiner zudem die Häufigkeitsklasse ist, desto weniger Wörter enthält sie: jede Klasse N enthält nach dem Zipf'schen Gesetz ca.  $2^N$  Wörter.

Die folgende Tabelle zeigt eine kategorisierte Auswahl jener Lemmata, die im seniorentreff-Korpus signifikant häufiger vorkommen als im Zeit-Korpus:

Lemma	Häufigkeitsklasse	
	seniorentreff.de	DIE ZEIT
<i>Wörter, die auf Sprechakte verweisen</i>		
danke	6	11
Glückwunsch	8	13
danken	6	10
Dank	6	10
bitte	6	9
bitten	6	8
Wunsch	6	8
leid	8	11
Entschuldigung	8	10
Bitte	8	10
entschuldigen	8	10
<i>Partikel: vor allem Modal- und Gradpartikel</i>		
mal	3	5
ja	3	5
sehr	3	5
wohl	4	6
gerne	5	7
leider	5	7
bestimmt	6	8
sicherlich	7	10
schade	7	10
meistens	7	9

<sup>15</sup> Bezogen auf die Wortformen im Deutschen sieht die Berechnungsformel so aus:  
 $HK(\text{Wort}) = \text{ganzzahliger Anteil}(\log_2(\text{anzahl}('der')/\text{anzahl}(\text{wort})))$ .

halt	7	9
total	7	9
<b><i>Kommunikation eröffnende oder schließende Wörter</i></b>		
hallo	5	12
Gruß	5	11
wünschen	5	8
Liebe	5	7
Hallo	6	15
herzlich	6	10
Spaß	6	9
freundlich	6	8
grüßen	7	10
Moin	8	15
Hi	8	14

Erwartungsgemäß handelt es sich dabei um Wörter, die auf Sprechakte verweisen, welche der Beziehungspflege dienen, um Modal- und Gradpartikel, sowie um Ausdrücke der Kontaktaufnahme und Kontaktbeendigung. Umgekehrt sind für das ZEIT-Korpus insbesondere Präpositionen (offenbar wegen der im Nominalstil häufiger auftretenden Präpositionalgruppen) und Inhaltswörter aus dem Bereich Politik und Wirtschaft signifikant. Die folgende Tabelle enthält eine Auswahl:

<b>Lemma</b>	<b>Häufigkeitsklasse</b>	
	<b>seniorentreff.de</b>	<b>DIE ZEIT</b>
<b><i>Präpositionen</i></b>		
in	2	1
von	2	1
aus	3	2
am	4	3
über	4	3
vor	4	3
seit	6	4
vom	5	4
unter	5	4
bis	5	4
gegen	5	4
neben	8	6
<b><i>Inhaltswörter</i></b>		
Jahr	4	3
deutsch	6	4

Land	6	4
groß	5	4
Prozent	9	5
politisch	8	5
Million	8	5
Euro	7	5
Staat	7	5
Stadt	7	5
Woche	7	5
Ende	7	5
Präsident	10	6
Universität	10	6
Krise	11	6
britisch	11	7
national	11	7
Kilometer	11	7
Klimaschutz	17	9
schwarz-gelb	17	9

Zwar sind die Ergebnisse bis zu einem gewissen Grad so zu erwarten, doch belegen sie beeindruckend die Notwendigkeit, neben Zeitungstexten auch persönlich adressierte nächsprachliche Texte in die Analyse einzubeziehen. Die Analyse liefert zudem einen Beleg dafür, dass Online-Diskussionsforen eine geeignete Quelle für eine prototypisch nächsprachliche Kommunikationskonstellation sind.

## 6 Exemplarische Analyse II: Temporale Stabilität

Temporale Stabilität ist ein wichtiges Kriterium für die Aufnahme eines Wortes in den Zentralen Wortschatz. So können sprachliche Moden, wie die inzwischen kaum noch verwendete Bezeichnung „Luder“ für einen bestimmten Typus Frau Anfang der 2000er Jahre, oder aus dem Zeitgeschehen motiviert hochfrequent auftretende Wörter wie „Millenium“ herausgefiltert werden.

Zur Berechnung der temporalen Stabilität bedienen wir uns der Maßzahl Gries' DP<sup>16</sup>. Bei dieser handelt es sich um einen normalisierten Wert (zwischen 0 und 1), der Signifikanzwerte für die Abweichung relativer Frequenzen in Teilkorpora zusammenfasst. Je höher der DP ist, desto ungleichmäßiger ist ein Lemma in den Teilkorpora eines Korpus verteilt. Um temporale Stabilität zu messen, teilen wir unser Korpus in zeitlich definierte Subkorpora ein.

Die folgende Analyse wurde anhand eines Korpus mit sämtlichen Texten der Ressorts Panorama und Politik von SPIEGEL ONLINE aus den Jahren 1998 bis 2010 durchgeführt. Das Korpus umfasst insgesamt rund 44 Millionen Wortformen. Die Lemmata verteilen sich wie folgt auf die Häufigkeitsklassen (nur HK 0-14):

Häufigkeitsklasse	Anzahl Lemmata
HK 0	4
HK 1	6
HK 2	13
HK 3	17
HK 4	26
HK 5	78
HK 6	163
HK 7	376
HK 8	660
HK 9	1066
HK 10	1747
HK 11	2867
HK 12	4534
HK 13	7003
HK 14	10127

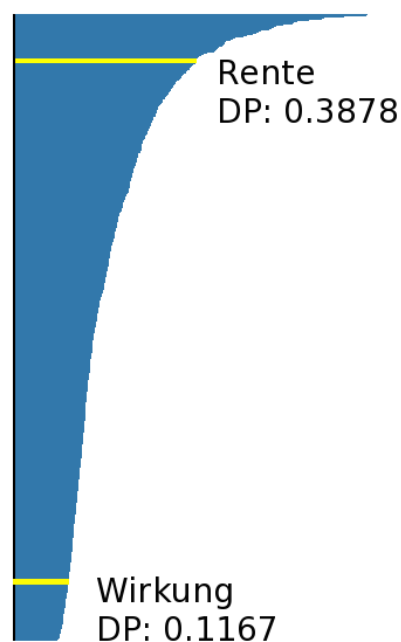
Die Verteilung zeigt, dass sich insgesamt 2409 Lemmata in den Häufigkeitsklassen 0 bis 9 finden. Auch wenn sich unter den Lemmata noch Eigennamen und weitere *named entities* befinden, ist doch zu erwarten, dass sich in der Häufigkeitsklasse 9 besonders viele Zweifelsfälle im Hinblick

---

<sup>16</sup> Vgl. Gries, Stefan Th. (2008): Dispersions and adjusted frequencies in corpora. In: International Journal of Corpus Linguistics 13, S. 403-437.

auf die Aufnahme in einen Grundwortschatz für Lernende finden. Denn solche Grundwortschätze umfassen meist zwischen 1000 und 2000 Lemmata. Gries' DP kann hier eine Entscheidungshilfe sein.

Visualisiert man die Verteilung der Gries' DP-Werte auf der Basis monatlicher Teilkorpora aller Lemmata der Häufigkeitsklasse 9, so zeigt sich, dass es teils erhebliche Unterschiede in der Stabilität der Lemmata gibt. Die folgende Abbildung soll das veranschaulichen:



*Abb.1: Verteilung der monatsweise berechneten Gries' DP-Werte der Lemmata der Häufigkeitsklasse 9 im SPIEGEL ONLINE-Korpus (Teilkorpus Politik und Panorama)*

Während sich am äußersten oberen Ende der Skala Namen wie „Guttenberg“ (DP: 0.7506) oder „Milosevic“ (DP: 0.7031) mit hohem DP finden, rangieren am unteren Ende mit den niedrigsten DPs der Häufigkeitsklasse Verben und Adjektive wie „antworten“ (DP: 0.0914), „rein“ (DP: 0.0951) oder auch Partikel wie „indem“ (DP: 0.0968). In der Grafik ist auch die Lage der Substantive „Rente“ und „Wirkung“, welche sich in ihrem DP-Wert stark unterscheiden, in der DP-Verteilung visualisiert. Auf-

grund der höheren Stabilität müsste bei der Wahl zwischen beiden Lemmata dem Wort „Wirkung“ der Vorzug gegeben werden. Allerdings ist hier noch einmal darauf hinzuweisen, dass sich diese Entscheidung so ergeben würde, wenn man ihr nur das Teilkorpus (Spiegel online, Panorama und Politik) zugrundelegt.

## **7 Fazit**

Im vorliegenden Aufsatz haben wir einen Überblick über unterschiedliche Ansätze zur Bestimmung des zentralen Wortschatzes des Deutschen gegeben. Dabei haben wir unser eigenes Projekt im frequenzorientiert-datengeleiteten Paradigma verortet und in Abgrenzung zu den bisherigen Ansätzen den Begriff der Frequenzorientierung in unterschiedliche Dimensionen (relative Frequenz, Stabilität, Produktivität) differenziert. Anhand einiger exemplarischer Analysen konnten wir illustrieren, welches Potenzial dieser Ansatz für eine transparente und valide Bestimmung des zentralen Wortschatzes haben kann. Es muss einer weiteren Publikation vorbehalten bleiben, die korpuslinguistische Operationalisierung der Kriterien „themenübergreifende Stabilität“ und „Produktivität“ darzustellen und das mathematische Modell zur Verrechnung der unterschiedlichen Faktoren sowie deren Gewichtung zu diskutieren.

# 日本の DaF におけるドイツ語基礎語彙へのアプローチ

岩崎克己

## 1 はじめに

本稿の目的は、広島大学において 2008 年に作られたドイツ語初学者用の改訂版語彙リスト（以下、「広大語彙リスト」と略す）<sup>1</sup>の単語を、コミュニケーション的なアプローチに基づいて作られた Glaboniat/u.a. (2005)<sup>2</sup>の語彙リスト、および Herder / BY-Korpus を基に頻度順のアプローチで作られた Tschirner (2008)<sup>3</sup>の語彙リストとそれぞれ比較することで、異なったアプローチにおける語彙リストの相違点と共通点について、具体的な事例をもとに何が言えるかを考察することである<sup>4</sup>。本稿ではまず、日本のドイツ語教育における基礎語彙の選定の際に、頻度順のアプローチのみに依拠しようとする場合の問題点について簡単に触れ、次に、「広大語彙リスト」について概略を紹介する。最後に、日本のドイツ語教育の文脈を踏まえたコミュニケーション・アプローチの観点から、上記の 3つの語彙リストに含まれる単語の比較を行い、それぞれの特徴について考察する。

## 2 基本語彙の選定方法をめぐる 3つのアプローチ

2011 年度独文学会秋季研究発表会におけるシンポジウム『ドイツ語基礎語彙：辞書学と外国語教育の観点から』（2011 年 10 月 16 日）の開催趣旨の中で、

---

<sup>1</sup> このリストの原型は、もともとは、広島大学外国語教育研究センター・カリキュラム専門部会の委託を受けて筆者が同僚の吉満たか子氏、Axel Harting 氏らとともに 2005 年から 2007 年にかけて作った基礎語彙リスト（800 語）と追加語彙リスト（350 語）である。ただし、本稿では、それを 2008 年 10 月に岩崎が個人的に改訂・補充した改訂版（約 850 語の基礎語彙リストと約 350 語の追加語彙リスト）の方を取りあげる。これらの成立の経緯については、詳しくは本稿の第 4 節参照。なお、リストそのものは、以下の URL からダウンロードできる。http://home.hiroshima-u.ac.jp/katsuiwa/grundwortschatz.htm (29.2.2012)

<sup>2</sup> Glaboniat, M./Müller, M./Schmitz, H./Rusch, P./Wartenschlag, L. (2005): Profile deutsch Niveau A1-A2 B1-B2 C1-C2. Berlin/München/ Wien/ Zürich/New York: Langenscheidt. なお、本稿で「Glaboniat/u.a. (2005)の語彙リスト」という場合、具体的には Glaboniat/u.a. (2005)付属の CD-ROM 中の Wortregister に含まれる単語や表現の全体を指す。

<sup>3</sup> Tschirner, Erwin (2008): Grund- und Aufbauwortschatz Deutsch als Fremdsprache nach Themen - Lernwörterbuch. Berlin: Cornelsen.

<sup>4</sup> 全データの比較については、「広大語彙リスト」の各単語を品詞ごとにアルファベット順に並べ、Glaboniat/u.a. (2005)における言語能力のレベル(A1～B2)と Tschirner(2008)における頻度順位をそれぞれ書き込んだ対照表を作成した。それらは、以下の URL からダウンロードできる。http://home.hiroshima-u.ac.jp/katsuiwa/grundwortschatz.htm (29.2.2012)

岡村三郎らは、基礎語彙の選定方法について、大きく以下の3つのアプローチがあることを、それぞれの代表的な文献をあげながら紹介している。

- 1) 頻度を選択基準とするコーパス言語学的なアプローチ (frequenzorientierter Ansatz)<sup>5</sup>、
- 2) コミュニケーションの場面を想定しそこでの重要性を選択基準とするコミュニケーション的アプローチ (kommunikativ-pragmatischer Ansatz)<sup>6</sup>、
- 3) 既存の教科書や辞書などを基にした統計的なデータと内観の重なりを基準とした辞書学的なアプローチ (lexikografischer Ansatz)<sup>7</sup>。

岡村らはこれらのアプローチの今日的な重要性として、1) コーパス言語学的なアプローチについては、Jones/Tschirner (2006)や Tschirner (2008)の出版以降ドイツ語圏におけるコーパス研究の隆盛とともにこの方法が注目を浴びてきていることを、2) コミュニケーション的アプローチについては、CEFR<sup>8</sup>に準拠した Glaboniat/u.a. (2005)による言語能力の記述や語彙リストが ZD (Zertifikat Deutsch) や SD (Start Deutsch) 等の国際的なドイツ語の資格試験の基準となり、それゆえ今日のドイツ語圏で出版されているドイツ語教科書の能力レベルの記述 (A1～C2) の事実上の指標にもなっていることを、3) 一種の折衷的な立場である辞書学的なアプローチについては、日本のドイツ語教科書や辞書における共通部分としてのいわゆる「重要語」が、日本における検定試験等の基礎として使われていることなどを指摘している。

---

<sup>5</sup> 岡村らが挙げているのは、註 2 で挙げた Tschirner (2008)および以下の文献。Jones, Randall/Tschirner, Erwin (2006): A Frequency Dictionary of German: Core Vocabulary for Learners. Routledge; Pfeffer, J. Alan (1970): Grunddeutsch: Basic (Spoken) German Dictionary for Everyday Usage. Englewood Cliffs, Prentice-Hall; Rosengren, Inger (1972, 1977): Ein Frequenzwörterbuch der deutschen Zeitungssprache. Die Welt, Süddeutsche Zeitung. Band 2. Lund: Gleerup.

<sup>6</sup> 岡村らが挙げているのは、註 1 で挙げた Glaboniat/u.a. (2005)および以下の文献。Baldeger, Markus/Müller, Martin/Schneider, Günther (1980): Kontaktschwelle Deutsch als Fremdsprache, Berlin, Langenscheidt (Europarat, Rat für kulturelle Zusammenarbeit); Deutscher Volkshochschulverband/Goethe-Institut (1972): Das Zertifikat Deutsch als Fremdsprache. Bonn/Frankfurt/München. (2. Aufl. 1977; 3. Aufl. 1985).

<sup>7</sup> 岡村らが挙げているのは、以下の2つの文献。Haderlein, Veronika (2007): Das Konzept zentraler Wortschätze: Bestandsaufnahme, theoretisch-methodische Weiterführung und praktische Untersuchung. Dissertation, LMU München: Fakultät für Sprach- und Literaturwissenschaften. (<http://edoc.ub.uni-muenchen.de/8066/>); Schnörch, Ulrich (2002): Der zentrale Wortschatz des Deutschen. Strategien zu seiner Ermittlung, Analyse und lexikographischen Aufarbeitung, Studien zur Deutschen Sprache 26, Tübingen: Gunter Narr Verlag. なお、岡村らのシンポジウム趣意書の中で挙げられた文献ではないが、同じく German as Foreign Language Department of Langenscheidt (1991): Langenscheidt Basic German Vocabulary. Berlin/München: Langenscheidt. 等もここで言う辞書学的なアプローチで作られた語彙集の典型である。

<sup>8</sup> Council of Europe (2001): Common European Framework of Reference for Languages: Learning, teaching, assessment. Cambridge: Cambridge University Press.



こうした議論を背景に、改めて日本における基礎語彙集の歴史を振り返ってみると、Steger/Keil (1974)<sup>9</sup>の日本語版である『ドイツ語基本単語 2000』(日本放送協会 1974)を皮切りに、『おぼえるための基本ドイツ単語 1200』(朝日出版社 1975)、『アオスレーゼ現代ドイツ単語 2300』(朝日出版社 1983)、『ドイツ語キーワード 1000』(Newton Press 1989)、『聴いて、話すための—ドイツ語基本単語 2000』(語研 1991)、『英語から覚えるドイツ単語(2280 語)』(創拓社 1994)、『ドイツ語がわかる重要単語 850』(東洋出版 1995)、『独検突破単語集 3・4 級必修合格 1600』(三修社 1995)、『今すぐ話せるドイツ語単語集(1600 語)』(ナガセ 2002)、『効率よく覚えるドイツ重要単語 2200』(白水社 2009)、『キクタンドイツ語入門編：独検 5 級レベル聞いて覚えるドイツ語単語帳(456 語)』(アルク 2010)、『新・独検合格単語+熟語 1800』(第三書房 2010) 等、過去 30 年ほどに限ってみても十数種類の基礎語彙集が出ている<sup>10</sup>。これらの多くは、1) 場面や分野別の配列、2) 文法説明の観点からの配列、3) アルファベット順の配列等を採用しているが、語彙の選定基準については明らかにしていないものが多く、頻度順の配列方式を採用している『ドイツ語キーワード 1000』を除けば、そのほとんどは、既存の複数の学習用辞書から重要単語の共通部分を抜き出すか、部分的には場面や分野などを想定したコミュニカティブなアプローチに準拠しつつも、全体としては、編集者の直感や経験に依拠して作られたものが多い。こうした「恣意的な」選定方式に対する批判として、最近では言語コーパス等の経験的なデータに基づくいわゆる「科学的な」アプローチを求める声もあり、上述のシンポジウムもそうした議論の一環として企画されたものである。

### 3 日本のドイツ語教育における基礎語彙選定の際の頻度順アプローチの問題点

筆者は、日本のドイツ語教育における基礎語彙選定の基準を科学的に問い直そうという議論それ自体は、きわめて正当な主張であると考えている。また、今日においては、コーパス言語学的手法を抜きにした語彙リスト研究はあり得ないであろう。ただし、単にコーパスを使って頻度を出し上位語を集めれば、日本におけるドイツ語学習者用の基礎語彙集が作れるというほど問題は単純ではない。たとえば、経験的なデータに基づく「科学的な」アプローチのイメージとしては、BNC (British National Corpus) を基準スケールとして作成された英

<sup>9</sup> Steger, Hugo/Keil, Maria (1974): Der deutsche Mindestwortschatz 2000. Ismaning/München.

<sup>10</sup> このほかにも単語集としては、『例文活用 ドイツ重要単語 4000』(白水社 2003)、『日本語から引く知っておきたいドイツ語(10000 語)』(小学館 2003)、『ドイツ単語 5500』(郁文堂 2000)、『読んでおぼえるドイツ単語 3000—英語からドイツ語へ』(朝日出版社 1987)、『ドイツ基本語 5000』(白水社 1971)等が出ているが、これらは収録語彙のサイズが 3,000 語から 10,000 語で、もはや基礎語彙集とは言えないため、ここでは含めない。

語における『JACET 8000 英単語』(桐原書店 2005)<sup>11</sup>のような、語彙リストなどが想定される。しかし、英語との大きな違いは、ドイツ語には、英語の BNC や Bank of English のような当該言語を真に代表していると言える大規模な均衡コーパスが、まだ存在しないことである。これが 1 つ目の問題である。均衡コーパスとは、そのコーパスに含まれるテキストを選定する際に、テキストの種類やジャンルが偏らないよう、あらかじめ一定の科学的な選択基準に基づいて様々な分野のテキストを一定量ずつバランス良く集めて作成されたコーパスのことを言う。それに対し、ドイツ語のコーパスは、たとえば IDS (Institut für Deutsche Sprache) が作成したドイツ語圏で最も有名な大規模コーパス IDS-Korpora mit COSMAS II (2210 万語)<sup>12</sup>でさえも、それに含まれるテキストの多くは一部の地方新聞や、同好会サークルのフォーラムなどの偏った資料に限られ、様々なジャンルのテキストをバランス良く集めた均衡コーパスとは言えない。これは IDS にはコーパス研究用の予算はあっても、コーパスのコンテンツそれ自体の取得のための予算がほとんど無いため、そのコンテンツは、著作権処理にお金のかからない資料や、宣伝などの目的でデータを無償で提供してくれる地方新聞等の一部のものに偏らざるを得ないからである。他方、Herder/BY-Korpus は、ドイツ語圏で唯一と言っても良い自称「均衡コーパス」であり、“Grund- und Aufbauwortschatz Deutsch als Fremdsprache” (Tschirner 2008) 作成の基準となったものだが、数億語レベルの BNC 等と比べると、今日のコーパスとしては収録語数は 420 万語とやや小さい。また、その内訳は、均衡コーパスとは言うものの、1) ドイツ語圏の新聞、2) 専門書、専門雑誌から取った研究論文、3) エンターテイメント小説を含む文学作品、4) テレビドラマの会話を含む会話テキストから、それぞれ約 100 万語ずつ、また、使用説明書や解説書、法律書、契約書、ダイレクトメールや広告等の実用テキストから約 20 万語を取って作ったもの(Tschirner 2008, 4)であり、IDS のコーパスよりは資料に多様性があるものの、集めやすさのせい、やはり新聞や研究論文等の一部のジャンルの比重が高過ぎるように思われる。

2 つ目の問題は、日本における教育の実態や日本の文脈における言語教育という特殊事情の配慮に関わる問題である。『JACET 8000 英単語』の場合も、「BNC (British National Corpus) を基準スケールとし、日本人の英語学習者を反映した言語資料... (中略) ...から作成したサブコーパスをからめて作製」(相澤他編 2005, 2)したと、「はしがき」に書かれているように、単に BNC の頻度順リスト British National Corpus Frequency List<sup>13</sup> (作成者: Adam Kilgarriff) を用いただけ

<sup>11</sup> 相澤一美/石川 慎一郎/村田年 他 編著(2005): JACET 8000 英単語 - 「大学英語教育学会基本語リスト」に基づく。桐原書店。

<sup>12</sup> <http://www.ids-mannheim.de/cosmas2/projekt/referenz/dokutexte.html> (29.2.2012).

<sup>13</sup> <http://www.kilgarriff.co.uk/bnc-readme.html> (29.2.2012).

でなく、「日本国内の各種の英語試験や英語教材などを独自にデータベース化した」(相澤他編 2005, 10)様々なサブコーパス(Uemura / Ishikawa 2004, 333)<sup>14</sup>を基に、それぞれのサブコーパスごとにまず単語の頻度順リストを作り、それらの複数の語彙リストを比較対照しながら順位を決めている。『JACET 8000 英単語』の場合は、日本人学習者に必要だと思われる単語をリストに加えるという直接的な方法ではなく、あくまで「日本人の英語学習者を反映した言語資料」群をサブコーパスとし、対数尤度比による頻度順位の補正を行うという間接的な方法を取ってはいる。しかし、その場合も、どの資料をどの程度加えるかという度合いに応じて、日本における英語教育の特殊事情への配慮の仕方とその比重が変わり、したがってアウトプットとして出てくる語彙リストの順位と語彙も変わるという意味で、BNCのみから作られた Kilgarriff の語彙リストに大きな修正が加えられていることに変わりはなく、日本の中学校や高校の教科書等を電子化したサブコーパスから得られた語彙が最終的には、全体の3割程度を占めている<sup>15</sup>。

3つめの問題は、「コーパスで得られた頻度順で基礎語彙の範囲を決めて良いか」という問いと、「頻度順をそのまま学習すべき順序と考えて良いか」という問いの2つを混同することの危険性である。『JACET 8000 英単語』のように、そもそも日本における語学教育の現状を反映した調整がすでに行われていたり、あるいは4000語からなる Tschirner (2008)のように、語彙リストに含まれる単語数がある程度あり、かつ、「最終的にこれらをマスターしよう」という程度のゆるやかな指針として使うのであれば、頻度順を学習順と混同してもそれほど大きな問題は起こらない。というのも、前者の場合は、初学者に必要な語彙がある程度上位に集められており、そうした調整の無い後者の場合も、4000語の規模になると、重要な単語は、頻度順での並び方は違っても全体としてある程度網羅されるからである。しかしそうした調整や配慮もないまま語彙リストの規模が数百あるいは千以下などと比較的小さくなっていくと話はまったく別である。というのも、コーパスの頻度順では、内容語よりも機能語の、また内容語の中でも名詞よりは動詞(特に機能動詞)の頻度がそれぞれ高くなる傾向があるため、基礎語彙に含まれる単語数が小さいと、名詞では本来必要な基本語の大部分が漏れたり、逆に機能語や動詞などでは最初に学ぶべき語としては重要性の低いものが大量に紛れ込んだりして、特定の品詞に語彙が偏る弊害が大きくなるからである。これらの点については、本稿の第4節で、冒頭に述べた3つの語彙リストの比較を通して具体的に述べるが、過去に日本で出版された基礎語彙集の多くが、せいぜい2000語以内であったことが示すように、日本にお

<sup>14</sup> Uemura Toshihiko/Ishikawa Shin'ichiro (2004): JACET 8000 and Asia TEFL Vocabulary Initiative. *The Journal of Asia TEFL (The Asian Association of TEFL)* 1(1). 333-347.

<sup>15</sup> これは、編集者の1人から筆者が直接聞いた話による。

けるドイツ語教育の大半を占める大学や高等専門学校の教養教育における第2外国語教育においては、平均的な学習時間は、1年間（約90時間＝90分授業×週2回×15週間×2 Semester）に過ぎず、そこで考えられる現実的な基礎語彙サイズは、1000語（±200語）程度である。上述したように、こうした限定された基礎語彙サイズで頻度順を機械的に適用すると、動詞はある程度含まれるものの、それとともに使うべき名詞はほとんど基礎語彙から抜け落ちてしまい、日常的な口頭表現の大部分がまったく不可能になってしまう。したがって、コーパスを使った頻度順による語彙選定を試みる場合にも、出てきたデータの上位語をそのまま日本のドイツ語教育における基礎語彙にできるわけではなく、コーパスを使った言語研究の場合と同様、そのデータを人間（研究者あるいは教員）が分析し、それを目的との関係で評価して使う作業（具体的には、複数の語彙リスト間の比較対照、単語の取捨選択、順位の入替え、間接的あるいは直接的な単語の補充等の加工作業）が不可欠である。この点を曖昧にしていると、「基礎語彙」としては使い物にならない現実離れしたデータが一人歩きするか、あるいは、出てきたデータを手直しする過程で再び無自覚な形で「恣意的な」直感が持ち込まれることになりかねない。では、機械による処理によって得られたデータを人間が分析し評価する際の基準とは何であろうか。筆者は、この段階の議論においては、学習者が日常生活のどの場面で、どのような目的で、ドイツ語をどのように用いるのかという言語機能の側から個々の単語の重要性を評価するコミュニケーションなアプローチが、基礎語彙選定へのもう一つのアプローチとして不可避になると考える。具体的な学習者を想定して考える限り、たとえ頻度順のアプローチから出発した場合でも、最終的にはいずれかの時点でコミュニケーションな観点を補完的な判定基準として取り入れることは不可欠なのではないだろうか。

ただし日本のドイツ語教育においてコミュニケーションの主体となる具体的な学習者とは、Glaboniat/u.a. (2005)において想定されているような、ドイツ語圏で生活すること、あるいはすでに生活していることを前提にドイツ語を長期的な視点で学んでいく一般市民ではない。その大部分は、ドイツ語圏から遠く離れた日本の大学の教養教育の枠組みの中でまずは1年間という期限を切ってドイツ語を学び始める日本人の大学生である。したがって、こうした日本におけるドイツ語学習の文脈を踏まえ、学習の主たる対象者である日本人学生の自己表現の要求や知的な関心を考慮し、彼らの生活空間に根ざした表現が可能な形で初級段階の語彙を考える必要がある。効率的な語彙学習という観点に限って言っても、単語の記憶・定着を促進する要因は、どれだけ自分と関連づけら

れるかという処理の深さ(Craik/ Lockhart 1972)<sup>16</sup>である。2005年に広島大学で筆者らが基礎語彙の作成を委託されたとき、CEFRに準拠した Glaboniat/u.a. (2005)をそのまま採用するという方針を採らず、何らかの形で日本のドイツ語教育の現状を反映しようと試みた理由もそこにあった。

#### 4 広島大学で作られた基礎語彙リストと追加語彙リスト

本稿冒頭の註でも述べたが、ここで取りあげる「広大語彙リスト」の原型は、2005年から2007年にかけて、広島大学の岩崎克己、吉満たか子、Axel Hartingの3人が、教養教育の枠組みでドイツ語を学ぶ大学1年生全体を対象として作成した、基礎語彙(Grundwortschatz)約800語と追加語彙(Aufbauwortschatz)約350語から成るドイツ語の語彙リストである。これは、広島大学の1年生向けの初修外国語授業である週2回(年間90時間)の「ベーシック・外国語」授業の到達目標の明確化の一環として、広島大学外国語教育研究センター・カリキュラム実施専門部会の委託を受け、初修外国語の各言語(ドイツ語・フランス語・スペイン語・ロシア語・中国語・韓国語)で作成された基語彙リストのドイツ語版である。

当時、岩崎、吉満、Hartingらの3人が、作成に当たり採用した方針は、まず第1にSD1やSD2(Start Deutsch 1/2)等の国際的なドイツ語の資格試験を意識しA1/A2レベルの単語をある程度カバーすることであった。しかしまた同時に、前節末でも述べた理由から、第2として、日本におけるドイツ語学習の文脈を考慮し、日本の大学で初めてドイツ語を学ぶ日本人大学生が自己表現に必要なと思われる語彙や彼らの大学生としての日常生活に関連する語彙、さらには彼らが日本文化との対照を通じてドイツ語圏の文化を学ぶために必要な語彙にも重点を置いた<sup>17</sup>。

実際の語彙選定に当たっては、まず、以下の資料を準備した<sup>18</sup>。

---

<sup>16</sup> Craik, Fergus/Lockhart, Robert (1972): Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior* 11, 671-684. なお、語彙の深さという観点については、以下も参照。

門田修平編著(2003): 英語のメンタルレキシコン- 語彙の獲得・処理・学習 -. 松柏社; 望月正道/相澤一美/投野由紀夫(2003): 英語語彙の指導マニュアル. 大修館書店; Nation, I. S. P. (2001): *Learning Vocabulary in Another Language*. Cambridge Applied Linguistics: Cambridge University Press.

<sup>17</sup> 日本のドイツ語学習の文脈において語彙を考えると、「効率的な文法学習に役立つ語彙」という観点も考慮されるべきであるが、初級文法の説明に適した語彙や用例の大部分は、ここで基本方針に挙げた語の範囲内でほとんどカバーされていたので、文法学習の観点を語彙選定の基準としてわざわざ取りあげる必要性は、実際にはほとんど無かった。

<sup>18</sup> 直接利用したわけではないが、先行事例として北大で作られていた以下の語彙リストも参照した。石川克知/佐藤俊一編(1996): 言語文化部研究報告叢書8<特集>ドイツ語基礎語彙研究, 北海道大学。

- 1) コミュニカティブ・アプローチに基づく「Glaboniat/u.a. (2005) の語彙リスト」において受容能力レベルが A2 までの単語（動詞・名詞・形容詞）約 1700 語のリスト
- 2) 『自己表現のためのドイツ語 (*Farbkasten Deutsch neu 1*)』(三修社 2003)、『場面で学ぶドイツ語 (*Szenen 1*)』(三修社 2006)、『コミュニケーションのためのドイツ語 (*Deutsch für Kommunikation*)』(三修社 1996)等の「日本人大学生の自己表現」に重点を置いていたいくつかの代表的な教科書に使われていた語彙を中心に集めた約 950 語のリスト<sup>19</sup>
- 3) ドイツ語技能検定試験 3 級・4 級に頻出する単語をまとめた約 1800 語のリスト<sup>20</sup>

1) の「Glaboniat/u.a. (2005) の語彙リスト」は、ドイツ語圏におけるコミュニケーションなアプローチを純粹に反映したリストであり、2) および 3) は日本のドイツ語学習者の文脈におけるコミュニケーションなアプローチと辞書学的なアプローチの両方が混在する形で作られたリストであった。実際の語彙選定手順としては、これらの 3 つのリストを対照し、それらすべてに含まれる単語、それぞれ 2 つずつのリストに含まれる単語、1 つのリストにしか含まれない単語という順で、各単語を並べた基礎資料を基に、基本方針の第 2 に挙げた日本人大学生に必要なと思われる語彙という観点から、3 人の合議で約 800 語をリスト化し、残ったものからさらに 350 語を選んで追加語彙とした。作成した語彙は、動詞、名詞、形容詞、およびその他という 4 つの品詞カテゴリーに分け、それぞれをアルファベット順に整理した形式と、コミュニケーションの場面や分野別に 19 種類に分類<sup>21</sup>した形式との、2 つの形でまとめた。また、それとは別に、約 50 個の重要な慣用表現 (*wichtige Redewendungen*) といわゆる教室ドイツ語 (*Klassendeutsch*) をまとめた。

当時、各言語ごとに作られた語彙リストは、2007 年度末の広島大学外国語教育研究センター・外国語企画会議への答申後、公的な形では公表されなかった

<sup>19</sup> *Farbkasten Deutsch neu 1* と *Szenen 1* の単語を取りあげた理由は、それらが日本人学習者の自己表現に重点を置いたドイツ語初級用教科書の中で当時最も売っていた教科書だからである。また故関口一郎氏が作られた *Deutsch für Kommunikation* の方は、それが、コミュニケーションに重点を置いたドイツ語初級用教科書の中では最初期のものであり、かつその後の同種の教科書のスタイルに多大な影響を与えたものだからである。

<sup>20</sup> これは、飯島一泰/清水朗編著 (1996): 独検合格らくらく 30 日. 郁文堂. の巻末の単語集、およびドイツ語技能検定試験 4 級の 2000 年～2005 年の出題問題に使用されていた単語の集計表を基に作った。

<sup>21</sup> 分類の枠組みは以下の 19 個。1. あいさつ、2. 人物・性質、3. 家族・交友、4. 自由時間・余暇、5. 買い物、6. 住まい、7. 日課・時間、8. 休暇、9. 街・屋外、10. 体・健康、11. 職場・学校、12. 衣服、13. 飲食、14. 教室活動、15. 思考・感情、16. 行為・活動、17. 環境・自然、18. 日常・所有、19. その他。

ため、この「広大語彙リスト」(2007年版)も、ドイツ語担当教員に試案として提示・配布されただけで、その利用は各人の自由に任された。作成者の一人である岩崎は、個人的にインターネット上で閲覧可能<sup>22</sup>にし、またこの間、自分の授業で使いながら不備を直し、2008年には必要と思われる語彙を50語程度補った改訂版<sup>23</sup>を作り利用してきたが、大学全体での扱いとしては、事実上店ざらし状態であった。なお、作成者自身が、個人的には利用しながらも、それを一般的な基準として広く主張することをためらった理由の1つは、これらのリストの妥当性や科学性の根拠に対して確固とした自信が無かったからでもある。本稿で、この「広大語彙リスト」(改訂版)と Glaboniat/u.a. (2005) や Tschirner (2008)の語彙リストとの比較を試みるのも、当時の自分たちの作業にどれだけの妥当性があったのかを別の尺度で見直してみたいという問題意識によるところが大きい。その際、「広大語彙リスト」と「Tschirner (2008)の語彙リスト」との比較は、頻度順アプローチという、当時筆者らが考慮していなかった全く異なる語彙選定のアプローチから見ると何が言えるかを考えるためであり、「Glaboniat/u.a. (2005)の語彙リスト」との比較は、同じくコミュニカティブなアプローチに基づきつつも、ドイツ語圏におけるドイツ語学習の文脈と日本におけるドイツ語学習の文脈の違いについて考えるためである。

## 5 3つのリストの比較

前節で「広大語彙リスト」の選定基準として、日本の大学で初めてドイツ語を学ぶ日本人学生に必要なと思われる語彙という曖昧な表現を用いたが、本節では、その内容をさしあたり次のように具体化し、それらに必要な語彙がそれぞれのリストにどの程度含まれているかを見ていくことにする。

- (1) 日本に住む大学生が、出身、住所、大学生活、言語、飲食の好み、ファッション、スポーツ、音楽、趣味、家族・友人等の情報を含む自己紹介ができ、
- (2) 簡単な日常的な状況記述(身の回りの持ち物や自分の部屋の状況についての説明、屋内あるいは周辺地域における空間的な記述等)ができ、
- (3) 価値意識、習慣、色彩感覚や美観等における日独文化の違いを大まかに説明でき、それらに対する判断等についても適切な形容詞を使って表現でき、

<sup>22</sup> <http://home.hiroshima-u.ac.jp/katsuiwa/grundwortschatz.htm> (29.2.2012).

<sup>23</sup> 2008年の改訂版では、元のリストの中の120語ほどの単語の扱いが変わっている。詳しい変更点は以下のURL参照。<http://home.hiroshima-u.ac.jp/katsuiwa/wortschatz/01102008.pdf> (29.2.2012).

- (4) 適切な話法の助動詞を用いて自分の技能や感情や要求や望み、必然性や可能性についての判断等を表現でき、
- (5) 日常の自分の生活（学生生活や休暇や1日の過ごし方等）について概略を時系列で説明でき、
- (6) 直近の過去に体験したこと等について簡単に報告することができる。

まず、「広大語彙リスト」の基礎語彙約 850 語の中からこれらの領域で必要とされるであろう単語のうちの代表的なものを具体的に取り挙げ、それに Glaboniat/u.a. (2005) における能力レベルの情報（受容レベルと産出レベルの順にハイフンでつないだもの）と Tschirner (2008) における頻度順位の情報を加えて列挙する。なお、すでに述べたように、「広大語彙リスト」（改訂版）の基礎語彙は約 850 語で、追加語彙 350 語も含めると約 1200 語になる。3つのリストはそれぞれの尺度が違いうえ、基礎的な単位となる語や表現のとらえ方も違う<sup>24</sup>ので、あくまで大雑把な目安でしかないが、「Glaboniat/u.a. (2005)の語彙リスト」における受容能力レベルが A1（A1-A1, A1-A2, A1-B1 等）までの語の数も約 850 語で、受容・産出の両方の能力レベルが A2（A2-A2）までの語の数を含めると同じく約 1200 語となるので、この目安は、Tschirner (2008)のそれぞれ上位 850 語と 1200 語という基準とともに3つの語彙リストのどの部分に特定の語が含まれているか、あるいは含まれていないかを考える際の指標となる。また、同じく、リストの語彙数から大雑把に見ると Glaboniat/u.a. (2005)における B1-B1（受容・産出いずれのレベルも B1）はおおよそ 2000 語レベル以上の語彙に相当する。

#### ・動詞・助動詞

まず、動詞・助動詞に関して言えば、自己発信、日常的な状況記述、価値判断等の表現に必要な動詞・助動詞の多くは、3つのリストともに上位 850 語の範囲に含まれている。

heißen（～と称する: A1-A1, 123 位）、kommen（来る: A1-A1, 61 位）、wohnen（～に住む: A1-A1, 382 位）、lernen（学ぶ: A1-A2, 203 位）、studieren（専門的に学ぶ: A1-A2, 438 位）、arbeiten（働く: A1-A1, 200 位）、trinken（飲む: A1-A1, 610 位）、essen（食べる: A1-A1, 657 位）、machen（する: A1-A1, 49 位）、sprechen（話す: A1-A1, 157 位）、spielen（球技等をする・[楽器を]演奏する: A1-A1, 197 位）、sehen（見

<sup>24</sup> たとえば besetzt は「広大語彙リスト」や Glaboniat/u.a. (2005)では形容詞であるが、Tschirner (2008)では besetzen という動詞の一変化形として扱われている。また、Glaboniat/u.a. (2005)では、同じ語でも用法ごとに別の能力レベルが付与されていたり、基本単位に単語だけではなく一部熟語が含まれたりしている。



る: A1-A1, 81 位)、lesen (読む: A1-A1, 325 位)、nehmen (取る: A1-A1, 139 位)、haben (持っている: A1-A1, 7 位)、sein (～である: A1-A1, 3 位)、bleiben (留まる: A1-A1, 112 位)、fahren (乗り物で行く: A1-A1, 169 位)、gehen (歩いていく: A1-A1, 69 位)、warten (待つ: A1-A1, 390 位)、treffen (会う: A2-A2, 289 位)、besuchen (訪れる: A2-A2, 705 位)、suchen (捜す: A1-A2, 295 位)、brauchen (必要とする: A1-A2, 201 位)、kaufen (買う: A1-A1, 583 位)、bekommen (もらう: A1-A2, 235 位)、finden (～を～と思う: A1-A2, 110 位)、gefallen ([人の] 気に入る: A1-A1, 538 位)、helfen (助ける: A1-A1, 408 位)、geben (与える: A1-A1, 57 位)、schreiben (書く: A1-A1, 247 位)、schicken (送る: A1-A1, 953 位)、kennen (知っている: A1-A1, 181 位)、wissen (知っている: A1-A1, 79 位)、vergessen (忘れる: A2-A2, 597 位)、glauben (信ずる・思う: A1-A2, 143 位)、denken (思う: A1-A2, 124 位)、sagen (言う: A1-A1, 46 位)、fragen (尋ねる: A1-A1, 175 位)、liegen (ある・横になっている: A1-A1, 118 位)、legen (置く・横たえる: A2-A2, 298 位)、stehen (ある・立っている: A1-A1, 87 位)、stellen (置く・立てる: A2-A2, 192 位)、sitzen (座っている: A1-A1, 263 位)、setzen (据える: A2-A2, 229 位)、hängen (掛かっている/掛ける: A2-A2, 690 位)、bringen (持ってくる: A1-A2, 162 位)、verstehen (理解する: A1-A1, 223 位)、versuchen (試みる: A1-A1, 224 位)、ankommen (到着する: A1-A1, 716 位)、*aufstehen* (起床する: A1-A2, 1176 位)、zurückkommen (戻って来る: A2-B1<sup>25</sup>, 1113 位)、können (～できる: A1-A1, 23 位)、wollen (～するつもりである: A1-A2, 65 位)、dürfen (～しても良い: A1-A2, 142 位)、müssen (～しなければならない: A1-A1, 45 位)、sollen (～すべきである: A1-A2, 63 位)、mögen (～かもしれない: A1-A1, 151 位)

上記の動詞のうち、下線の付いている 8 つの動詞とイタリックになっている 2 つの動詞は、それぞれ、「Glaboniat/u.a. (2005)の語彙リスト」と「Tschirner (2008)の語彙リスト」においてのみ、1200 語レベルには含まれるが 850 語レベルには含まれない例である。しかし、それ以外の動詞は、3 つのリストに共通して上位 850 語以内に含まれている。このことからわかるように、動詞・助動詞に関しては、どのアプローチを採ろうとも、必要な動詞の大多数はカバーされると言える。

ただし、「Tschirner (2008)の語彙リスト」については、明らかに重要動詞だと思われるのに、上位 1200 語に入らない例がいくつか見られる。たとえば、学生が自分の趣味について語る際、その多くは音楽鑑賞を挙げるがその際に必要な hören (聴く: A1-A1, 1560 位) や、趣味に関する話題で 1 クラスに必ず何人かは使う malen (絵を描く: A1-A2, 1859 位) 等がそれである。また携帯電話は今や

<sup>25</sup> 正確に言うと、zurückkommen は「Glaboniat/u.a. (2005)の語彙リスト」の中には直接含まれていないが、前綴り zurück-の能力レベルの記述として A2-B1 が登録されている。

現代人の必需品であるが、それを使って行う *anrufen* (電話を掛ける: A1-A2, 1388 位) や *telefonieren* (電話で話す: A1-A1, 3357 位) も、また *waschen* (洗う: A1-A1, 1783 位)、*mitbringen* (持って行く: A1-A1, 1802 位) 等の日常的な動詞も上位 1200 語には含まれていない。さらに、*abfahren* (出発する: A1-A2, x) や *jobben* (バイトする: x, x) に至っては、そもそも 4000 語のリストにさえ入らない。これらは、*jobben* を除けば、「広大語彙リスト」でも「Glaboniat/u.a. (2005)の語彙リスト」でもともに上位 850 語に含まれているので、その違いが目立つ例である。他にも、日本人の学生が趣味などの表現でよく使う、*zeichnen* (イラスト等を描く: B1-B1, 1811 位)、*sammeln* (収集する: B1-B1, 1326 位)、*angeln* (釣りをする: B1-B2, x) や日課の表現などで使われる *aufräumen* (片付ける: B1-B1, 3963 位) 等も、「Tschirner(2008)の語彙リスト」では上位 1200 語には入らない<sup>26</sup>。

それに対し、「Tschirner(2008)の語彙リスト」の上位語には、初級者用の語彙と言われると、多くの教員が違和感を抱くであろう動詞が多く含まれている。それらの一部を試しに上位 1000 語の範囲内で 20 個ほどリストアップしてみると、以下のものがある。

*gelten*(x, 185 位), *bestehen*(B1-B2, 210 位), *betreffen*(x, 388 位), *vergehen*(x, 400 位), *ergeben*(x, 432 位), *erfolgen*(x, 570 位), *vorliegen*(x, 681 位), *erhöhen*(B2-B2, 731 位), *vorhanden*(x, 733 位), *beziehen*(x, 718 位), *beteiligen*(x, 770 位), *umfassen*(x, 785 位), *aufbauen*(x, 832 位), *überzeugen*(x, 872 位), *drohen*(x, 875 位), *betreiben*(x, 928 位), *vertreten*(x, 939 位), *benötigen*(x, 968 位), *verhindern*(x, 973 位), *erwähnen*(x, 985 位)。

上記の動詞群は、*bestehen* と *erhöhen* を除くと括弧内の具体的な能力レベルの記述がx印であることからわかるように、Glaboniat/u.a. (2005)の判断でもそのほとんどが B2 レベルにさえ含まれない。上記の動詞は、主に論文や専門書、新聞などの書き言葉の論説文や法律の条文に多く出てくる動詞である。すでに述べたように、日本のドイツ語教育の大半を占める大学や高等専門学校の教養教育における学習時間は、1 年間 90 時間程度であり、そこで考えられる基礎語彙数は、1000 語 (±200 語) 程度である。このレベルの学習者が読めるとはとうてい考えられない新聞や専門書が大半を占めるコーパスの頻度順で最初の 1000 語レベルの基礎語彙を機械的に決めようとしたら何が起こるかをこの例は如実に示している。

---

<sup>26</sup> ただし、これら 3 つの動詞は、「Glaboniat/u.a. (2005)の語彙リスト」の能力記述でも B1-B1 レベルであり、したがって 2000 語レベルの語彙である。その逆に、「Glaboniat/u.a. (2005)の語彙リスト」では 850 語レベルだが、「広大語彙リスト」には含まれない動詞には *heizen* (暖房する: A1-A2, x)、*küssen* (キスする: A1-A2, 2629 位) 等がある。ただし、これら以外では、すでに述べたように両者の上位語には大きなずれはない。

・名詞

しかしながら、頻度順アプローチの問題点が最も顕著に浮かび上がるのは、動詞ではなく、むしろ名詞の場合である。以下では、「広大語彙リスト」に含まれている名詞を取りあげ具体的に検討していくが、日本人の学生が、家族関係の表現で必要とする Großvater(祖父: A1-A1, 3642 位)、Großmutter(祖母: A1-A1, 2834 位)、Onkel(おじ: A2-B1, 2852 位)、Tante(おば: A2-B1, 1663 位)、Geschwister(兄弟姉妹: B1-B2, 3094 位)などの多くは、「Tschirner(2008)の語彙リスト」では上位 2500 語にさえ入らない。飲食物を表す名詞は上位 1200 語内には 1 つもなく、そもそも含まれているのは、Kaffee(コーヒー: A1-A1<sup>27</sup>, 1423 位)、Fisch(魚: A1-A1, 1626 位)、Fleisch(肉: A1-A1, 1627 位)、Wein(ワイン: A2-B1, 1643 位)、Bier(ビール: A2-B1, 1716 位)、Brot(パン: A1-A1, 1760 位)、Ei(卵: A1-A1, 2608 位)、Tee(紅茶: A1-A1, 2717 位)、Milch(牛乳: A1-A1, 2789 位)、Käse(チーズ: A1-A1, 3615 位)、Apfel(りんご: A1-A1, 3842 位)、Zwiebel(タマネギ: A2-B1, 3387 位)、Kartoffeln(じゃがいも: A1-A1, 3314 位)のみである。しかもその半数は 2500 語以降である。Salat(サラダ: A1-A1)、Reis(米・ご飯: A1-A1)は、「Tschirner(2008)の語彙リスト」には登場しない。

学生の中には何らかの楽器のできるものが多いが、楽器でリストに入っているのは、Klavier(ピアノ: Δ<sup>28</sup>, 2779 位)のみで、それ以外は、Gitarre(ギター: -)や Geige(バイオリン: -)を含めすべて含まれていない。スポーツをする学生は多いが、Fußball(サッカー: A1-A2, 1500 位)、Tennis(テニス: A2-B1, 2952 位)以外は、スポーツの名称はそもそも含まれていない。若い世代が特に気を遣うファッションや身の回りの持ち物についても、Kleid(ワンピース/衣類: A2-A2, 1684 位)、Schuhe(靴: A1-A1, 2117 位)、Hose(ズボン: A1-A1, 2426 位)、Hemd(シャツ: A1-A1, 2537 位)、Mantel(コート: A1-A1, 2548 位)、Hut(帽子: A2-A2, 2693 位)、Rock(スカート: A2-A2, 3230 位)、Anzug(スーツ: B1-B1, 3544 位)、Jacke(上着: A1-A1, 3895 位)とほとんどすべて 2000 語以降であり、Jeans(ジーンズ: A1-A1)、Socken(ソックス: A2-A2)、Strümpfe(ストッキング: A2-B1)はそもそもリストに登場しない。持ち物についても、Uhr(時計: A1-A1, 351 位)以外は、1500 語以降に Tasche(かばん: A1-A2, 1641 位)、Schlüssel(カギ: A1-A1,

<sup>27</sup> 「Glaboniat リスト」の語彙リストには冷たい飲み物には個々に能力レベルの記載があるが、温かい飲み物に関しては、Heiße Getränke というカテゴリー情報のみで、具体的な能力レベルの記載はない。しかし、関連する語彙のレベル記述から、少なくとも、Kaffee と Tee に関しては、A1-A1 と見なしてかまわないと思われる。同様に、「Glaboniat リスト」の語彙リストには Ei は含まれていないが、これは、単純な記載ミスだと思われる。Fisch/ Fleisch/ Käse 等の基本食品がいずれも A1-A1 であることから Ei も A1-A1 と見なした。

<sup>28</sup> 「Glaboniat リスト」の語彙リストでは楽器についても Klavier/ Gitarre/ Geige など、9 個ほど登録されているが、Musikinstrument というカテゴリー情報のみで、具体的な能力レベルの記載はない。

2325 位)、Handy (携帯電話: A2-A2, 2470 位)、Brille (メガネ: A2-A2, 3428 位)、Heft (ノート: A1-A1, 3603 位) がかりうじて入っているだけで、Taschentuch (ハンカチ: B1-B2)、Portmonee (財布: A2-A2)、Lehrbuch (教科書: -)、Wörterbuch (辞書: A1-A2)、Stift (筆記用具/ペン: A1-A2)、Bleistift (鉛筆: A1-A2)、Kugelschreiber (ボールペン: A1-A2)、Studentenausweis (学生証: -)、Regenschirm (傘: -)、などの学生の持ち物の多くは、4000 語のリストにさえ含まれていない。日本人学生が自宅やアパートの自分の部屋に所有しているものについても、上位 1000 語に入っているのは、Tisch (机: A1-A1, 496 位) と Bett (机: A1-A1, 656 位) と Computer (コンピュータ: A1-A1, 741 位) のみで、その他は Stuhl (いす: A1-A1, 1536 位)、Fernseher (テレビ: A1-A1, 2419 位)、Regal (本棚: A2-A2, 2644 位)、Schrank (タンス: A1-A1, 3232 位)、Lampe (スタンド: A2-A2, 3899 位) と、軒並み順位は低く、Waschmaschine (洗濯機: A2-A2) や Kühlschrank (冷蔵庫: A1-A1) はリスト落ちである。道案内などで大事な情報は、方向を指示する表現<sup>29</sup>と案内の目安となる交差点や建物や店舗名などであるが、それについても、上位 1000 語に入っているのは、Bank (銀行: A1-A1, 511 位) のみで、それ以外は、Flughafen (空港: A1-A2, 1935 位)、Post (郵便局: A1-A1, 1944 位)、Bahnhof (駅: A1-A2, 1956 位) Café (カフェ: A1-A1, 2349 位) とほぼ 2000 語レベルであり、Supermarkt (スーパー: A1-A1)、Kaufhaus (デパート: B1-B1)、Buchhandlung (本屋: B1-B1)、Bäckerei (パン屋: B1-B2)、Kreuzung (交差点: A1-A1)、などは 4000 語のリストに含まれない。

このように、Tschirner (2008) のようなドイツ語圏で現在利用可能なコーパスを用いて作った語彙リストの頻度順をそのまま適応して、その上位語を日本のドイツ語教育における初級者用の基礎語彙 (1000 語程度) としようとするれば、核となる動詞はあってもそれと使える名詞はほとんど含まれないため、学習者は、核家族の範囲内でしか家族について語れず、自分の趣味については、「読書」以外何も挙げられず、飲食物の好み、スポーツについては何 1 つ言えず、自分の持ち物については「時計」のみ、自分の部屋の中にあるものについては、「机」と [ベッド] と「コンピュータ」について触れられるだけ、道案内では、方向を表す表現はいっさい使えず、目安になる建物は「銀行」だけという悲惨な結果になってしまう。

ただし、これは、すでに第 2 節でも触れたように、コーパスの頻度順という指標特有の性質から必然的に生じる現象でもある。そのことを、たとえば spielen (A1-A1, 197 位) という動詞を例に取って考えてみよう。インターネット上ですこし捜せば、以下のような spielen を含む例文は次々と見つかる。

<sup>29</sup> 方向を表す副詞も Tschirner (2008) では上位 1200 語には入らず、links (左へ: A1-A1, 1446 位)、rechts (右へ: A1-A1, 1449 位) はなんとか出てくるが、geradeaus (まっすぐに: A1-A1) はそもそもリストにない。

Tim spielt Fußball.  
Tennis kann ich bis ins hohe Alter spielen.  
Lasst uns Baseball spielen!  
In seiner Freizeit spielt Harry Redknapp gerne Golf.  
Wer kann Geige spielen?  
Wie gut spielst du Gitarre?  
Er spielt gern mit dem Hund.  
Mein Großvater spielt Schach.  
Meine Mutter spielt mit dem Gedanken, in ein Altersheim zu ziehen.  
Mein Neffe spielt gerne mal bei uns am Computer.  
Nein, das spielt keine Rolle.

ここだけで、spielen、は延べ数で 11 回使われているが、他の名詞は全て 1 回だけである。したがってこの例文から得られる使用頻度では spielen のみが上位に入り、他の名詞の重要性は統計的な数値としては表れない。しかし、忘れてはならないのは、spielen だけが単独で出てくることはほとんどなく、これらのどれかの名詞（正確には名詞句やその名詞を含む前置詞句）と一緒に使われない限り、実際には spielen 自体も使えないことである。したがって、語彙リストの範囲が、これらの共起する名詞の順位までカウントされるほど十分に拡大されない限り、言語の使用実態は見えてこないし、学習の目安となる語彙リストとしては意味をなさない。その拡大範囲が、どの程度になるかは一概に言えないが、上記の例で使われている名詞 17 個のうち、「Tschirner (2008)の語彙リスト」に含まれるのは、Fußball (1500 位)、Tennis (2952 位)、Alter (722 位)、Freizeit (1156 位)、Hund (1048 位)、Mutter(228 位)、Großvater(3462 位)、Gedanke (531 位)、Computer (2952 位)、Rolle (398 位) の 10 個のみであり、これらは、3000 語の範囲でカバーできるが、残りの 7 つはそもそも 4000 語の範囲に入らない。そのことから考えても、言語教育のための基礎資料という観点で言うと、おそらくもう少し大きな規模が必要になるであろう。その意味で、コーパスを使った頻度順語彙リストを 1000 語ずつ 8000 語レベルまで分類した『JACET 8000 英単語』の試みは、頻度順リストの仕様やサイズを考えるうえで示唆的である。ただし出現回数における差が有意となるだけの数のデータをその順位まで集めようとするれば、少なくとも Herder/ BY-Korpus よりは大きな「均衡コーパス」が必要になるであろう。

なお、これまで述べてきた例とは逆に、動詞の場合と同様、「Tschirner (2008)の語彙リスト」の上位語には、初級者用の語彙と言われると、多くの教員が違和感を抱くであろう抽象的な名詞が非常に多い。紙幅の関係で、ここでも全ては挙げられないので、その一部を上位 1000 語の範囲内で 20 個ほどリストアッ

プしてみる。これらも、政治経済や学問研究の分野の論文・論説文などで多用されそうな表現である。

Grund (根拠: B1-B1, 231 位)、Bereich (領域: B1-B1, 259 位)、Entwicklung (発展: ×, 305 位)、Abbildung (図: ×, 441 位)、Rahmen (枠組み: ×, 462 位)、Regel (規則: B1-B2, 486 位)、Beziehung (関係: B1-B1, 494 位)、Verhältnis (状況: B2-B2, 519 位)、Zusammenhang (関連: ×, 553 位)、Angebot (申し出: B1-B2, 586 位)、Anspruch (請求: ×, 709 位)、Anwendung (適用: ×, 717 位)、Verfahren (手続き: ×, 745 位)、Maßnahme (処置: ×, 757 位)、Bevölkerung (人口: B1-B1, 771 位)、Bedingung (条件: B1-B1, 806 位)、Untersuchung (調査: B1-B1, 830 位)、Schwierigkeit (困難: B1-B2, 868 位)、Grundlage (基礎: ×, 971 位)、Zustand (状態: ×, 975 位)

以上、主に「広大語彙リスト」と「Tschirner (2008)の語彙リスト」の比較を行ってきたが、次に、「広大語彙リスト」と「Glaboniat/u.a. (2005)の語彙リスト」の中の名詞の重なり具合を見てみよう。「広大語彙リスト」850 語に含まれる名詞 455 個のうち「Glaboniat/u.a. (2005)の語彙リスト」の上位 850 語程度に相当するであろうと思われる、受容レベル A1 の名詞は 297 個 (約 65.4%) である。また受容レベル A2 の名詞も 127 個 (28.0%) であり、両者を合わせると 93% を超え、全体として語彙レベルがかなり重なっているのがわかる。それに対し、上位約 2000 語以降に相当すると思われる、受容レベルが B1 以上の名詞は、455 個中、以下にあげる 30 個 (約 6.7%) しか無い。

Angestellter (会社員: B1-B1, 2344 位)、Anzug (スーツ: B1-B1, 3544 位)、Beamter (公務員: B1-B1, 1891 位)、Buchhandlung (書店: B1-B1, ×)、Dame (婦人: B1-B1, 1353 位)、Hals (首・のど: B1-B1, 1823 位)、Kaufhaus (デパート: B1-B1, ×)、Neujahr (新年: B1-B1, ×)、Prüfung (試験: B1-B1, 1208 位)、Schauspieler/-in (俳優: B1-B1, 2205 位)、Taschentuch (ハンカチ: B1-B2, ×)、Zeitschrift (雑誌: B1-B2, 2496 位)、Bäckerei (パン屋: B1-B2, ×)、Geschichte (歴史学・お話: B1-B2, 275 位)、Geschwister (兄弟姉妹: B1-B2, 3094 位)、Kellner / -in (ウェイター: B1-B2, ×)、Rathaus (市役所: B1-B2, 3122 位)、Roman (小説: B1-B2, 1038 位)、Brust (胸: B2-B2, 1676 位)、Federmäppchen (筆入れ: B2-B2, ×)、Schal (マフラー: B2-B2, ×)、Naturwissenschaft (自然科学: B2-×, 3777 位)、Ski (スキー: B2-×, ×)、Lehrbuch (教科書: ×, ×)、Radiergummi (消しゴム: ×, ×)、Regenschirm (雨傘: ×, ×)、Schulter (肩: ×, 1065 位)、Studentenausweis (学生証: ×, ×)、Tempel (寺・寺院: ×, ×)、Schinto-Schrein (神社: ×, ×)

その逆に、「Glaboniat/u.a. (2005)の語彙リスト」では 850 語レベルだが、「広大語彙リスト」には含まれない名詞には、以下のものがある。これらは、食品や香辛料名あるいは、実際にドイツ語圏で生活する際には必要になる概念を表す名詞などである。

Rindfleisch (牛肉: A1-B1, ×)、Kalbfleisch (仔牛の肉: A1-B1, ×)、Konfitüre (粒入りジャム: A1-B1, ×)、Pfeffer (胡椒: A1-A1, ×)、Essig (酢: A1-A2, ×)、Eingang (入口: A1-B1, 2674 位)、Ausgang (出口: A1-B1, 3553 位)、Dokument (公文書: A1-B1, 3179 位)、Staatsangehörigkeit (国籍: A1-B2, ×)、Postleitzahl (郵便番号: A1-B2, ×)

#### ・形容詞

形容詞には、大きく分けると、groß/ klein や lang/ kurz のように、外的世界の物理的性質を記述し、それゆえ、ある程度の客観性を反映した形容詞と、interessant/ langweilig や schön などのようにもともと個人の判断や価値意識によって成り立ち、それゆえ、主観性の強い形容詞がある。「主観性の強い形容詞」は、sein とともに賓辞として使うと、文脈によっては子供っぽいニュアンスを与えるため、通常は「(~を) ~と思う」という意味の finden 等と組み合わせて使用することが多い。価値判断の形容詞は対象や出来事に意味づけする機能があるので、精神年齢の比較的高い学習者にとっては重要な語彙である。また、対象の抽象的な記述に使うケースが多い「客観性を反映した形容詞」とは異なり、印象に残る練習が組み立てやすく(=「処理の深さ」<sup>30</sup>を追求しやすく)、学習の成果が上がりやすい形容詞でもある。実際、筆者が行った簡易調査の結果から見ても初級学習者の語彙の中では、定着率が比較的良好(岩崎 2012, 36)<sup>31</sup>。しかし、「広大語彙リスト」(斜体は追加語彙に当たる形容詞)に含まれる、「主観性の強い形容詞」の大半は、「Tschirner(2008)の語彙リスト」では、上位 1200 語には入らず、その範囲の語彙では多様な価値判断を表現することは難しい<sup>32</sup>。

bequem (快適な: B1-B1, 2460 位)、billig (安い: A1-A2, 1675 位)、cool (いける・かっこいい: B2-B2, ×)、elegant (優雅な: B2-B2, 3084 位)、freundlich (友好的な・親切な: A2-A2, 1555 位)、gemütlich (居心地よい: A2-A2, 3456 位)、hübsch (かわいい: A1-A2, ×)、intelligent (知的な: A2-A2, 3470 位)、langweilig (たい

<sup>30</sup> 註 16 参照。

<sup>31</sup> 岩崎克己 (2012): 日本の初修外国語教育における基礎語彙へのアプローチ。『広島外国語教育研究』第 15 号. 21-48。

<sup>32</sup> また、自分自身の心理状態や体調を表す形容詞も、glücklich 以外は、同様に Tschirner (2008) のリストでは、上位語には入らない。müde (疲れた・眠い: A2-A2, 2077 位)、traurig (悲しい: A1-A1, 1874 位)、froh (楽しい: A2-A2, 1421 位)、fröhlich (陽気な: A2-B1, 3190 位)。

くつな: A1-A1, 2925 位)、musikalisch (音楽的な: Δ, 2268 位)、nett (親切な・感じの良い: A1-A1, 1486 位)、sportlich (スポーツの・スポーティーな: A2-B1, 1905 位)、süß (かわいい・甘い: A1-A1, 3038 位)、sympathisch (好感の持てる: A1-A2, ×)、wunderbar (すばらしい: A1-A2, 1886 位)

angenehm (心地よい: B1-B1, 1751 位)、dumm (ばかな・愚かな: A1-A1, 2518 位)、faul (怠惰な・腐った: B1-B1, ×)、fleißig (勤勉な: B1-B1, ×)、herrlich (すばらしい: B1-B1, 3011 位)、klug (賢い: ×, 3017 位)、komisch (おかしな・変な: A2-B1, 2265 位)、lustig (愉快的な・陽気な: A2-B1, 2266 位)、nützlich (役に立つ: B1-B2, 3117 位)、romantisch (ロマンチックな: ×, ×)、schick (シックな: B1-B1, ×)、schrecklich (ひどい・恐ろしい: A1-A2, 2116 位)、unmöglich (不可能な・ひどい: A2-A2, 1745 位)

それと対照的なのが、以下のような、物理的性質の記述に用いる「客観性を反映した形容詞」である。これらは、「Tschirner(2008)の語彙リスト」でも上位語に含まれ、3つのリストのいずれにおいても 850 語の範囲で必要な語彙をほぼカバーできる。

alt (古い・年をとった: A1-A1, 116 位)、breit (幅広い: A2-B1, 730 位)、eng (狭い: A2-A2, 593 位)、früh (早い: A1-A2, 324 位)、groß (大きい・背が高い: A1-A1, 74 位)、hoch (高い: A1-A2, 129 位)、klein (小さい・背が低い: A1-A1, 114 位)、kurz (短い: A1-A2, 205 位)、lang (長い: A1-A2, 95 位)、langsam (ゆっくりとした: A1-A1, 606 位)、neu (新しい: A1-A1, 80 位)、schnell (速い: A1-A1, 234 位)、spät (時間的に遅い: A1-A1, 171 位)、stark (強い: A1-A2, 207 位)、tief (深い: A1-A2, 508 位)、weit (離れた・だぶだぶの・かなり: A1-A1, 122 位)

さらに、「Tschirner (2008)の語彙リスト」の上位語には、動詞や名詞の場合と同様、初級者用の語彙と言われると、多くの教員が違和感を抱くであろう、政治経済、学問研究の分野でよく使われる中級の語彙や、さらにそれ以上のレベルの形容詞も多く含まれている。ここでも、紙幅の関係で全ては挙げられないので、後者の例だけ、その一部を上位 1000 語の範囲内で 20 個ほどリストアップする。これらも、名詞の場合と同様に、主に論文や論説文などで使われる書き言葉で、そのほとんどが抽象的な意味合いを持って副詞的に使われる形容詞である。



allgemein (全般に: ×, 341 位)、gering (ごくわずかの: ×, 468 位)、tatsächlich (実際に: B1-B2, 515 位)、zusätzlich (そのうえ: B1-B2, 520 位)、relativ (相対的に: ×, 546 位)、speziell (特に: ×, 573 位)、entscheidend (決定的な: ×, 588 位)、ständig (しょっちゅう: B1-B2, 614 位)、notwendig (必然的な: B1-B2, 620 位)、grundsätzlich (根本的に: ×, 723 位)、künftig (将来的に: ×, 742 位)、zahlreich (数多くの: ×, 841 位)、erfolgreich (成功している: ×, 843 位)、regelmäßig (規則的な: B1-B2, 860 位)、staatlich (国の: B1-B1, 869 位)、ehemalig (前の: ×, 876 位)、erforderlich (必要な: ×, 893 位)、lediglich (ただ: ×, 898 位)、unabhängig (~に依存せず: ×, 902 位)、unmittelbar (直接的に: ×, 909 位)

以上、形容詞についても、主に「**広大語彙リスト**」と「**Tschirner (2008)の語彙リスト**」の比較を行ってきたが、本節の最後に、「**広大語彙リスト**」と「**Glaboniat/u.a. (2005)の語彙リスト**」の重なり具合も見てみよう。形容詞の場合も、「**広大語彙リスト**」約 850 語に含まれる形容詞 103 個のうち「**Glaboniat/u.a. (2005)の語彙リスト**」の上位 850 語程度に相当するであろうと思われる、受容レベル A1 の形容詞は 74 個 (約 72.5%) である。また、受容レベル A2 の形容詞も 22 個 (21.6.0%) であり、両者を併せると 94% を超え、形容詞の場合も名詞の場合と同様、全体として語彙レベルがかなり重なっている。それに対し、上位約 2000 語以降に相当すると思われる受容レベルが B1 以上の形容詞は、103 個中、以下にあげる 6 個 (約 5.9%) しか無い。

bequem (快適な: B1-B1, 2460 位)、persönlich (個人的な: B1-B1, 483 位)、rund (丸い: B1-B1, 318 位)、pünktlich (時間どおりの: B1-B2, ×)、cool (いけてる: B2-B2, ×)、elegant (優雅な: B2-B2, 3084 位)

その逆に、「**Glaboniat/u.a. (2005)の語彙リスト**」ではだいたい 850 語レベル (斜体字の形容詞は一部 850 語レベルを超える) だが、「**広大語彙リスト**」には含まれない形容詞には以下のものなどがある。これらも、名詞の場合と同様、その多くは食品関係の語彙か、ドイツ語圏では良く聞くが、ドイツ語を学習する教室ではあまり聞かれない語彙である。

fein (細かな: A1-A1, 1238 位)、blöd (ばかな: A1-A1, 2669 位)、aktuell (話題の: A1-A2, 779 位)、zäh (堅い: A1-B1, ×)、gebraten (焼いた: A1-B1, 3859 位)、haltbar (日持ちする: A1-B1, ×)、pauschal (全部込み: A2-A2, ×)、gratis (無料: A2-B1, ×)

## 6 まとめとして

第4節では、日本のドイツ語教育の文脈を踏まえたコミュニカティブ・アプローチの観点から、3つの語彙リストに含まれる単語の比較を行ってきた。それからわかったことを簡単にまとめると、以下のようなになる。

- 1) 動詞・助動詞に関しては、どの語彙リストでも必要な語彙の多くが上位語の範囲（850語あるいは1200語程度まで）でカバーされている。
- 2) 名詞に関しては、「Tschirner (2008)の語彙リスト」の上位語の範囲では、日常的な口頭表現に必要な語彙の多くが含まれない。この観点では、「広大語彙リスト」や「Glaboniat/u.a. (2005)の語彙リスト」の方は必要な語彙をある程度カバーしている。
- 3) 形容詞に関しては、2つのタイプがあり、「客観性を反映した形容詞」に関しては、どの語彙リストでも必要な語彙の多くが上位語の範囲でカバーされるが、[主観性の強い形容詞]に関しては、「Tschirner (2008)の語彙リスト」の上位語では対応できず、「広大語彙リスト」や「Glaboniat/u.a. (2005)の語彙リスト」の方はカバーできる。
- 4) 「広大語彙リスト」と「Glaboniat/u.a. (2005)の語彙リスト」の比較では、一般的な日常表現に関する名詞については、どちらも大きな差はない。
- 5) ドイツ語圏で生活する際には必要となると思われる語彙は、やはり「Glaboniat/u.a. (2005)の語彙リスト」では上位語でもカバーされているが「広大語彙リスト」には少なく、逆に、日本の文化（Tempel / Neujahr 等）や大学生の持ち物や大学生活に関する話題に関係する語彙等は、後者のリストではカバーされているが前者のリストには少ない。
- 6) 頻度順のアプローチで作られた「Tschirner (2008)の語彙リスト」の上位語の中には、動詞、名詞、形容詞のすべてについて、論説文や学問的なテキストを読むのに必要だと思われる一群の中級以上の語彙が含まれている。

以上をまとめると、おおよそ次のような表になる。

	語彙の種類	広大	Glaboniat	Tschirner
1	世界のどこにしよう一般的な日常生活や自己表現に必要な動詞	○	○	○
2	世界のどこにしよう一般的な日常生活や自己表現に必要な名詞	○	○	×
3	ドイツ語圏に滞在するなら重要になるとと思われる名詞	△	○	△
4	ドイツ語文化圏とアジアや日本文化圏の対照に役立つ名詞	○	△	×

5	大学生生活や学生の日常（持ち物・服装・バイト等）に関する名詞	○	△	×
6	客観性のある程度反映した形容詞	○	○	○
7	主観性の強い形容詞	○	○	×
8	ドイツ語圏に滞在するなら重要になると思われる形容詞	△	○	×
9	ドイツ語文化圏とアジアや日本文化圏の対照に役立つ形容詞	○	△	×
10	政治経済や科学分野の論説文や論文等に必要と思われる動詞	×	×	○
11	政治経済や科学分野の論説文や論文等に必要と思われる名詞	×	×	○
12	政治経済や科学分野の論説文や論文等に必要と思われる形容詞	×	×	○

表1：3つのリストそれぞれの上位語に含まれる語彙

筆者は、上記表1の番号で言うと、コミュニケーション能力のコアの部分に関わる1、2、6、7の語彙が、基礎語彙の中にはまず含まれるべきであり、その次に、5、4、9の語彙が来るべきであり、3、8の語彙はその後に回すべきであると考えている。というのも、本稿第2節で述べたように、日本のドイツ語教育における学習者の多くは、ドイツ語圏から遠く離れた日本の大学の教養教育の枠組みの中でまずは1年間という期限を切ってドイツ語を学び始める日本人の大学生であり、最初から留学を考えていたり、ドイツ語関連の専門へ進もうと考えている学生は、ほとんどいないからである。したがって、日本で大学生活を送る学習者の知的な関心や日常生活に密着した語彙、ドイツの社会や言語文化を取りあげる場合もそれを日本の文化と対照させて語れるようなエモーショナルな側面を含んだ語彙にまず重点を置き、学習の動機付けを図るべきである。というのも、それらを通して、ドイツ語やドイツの言語文化に学生が関心を持って初めて、留学等の次のステップが生まれるからである。実際、筆者が勤務する広島大学でも、2年次以降にドイツ語関連の専門を選択したりドイツ留学を目指す学生の大半は、教養教育における1年次のドイツ語学習を通してドイツ語圏の言語文化に関心を持ったことをその志望動機に挙げている。なお、教養教育の枠組みを越えてドイツ語学習を続け、中級以上の段階に至った場合には、表1の10～12の領域の語彙をシステムティックに教えていくことも重要であろう。

最後に、日本独文学会あるいはドイツ語教育部会レベルでのドイツ語基礎語彙集作成プロジェクトとでも言うべきものに対する期待を述べ、本稿を終えた

い。筆者には、上述したような日本人の学習者の典型的な学習プロフィールを念頭に置きつつ、初級から中級（さらには上級）に至るドイツ語の語彙リストを、そろそろ学会として作成すべき時期に来ているように思われる。その場合は、『JACET 8000 英単語』が BNC を基準スケールとしたように、たとえば Tschirner (2008)の語彙リストや他の（理想的にはもう少し規模が大きく均整の取れた）ドイツ語コーパスから作られた語彙リストを基準リストとし、それに、他のアプローチで作られた複数の語彙リストやデータを参照しながら何人かの専門家が合議をしながら作って行くことになるだろう。そうしたドイツ語教育の分野における語彙リストの研究の進展に本稿が少しでも役に立てば幸いである。

# WORTSCHATZLERNEN AUS DER DIDAKTISCHEN PERSPEKTIVE

Angelika WERNER / Shiro KATTO

## 1 Einleitung

Spracherwerb und somit auch Wortschatzerwerb ist ein Lernprozess, der genau wie andere Lernprozesse bei einem selbständigen Individuum abläuft, in Auseinandersetzung mit der Umwelt und dem bisherigen Wissen der Lerner. Nach den heutigen gültigen konstruktivistischen Ansätzen sind es dementsprechend emergente selbstorganisierte Prozesse<sup>1</sup>. Eine Theorie im konstruktivistischen Bereich liegt nicht vor, auch nur wenige Longitudinalstudien, jedoch Einzeluntersuchungen zu verschiedenen Themen, besonders im L1-Bereich.

Für den Fremdsprachunterricht ergibt sich dadurch die lernerorientierte Frage: Welche Einheiten sind für den Lerner überhaupt „Wortschatz“ (Lemmata)?

Allgemeine Übereinkunft besteht bei einfachen Wörtern und folgenden Einheiten, die auch in Lernmaterialien jeglicher Art zu finden sind:

- Lexem- und Wortfamilien *gehen, ging, gegangen* oder *gehen, gehe, ...*
- Verbundene Lexeme und Komposita *Dokkyo Universität, Lautsprecher*
- Chunks und Phraseme *sag mal, nach Hause gehen*

---

<sup>1</sup> Schmidt, Claudia (2010): Kognitivistische/ Konstruktivistische/ Konnektionistische Ansätze. In: Krumm/Fandrych/Hufeisen/Riemer (Hrsg.): Deutsch als Fremd- und Zweitsprache. Ein internationales Handbuch. Berlin. S. 807-817.

- Phraseologismen, Phrasen *Herzlichen Glückwunsch zum Geburtstag!*
- Wortfelder wie Antonymien *groß-klein*
- Hyperonymien *Hund-Dackel, -Bernhardiner*
- Kollokationen *blond-Haare* usw.

Der Fremdsprachenwortschatz eines Lerners ist nur als momentaner zu beschreiben, er ist flexibel und dynamisch, dauernder Änderung unterworfen. Er ist zwangsläufig an die anderen Fähigkeiten und Fertigkeiten des Lerners gekoppelt, er ist angepasst an die Interlanguage und Person des Lerners. Hierbei sind Faktoren wie Niveau, Alter, Lerntyp, Lerntradition, Lernumgebung, usw. von entscheidender Wichtigkeit. Deshalb muss Wortschatz-Lernen - besonders beim instruierten Lernen - flexibel und vielseitig, interessant und effektiv sein. Allerdings ist es in weitem Maße individuell unterschiedlich. Daraus folgt: Es müssen genügend differenzierte Input- und Output-Gelegenheiten gegeben werden, also Lern- und Übungsmöglichkeiten rezeptiver und produktiver, mündlicher und schriftlicher Art. Nicht nur Lexem-Wissen (statisch), sondern auch prozedurales Wissen (grammatisch und pragmatisch, usw.), mit anderen Worten Sprachgebrauch in Situationen, ist das Lern- und Übungsziel. Lernen ist immer ein langwieriger komplizierter Vernetzungsprozess und muss durch Anwendung vertieft werden (rezeptiv und produktiv, explizit und implizit), i.e. durch ständige Wiederholung gefestigt. In Lernmaterialien (den Langenscheidt Vokabeltrainern etwa<sup>2</sup>) und auch Lehrerhandreichungen gibt es oft Lerntipps, wie: unbekannte Vokabeln herauszuschreiben, Vokabeln wöchentlich zu wiederholen, usw. Diese Lerntipps reichen jedoch nicht für ein effektives Wortschatzlernen,

---

<sup>2</sup> In digitaler Form z.B. [http://www.langenscheidt.de/reihe/833/Langenscheidt\\_Vokabeltrainer\\_60](http://www.langenscheidt.de/reihe/833/Langenscheidt_Vokabeltrainer_60), auf CD-Rom oder aber in Papierform etwa: Rohrman, Lutz/Lemcke, Christiane (2008): Wortschatz Intensivtrainer A2. Langenscheidt.

Lerner brauchen Anleitungen für eine systematische Wortschatzarbeit in Lernmethodik und Lernstrategien.

Für die Konzeptualisierung der Wortschatzvermittlung im gesteuerten Lernkontext gehen Antos / Lewandowska (2007)<sup>3</sup> davon aus, dass für die praktische Umsetzung Fragestellungen, Erkenntnisse und Ergebnisse aus vier Dimensionen bzw. Welten berücksichtigt und angemessen aufeinander bezogen werden sollten: (1) Aussagen über die Beschaffenheit des mentalen Lexikons, (2) die Lerntheorie, (3) die Interkulturalität beim Wortschatz und (4) die jeweiligen institutionellen Bedingungen. Auf dieser Basis soll hier eine Wortschatzlernstrategie für japanische Lernende, die Deutsch als zweite Fremdsprache an japanischen Universitäten lernen, vorgestellt werden. Wir haben dazu eine Lerner-Wortschatz-Karte, die von uns so genannte „EO-Karte“ (= ErfahrungsOrientierte Karte), entwickelt. Diese lässt sich unserer Ansicht nach sehr gut in den handlungsorientierten Fremdsprachenunterricht integrieren.

## **2 Wörternetz-Strategie**

Als wegweisende wissenschaftliche Arbeit über die Lernerperspektive bzgl. fremdsprachlichen Wortschatzlernens soll hier die u. E. effektive Wortschatzlernstrategie von Christiane Neveling (2004a, b)<sup>4</sup> kurz vorgestellt werden.

---

<sup>3</sup> Antos, Gerd/Lewandowska, Anna (2007): „Amsel, Drossel, Fink und Star“ – Wortschatzerwerb und seine Vermittlung im Spannungsfeld von vier „Forschungswelten“. In: Krumm, Hans-Jürgen/Tselikas, Paul (Hrsg.): Theorie und Praxis - Österreichische Beiträge zu Deutsch als Fremdsprache: Schwerpunkt: Wortschatz. Innsbruck. S. 11-22.

<sup>4</sup> Neveling, Christiane (2004a): Wörternetze als Abbilder mentaler Lexikonstrukturen – Untersuchungen zur Güte eines neuen Forschungsverfahrens. In: Fremdsprachen lehren und lernen 33. S. 128-146; Neveling, Christiane (2004b): Wörterlernen mit Wörternetzen – eine kognitiv-affektive Strategie. In: Börner, Wolfgang/Vogel, Klaus (Hrsg.): Emotion und Kognition im Fremdsprachenunterricht. Tübingen. S. 190-216.

Das Ergebnis ihrer Umfrage bei 453 deutschen Schülern, die Französisch lernen, ist zum einen der negative Eindruck bei den Lernern über Wortschatzlernen wie monoton, langweilig und lästig (Abb.1<sup>5</sup>), zum anderen hat sie mechanische Lernverfahren wie Listenlernen und das Nichtvorhandensein variiertes Wortschatzlernstrategien und -techniken bei Probanden festgestellt. Neveling führt das schlechte Behalten der Wörter auf einen „Mangel an Affektivität und an ausreichendem wie nützlichem prozeduralen Wissen“ zurück<sup>6</sup>.

Neveling hat zum Zweck des langfristigen Speicherns und Abrufens Wörternetz-Strategien erarbeitet, denen Erkenntnisse über das mentale Lexikon zugrunde liegen und deren Speichereffektivität und kognitiv-affektive Befriedigung bei Probanden empirisch validiert worden sind<sup>7</sup>.

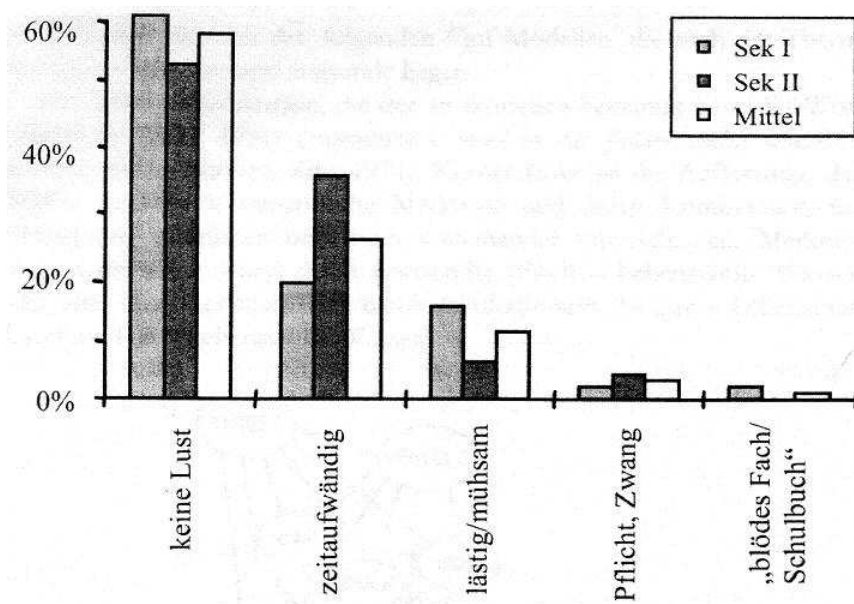


Abb.1: Ergebnis der Umfrage zum Wortschatzlernen

<sup>5</sup> Neveling (2004b). S. 191.

<sup>6</sup> Ebd. S. 192.

<sup>7</sup> Vgl. Neveling (2004b).



Diese Strategien werden in folgenden Schritten praktiziert: Der Lerner ordnet die zu lernenden Wörter nach den Strukturen des mentalen Lexikons (s.u.) auf dem Papier an (Ordnungsstrategien), dann werden die Wörter durch Linien miteinander verbunden und durch Kreise und Farben illustriert. Daraufhin werden diese Netze durch Hervorhebungen, Symbole, verschiedenartige Umrahmungen dargestellt (Elaborationsstrategien). Das so entstandene Wörternetz wird durch Wiederholungsstrategien und Abrufstrategien<sup>8</sup> gefestigt.

### **3 Mentales Lexikon**

#### **3.1 Begriff und Struktur**

Das mentale Lexikon als Modell der Psycholinguistik, das erklären soll, wie der Wortschatz einer Sprache im menschlichen Gedächtnis gespeichert ist und abgerufen werden kann, ist nach Raupach (angelehnt an Schwarz [1992]<sup>9</sup>), als Teil des Langzeitgedächtnisses anzusehen, welches als „Nahtstelle zwischen spezifischen Sinneseindrücken (bei der Sprachperzeption: *input*) oder motorischen Mustern (bei der Sprachproduktion: *output*) auf der einen Seite und mental repräsentierten Wissensstrukturen auf der anderen Seite“<sup>10</sup> funktioniert. Wörter sind durch enge Beziehungen im Langzeitgedächtnis so repräsentiert, dass man in (kommunikativen) Situationen ohne weiteres, wenn auch nicht immer erfolgreich, dort auf sie zugreifen kann, wo sie evoziert werden.

Was die Struktur des mentalen Lexikon angeht, ist es in mehrere Teilnetze mentaler Repräsentationen gegliedert, die interlexikalisch miteinander komplex vernetzt sind. Obwohl bei der Fragestellung, welche Teilnetze bzw. Felder für das mentale Lexikon angenommen werden sollen, bisher keine Einigkeit besteht, halten wir folgende Teilnetze

---

<sup>8</sup> Ebd. S. 197f.

<sup>9</sup> Schwarz, Monika (1992): Einführung in die Kognitive Linguistik. Tübingen.

<sup>10</sup> Raupach, Manfred (1997): Das mehrsprachige mentale Lexikon. In: Börner, Wolfgang/Vogel, Klaus (Hrsg.): Kognitive Linguistik und Fremdsprachenerwerb. Tübingen. S. 19-37. S. 21.

nach Neveling (2004a, 2004b), denen die Erkenntnisse von Kielhöfer (1994)<sup>11</sup> zugrunde liegen, für heuristisch sinnvoll:

**Sachnetz:** Logischer Ablauf und Zusammenhang einer Sache, was auf in einem soziokulturellen Milieu herausgebildeten Weltwissen beruht: *in der Bäckerei; grüßen, ein Brötchen, wählen, kaufen und bezahlen, usw.*

**Begriffnetz:** Hyponyme: *Brot; Semmelbrot, Baguette, Pumpernickel, usw.*

**Wortmerkmalsnetz:** Synonyme und Antonyme: *lecker, fein und köstlich* (Synonym), *lecker- scheußlich; fein- fade* (Antonym)

**Syntagmatisches Netz:** Semantische Zusammenpassung, Kollokationen und Redewendungen: *Brot backen, Brot herstellen, Brot wegschmeißen*

**Wortfamiliennetz:** Von einem Wort abgeleitete lexikalische Einheiten: *finden; erfinden, befinden, Befund und ausfindig usw.*

**Klangnetz:** Lexikalische Einheiten mit auditiven Ähnlichkeiten: *Haus und Hof, hier und heute*

**Affektives Netz:** „Affektive Ladung“<sup>12</sup>, welche ein Wort als Nebenbedeutungen hat, mit Emotionen verknüpft: *Freude; freuen, hoch- und himmeljauchzend*

Die Vernetzung der Teilnetze muss neben interpersonalen Gemeinsamkeiten soziokulturelle und kognitiv individuelle Unterschiede und Kreativität aufweisen, wobei verschiedene Semantisierungen und Assoziationen durch Welt- und Spracherfahrungen einen Knoten im Netz bilden. Zum mentalen Lexikon gehört nicht nur das vorgefertigte sprachliche Regelwissen hinsichtlich der phonologischen, orthografischen und morphologisch-syntaktischen (grammatischen) Informationen, sondern

---

<sup>11</sup> Kielhöfer, Bernd (1994): Wörter lernen, behalten und erinnern. In: Neusprachliche Mitteilungen (4/1994). S. 211-220.

<sup>12</sup> Ebd. S. 214.

auch „Verfahrensweisen für die Benutzung sprachlicher Produkte“<sup>13</sup>. Die Vernetzung im mentalen Lexikon ist nie fixiert, sondern wird durch physiologische und kognitive Reifung zum einen, Zuwachs an Welt- und Spracherfahrung sowie Welt- und Sprachwissen zum anderen, in einem Individuum ständig umstrukturiert<sup>14</sup>. Bei einer erfahrenen Person im Erwachsenenalter z. B. sind demnach Wörter mental anders konzeptualisiert als bei einer im Kindesalter.

Je mehr Teilnetze sich auf eine gelernte lexikalische Einheit beziehen, desto leichter kann sie abgerufen werden und dies wirkt sich sogar auf das langfristige Behalten im Gedächtnis positiv aus. Im Vergleich dazu sind Behalten und Abruf von isolierten Wörtern eher schwierig<sup>15</sup>. Aus diesem Grund müsste mehrkanaliges und ganzheitliches Wortschatzlernen im Fremdsprachenunterricht den großen Vorteil haben, dass es den Zugriff auf gelernte lexikalische Einheiten durch vielfältige Vernetzungen erleichtert<sup>16</sup>. Dies weist auch darauf hin, dass verschiedene Teilnetze im mentalen Lexikon gleichzeitig ko-aktiviert werden, wenn das gelernte Wort rezeptiv und produktiv verarbeitet wird. Wenn also ein Wort aktiviert wird, werden gleichzeitig auch andere semantisch oder form-ähnliche, subjektiv vernetzte Wörter, zu denen möglicherweise auch Synonyme, Antonyme und Hyponyme gehören, und Konzepte als Weltwissen ko-aktiviert<sup>17</sup>. Diese Überlegung, dass mehrere Teile im Gedächtnis ko-aktiviert werden, ist kompatibel mit dem konnektionistischen Lernmodell<sup>18</sup>. Daraus lässt sich ableiten, dass ein mentales Lexikon sehr dynamisch, flexibel und völlig anders systemati-

---

<sup>13</sup> Wolff, Dieter (2002a): Fremdsprachenlernen als Konstruktion – Grundlagen für eine konstruktivistische Fremdsprachendidaktik. Frankfurt am Main. S. 122.

<sup>14</sup> Kielhöfer (1994).

<sup>15</sup> Neveling (2004a, 2004b); Kielhöfer (1994).

<sup>16</sup> Bohn, Rainer (1999): Probleme der Wortschatzarbeit. Berlin. S. 71, 87-94; Dies betrifft auch das so genannte *TOT-Phänomen*, ein Wort das einem „auf der Zunge liegt“.

<sup>17</sup> Wolff (2002a). S. 123.

<sup>18</sup> Schmidt (2010).

siert ist als ein papiernes Wörterbuch, dessen Einträge in alphabetischer Reihenfolge stehen und keineswegs neu kreiert werden<sup>19</sup>.

### **3.2 Mehrsprachigkeit im mentalen Lexikon**

Der Mensch erweitert und bereichert vom Kindesalter an die Einträge im mentalen Lexikon der L1 durch Generalisierung und Ausdifferenzierung der lexikalischen Einheiten, so dass sie im alltäglichen Sprachgebrauch in höchstem Maße automatisiert werden, bis zum Sprachniveau eines Erwachsenen. Wenn es dahingegen um das Fremdsprachenlernen bei Erwachsenen geht, ist zu berücksichtigen, dass schon (kulturelles) Weltwissen und sprachliche Konzepte mental repräsentiert sind, die aus der L1 stammen. Diesen werden dann fremdsprachliche Wörter hinzugefügt und damit sind mehrsprachige bzw. interlinguale mentale Lexika weitaus komplizierter als monolinguale.

Sprachlernende einer Fremdsprache auf Anfängerniveau gehen immer in der Wortschatzarbeit von ihrer L1 (und den bereits gelernten Fremdsprachen) aus, während fortgeschrittene Lernende eine gewisse Unterscheidung zwischen Konzepten lexikalischer Einheiten vornehmen, selbst wenn das nie bedeutet, dass mentale Lexika der jeweiligen Sprachen voneinander unabhängig vorhanden sind<sup>20</sup>. Dabei ist entscheidend, wie man sich bislang mit der fremdsprachlichen Wortschatzarbeit beschäftigt hat, ob man die Fremdsprache über Bedeutungsgleichungen im zweisprachigen Wörterbuch oder über jeweils einsprachige Bedeutungserklärungen gelernt, oder Wortschatzarbeit kontextabhängig oder -unabhängig praktiziert hat, usw. Von Bedeutung für die Wortschatzarbeit ist, dass der Lernende Konzepte lexikalischer Einheiten zwischen seiner L1 und einer Fremdsprache auseinanderhalten und miteinander

---

<sup>19</sup> Aitchison, Jean (2003): *Words in the Mind: An Introduction to the Mental Lexicon*. Malden.

<sup>20</sup> Wolff, Dieter (2002b): *Das mentale Lexikon - Grundlage der Sprachkompetenz in der Muttersprache und der Fremdsprache*. In: *Der Fremdsprachliche Unterricht Englisch* (1/2002). S. 11-14.

koordinieren kann, wie das häufig bei bilingualen Sprechern zu beobachten ist, ein Prozess, bei dem es sich um die Modifizierung der bereits in der L1 gebildeten Konzepte handelt.

#### **4 Autonomiefördernde Überlegung: „Lernen lernen“**

Das Konzept „Lernen lernen“ setzt voraus, dass man für sich effektive Lernerstrategien und -techniken ausfindig macht, ausprobiert und weiter entwickelt. Dieser lange Lernprozess lässt sich jedoch durch Unterstützung und Beratung der Lehrperson gezielt als Gegenstand im Unterricht aufgreifen, indem Lernenden die Gelegenheit gegeben wird, darüber nachzudenken, auf welche Art und Weise man eigentlich das Lernen organisieren kann. Anhand einer erweiterten Perspektive und mit Informationen über potenzielle Lernmöglichkeiten strategischer bzw. technischer Art sollten Lernende dann schrittweise unabhängig von Lehrperson und Unterricht werden.

Autonomie in der Wortschatzarbeit ist nach Bohn (1999)<sup>21</sup> die Fähigkeit, „sich Wörter einer fremden Sprache weitgehend selbstständig (selbsttätig, selbstbestimmt) anzueignen“ und die Voraussetzung dafür ist, „dass wir wissen, wie wir Wörter am besten lernen.“

Die Relevanz der Autonomieförderung und des lebenslangen Lernens hat Auswirkung auf das Lernziel des DaF-Unterrichts an japanischen Universitäten, welche quantitativ die zentralen Orte des Deutschlernens in Japan sind. Aufgrund der zeitlichen und realistischen Bedingungen für eine Zweitfremdsprache wie Deutsch hält es etwa Degen für angebracht, „einen relativ großen Teil der Unterrichtszeit für die explizite und implizite Vermittlung von Autonomie zu investieren und [...] im Rahmen der Allgemeinbildung sogar zum Hauptziel des Sprachkurses

---

<sup>21</sup> Bohn (1999). S. 94.

zu machen.“<sup>22</sup> Einerseits könne das außerunterrichtliche Lernen, das wegen des geringen Zeitbudgets des Unterrichts in den Vordergrund gerückt werden müsse, bei denjenigen effizient gefördert werden, die sich für die Zielsprache interessieren. Andererseits könnten sich aber auch die Studenten dadurch mit Lernerstrategien als Werkzeug versehen, was ebenso für das Lernen anderer Fremdsprachen von Nutzen ist.

## **5 Erfahrungsorientierte Wortschatzlernkarte**

### **5.1 Was ist eine EO-Karte?**

Aus diesen Überlegungen zum Wortschatzerwerb haben wir als Vorschlag für eine effiziente Lernerstrategie für das Wortschatzlernen ein Hilfsmittel erstellt: die erfahrungsorientierte Wortschatzlernkarte, die EO-Karte. Sie ist als Wortschatzlernmaterial konzipiert, das zwischen Wortkartei, Vokabelheft und Wörterbuch liegt. Die EO-Karte zielt in der Wortschatzlernstrategie darauf ab, (1) Lernenden bewusst zu machen, wie man fremdsprachliche Wörter anders als durch mechanisches Wiederholen lernen kann und (2) Lernende unter dem Konzept „Lernen lernen“ dazu zu befähigen, das selbständige Wortschatzlernen schrittweise gestalten zu können. Die erstere Zielsetzung hat dabei instruktivistische Dimension, die letztere konstruktivistische. Beide Dimensionen sind vereinbar, weshalb sie hier kombiniert werden.

Zu (1): Die EO-Karte bedingt explizites bzw. intentionales Lernen<sup>23</sup>, das sich absichtlich mit dem Lerngegenstand beschäftigt, um sich diesen ins Gedächtnis einzuprägen. Es unterscheidet sich damit von implizitem bzw. inzidentiellem Lernen, wo beim Lesen, Hören und Reden beiläufig

---

<sup>22</sup> Degen, Ralph (2005): Vorschläge zur expliziten Förderung autonomen Lernverhaltens im Fremdsprachenunterricht an japanischen Hochschulen. In: Balmus/Oebel/Reinelt (Hrsg.): Herausforderung und Chance. Krisenbewältigung im Fach Deutsch als Fremdsprache in Japan. München. S. 129-147. S. 132.

<sup>23</sup> Intentionales Wortschatzlernen fördert nach Bohn (1999), S. 80 den Vermittlungswortschatz.

Lernen evoziert wird. Beim Fremdsprachenlernen ist es von Nutzen, beide Verfahren in den Unterricht zu integrieren.

Die EO-Karte soll nicht nur als Strategie zum Wortschatzlernen, sondern auch als explizite Erläuterung dienen, wie man Wörter einer Fremdsprache lernen kann, als instruktivistische Dimension, die den Benutzer unaufdringlich beim Lernen anleiten soll. Indem nach der Teilnetzidee des mentalen Lexikons verschiedene Zugriffsmöglichkeiten in die EO-Karte implementiert werden, soll das Wortschatzlernen damit über einfaches mechanisches Wiederholen hinausgehen und dem Lernenden bewusst machen, dass ein neu gelerntes Wort einer Fremdsprache mit vorhandenen Wissensbeständen und Erfahrungen interagiert und in vielfältiger Weise in die Wissensstruktur integriert werden kann.

Zu (2): In enger Korrelation mit (1) ist die Anfertigung einer EO-Karte eine intensive Beschäftigung damit, verschiedene Zugangsmöglichkeiten in freier eigener Entscheidung auszuprobieren, insofern dass der Lernende schrittweise die Verantwortung für das eigene Lernen übernimmt, um dadurch den für sich angemessenen Lernstil und -vorzug in der Wortschatzarbeit herauszufinden. Dies betrifft den Weg zur Lernerautonomie. Hier steht die konstruktivistische Dimension im Vordergrund, die aktive und initiiierende Tätigkeiten von der Lernerseite erfordert.

## **5.2 Struktur der EO-Karte**

Auf der EO-Karte sind Einträge verschiedener Komponenten vorgesehen. Im Prinzip sollte man dem Lerner den kreativen Umgang damit und die Entscheidung überlassen, welche Informationen über ein Wort eingetragen werden. Die EO-Karte (s.Abb.2) besteht aus folgenden 10 Komponenten:

**Stichwort oder idiomatischer Ausdruck (見出し):** Man kann nicht nur Wörter, sondern auch idiomatische Ausdrücke und Redewendungen in dieses Feld eintragen.

**Aussprache (発音):** Dies dient zur Sicherung des phonetischen Wissens, wobei die Lesungsart entweder in IPA oder in Katakana-Schrift (oder auch individuell frei) notiert wird.

**Bedeutungsangabe (意味):** Die Bedeutungsangabe des Wortes muss nicht „neutral“ aus dem papiernen Wörterbuch abgeschrieben werden, die Bedeutung kann aus einem konkreten Gebrauchskontext stammen.

**Beispielsatz (例文):** Der Beispielsatz kann kreativ sein, am besten stammt er aus einem erfahrenen Gebrauchskontext.

**Verweis (参照):** Rechts oben in der Karte ist eine Leerstelle für einen Verweis im eigenen Kartensystem angebracht, z. B. für den Anfangsbuchstaben bei einem alphabetischen System.

**Präfix und Suffix (接頭辞と接尾辞):** Suffixe und Präfixe kann man in die Pfeile eintragen, damit diese kleinen Bedeutungsträger besonders bewusst gemacht werden.

**Kollokation (コロケーション):** Kollokative Verbindungen mit dem eingetragenen Wort werden in diesem Kästchen festgehalten.

**Bild (イメージ図):** Man kann dort ein Bild (in Farbe) malen, das aus dem Wort imaginativ abgeleitet wird. Anwendungspotenzial ist auch die Schlüsselwortmethode (= keyword methode)<sup>24</sup>.

**Assoziogramm (連想):** In das Assoziogramm lassen sich nicht nur Synonyme und Antonyme, sondern auch phonetische Ähnlichkeiten zwischen L1 und L2 sowie subjektive Vorstellungen implementieren.

**Erfahrung (経験):** Man kann in diesem Kästchen eigene situative Erfahrungen schriftlich oder bildlich festhalten, die beim Gebrauch des in der EO-Karte einzutragenden Wortes gemacht wurden. Erfahrun-

---

<sup>24</sup> Vgl. Schmitt, Norbert (2000): Vocabulary in language teaching. Cambridge; Bohn (2000). S. 99.



gen als narrative Wiedergaben der Vergangenheit bei einer Person zu aktivieren, aktiviert auch das Gedächtnis mit den erlebten Wahrnehmungen, Sinneseindrücken, Emotionen und Gefühlen einer Situation. Wenn das zu lernende Wort mit einer Erfahrung zusammengebracht und ein Teil davon auf die EO-Karte als kurzer Text oder Foto eingetragen wird, kann man das Lemma besser im Gedächtnis aufrufen.

Als Format der EO-Karte bietet sich die Größe A5 als für Notizen angemessene Größe an.

### **5.3 Aspekte der EO-Karte**

**1) Sichtbarmachung des Lernprozesses:** Ein von den Lernenden selbst angefertigtes „Wörterbuch“ in Form von EO-Karten hat den Vorteil, den individuellen Lernprozess in der Wortschatzarbeit sichtbar zu machen; einfacher gesagt, die EO-Karte kann als Unterstützungsmittel zur Reflexion, wie etwa ein Portfolio oder ein Lerntagebuch eingesetzt werden. Wie Lernende sich mit dem Wörterlernen befasst haben, kann man keineswegs ausschließlich an den Ergebnissen von Wortschatztests erkennen, wo es darum geht, ob Lernende bestimmte Wörter bis zum bestimmten Zeitpunkt des Testablegens auswendig gelernt haben. Die Sichtbarmachung des Lernprozesses mit der EO-Karte gibt Lernenden die Gelegenheit, über eigene Lernpräferenzen bzw. den eigenen Lernstil zu reflektieren. Außerdem sind Einträge in der EO-Karte auch nützlich für die Sprachlernberatung, die Lernende zur Reflexion des eigenen Lernens anregen und sie motivieren sollen<sup>25</sup>.

---

<sup>25</sup> Kleppin, Karin (2003): Sprachlernberatung: Zur Notwendigkeit eines eigenständigen Ausbildungsmoduls. In: *ZFF* (14/1), S. 71-85; Mehlhorn, Grit/u.a. (2009): Studienbegleitung für ausländische Studierende an deutschen Hochschulen. München.

The diagram shows the layout of an EO-Karte. At the top left, there are two arrows: one pointing left labeled '接頭辞' (prefix) and one pointing right labeled '接尾辞' (suffix). Below these are fields for '発音:' (pronunciation), '見出し:' (headline), '意味:' (meaning), and '例文:' (example sentence). To the right is a large thought bubble labeled 'イメージ図' (image diagram). Below the thought bubble is a dashed box labeled 'コロケーション:' (collocation). At the bottom left is a box labeled '連想:' (association) containing a sun-like diagram with rays. At the bottom right is a box labeled '経験 (この単語を学んだ状況・内容など):' (experience).

Abb.2: EO-Karte

**2) Selbstentscheidung:** Die EO-Karte bietet Lernenden einen Anlass zur Selbstentscheidung. Sie haben zumindest die Auswahl, welche Einträge (aus gelernten Wörtern) gemacht werden. In der EO-Karte sind außerdem verschiedene Komponenten aus den Teilnetzen des mentalen Lexikons für Einträge vorgesehen. Der Eintrag eines Wortes in der EO-Karte muss jedoch nicht vollständig sein, so wie dies auch für das mentale Lexikon gilt. Vielmehr ist es von Bedeutung, dass Lernende in der bewussten Beschäftigung damit selber Entscheidungen treffen, welche Komponenten für ein erfolgreiches Lernen ausgefüllt werden sollten.

**3) Verarbeitungstiefe:** Mit der aus der Psychologie stammenden Hypothese der Verarbeitungstiefe (depth-of-processing hypothesis) ist die Tiefe der kognitiven Prozesse beim Lernen eines Individuums gemeint. Je tiefer die kognitive Verarbeitung ist, desto besser kann das Gelernte

im Gedächtnis langfristig verankert werden. Unserer Ansicht nach kann die EO-Karte als explizite Wortschatzlernstrategie das Entfaltungspotenzial für diese Verarbeitungstiefe entfalten, weil ein neu zu lernendes Lemma einer Fremdsprache in Verbindung mit verschiedenen Komponenten gelernt wird, die eine relativ tiefe mentale Verarbeitung erfordern.

**4) Wörterbuchdidaktik:** Beim Fremdsprachenlernen ist die Wörterbuchbenutzung unerlässlich und als kompetente Informationsbeschaffung wichtig für die Autonomieförderung<sup>26</sup>. Dennoch ist für Lernende häufig nicht selbstverständlich, mögliche Potenziale eines Wörterbuchs zu nutzen und für eigenes Fremdsprachenlernen effektiv zu machen<sup>27</sup>.

**5) Erfahrungsorientiertheit:** Die EO-Karte lässt sich auch als erfahrungsorientiertes Produkt betrachten, etwa in der Projektarbeit, wobei Erfahrungen eines Lernenden in einer sichtbaren Form festgehalten werden. Der wichtige Punkt hierbei ist, dass es bei der EO-Karte nicht um vorgegebene Einträge geht. Es ist wenig lernförderlich, wenn der Lernende den Eintrag aus einem papiernen Wörterbuch bloß abschreibt, ohne mit diesem Wort etwas erfahren zu haben. In diesem Sinne kann die EO-Karte einen Teil des prozessorientierten Lernens (wie etwa auch Projektarbeit und Stationenlernen) ausmachen. Es sollten dabei Freiräume für offene Handlungsaktivitäten bereitgestellt werden, damit sich die Lernenden kognitiv und emotional dynamischer als im klassischen Frontalunterricht verhalten. Ob das in der Praxis gelingt, hängt größtenteils von der Unterrichtsgestaltung durch die Lehrperson ab, was gewisse organisatorische Bemühungen erfordert. Jedoch kann die EO-Karte auch für Lehrpersonen hilfreich sein, um die Leistung der jeweiligen

---

<sup>26</sup> Zöfgen, Ekkehard (2010): Wörterbuchdidaktik. In: Hallet, Wolfgang/König, Frank G. (Hrsg.): Handbuch Fremdsprachendidaktik. Seelze. S. 107-111.

<sup>27</sup> Köster, Lutz (2001): Wortschatzvermittlung. In: Helbig/Götze/Henrici/Krumm (Hrsg.): Deutsch als Fremdsprache. Ein internationales Handbuch (2). Berlin. S. 887-893.

Lernenden zu evaluieren, denn Leistungsmessung ist in der Regel mit dem gesteuerten Fremdsprachenlernkontext verbunden.

Als Nachteile für eine Unterrichtsarbeit mit EO-Karten wären anzuführen, dass die EO-Karte nicht allen Lernstilen entgegenkommen kann, besonders nicht bei Lernenden mit auditiv orientierten Präferenzen. Das Material der EO-Karte müssen die Lernenden selber anschaffen, was Kosten verursacht. Viel wesentlicher ist aber, dass viele Lernende einer EO-Karte anfangs sehr skeptisch gegenüberstehen und die Herstellung als viel zu aufwändig, zeitintensiv und umständlich ansehen.

## **6 Perspektive**

Aus der didaktischen Perspektive ist zu beachten, dass beim Einsatz neuer Strategien und Techniken in der Praxis deren Benutzung Zeit zur Gewöhnung bzw. zum Kennenlernen verlangt und evtl. Konflikte mit den subjektiven Theorien, den Lernstilen und -typen der Lerner entstehen. Ein Fremdsprachenunterricht, in dem man in vielfältiger Weise mit Wissen, Perspektiven, Materialien und anderen Personen interagiert, sorgt bei den Beteiligten immer für Erfahrungen. Das Unterrichtsdesign einer solchen Lernumgebung für Erfahrungen aus aktiven Handlungen ist größtenteils auf eine instruierende Lehrperson angewiesen. Lerner aber halten neue Strategien und Techniken erst dann für nützlich, wenn sie deren Effektivität tatsächlich nachvollzogen haben. Es ist also nicht ausreichend, wenn eine Lehrperson im Unterricht über die neue Strategie und Technik nur informiert; vielmehr ist die aktive Beteiligung der Lerner erforderlich, an Gebrauchssituationen bzw. Aufgaben, die von der Lehrperson erarbeitet werden. Inhaltsreiche Erfahrungen im Unterricht und Lerneffekte beim Gebrauch der EO-Karte beeinflussen sich dann gegenseitig positiv.

Wir sind in diesem Aufsatz auf die Potenziale der EO-Karte für die Entfaltung des autonomen Wortschatzlernens eingegangen. Die

EO-Karte als Vorschlag einer Strategie entstammt jedoch theoretischen Überlegungen, deswegen sollte sie weiterhin, sowohl quantitativ als auch qualitativ, empirischen Überprüfungen auf Validität und Reliabilität der Lerneffekte in der Praxis unterzogen werden. Dabei kann zum einen der individuell differente, kreative Gebrauch der EO-Karte, zum anderen deren Einsatz in kooperativen Lernformen als wertvoll für weitere Analysen angesehen werden, da dessen Dokumentierung sowohl den Lernprozess eines einzelnen Lerners verrät, als auch kooperative Lerntendenzen aufzeigt.