

**Treffungen, Sinkung und Benützung -**  
**Korpuslinguistische Untersuchung des Erwerbs von**  
**derivationsmorphologischen Wortbildungsregularitäten bei fortgeschrittenen**  
**Lernern des Deutschen als Fremdsprache**

**Magisterarbeit**  
**zur Erlangung des akademischen Grades**  
**Magister Artium (M.A.)**  
**im Fach Germanistische Linguistik**

**Humboldt-Universität zu Berlin**  
**Philosophische Fakultät II**  
**Institut für Germanistik**

**eingereicht von:**

**Michael Berth**  
**geb. am 31.10.1978**  
**in Reutlingen**

**Wissenschaftliche Betreuerin: Prof. Dr. Anke Lüdeling**

**Berlin, den 17. September 2009**

# **Inhaltsverzeichnis**

## **1. Einleitung (5)**

## **2. Theoretischer Hintergrund (9)**

### **2.1. Zur Theorie der Wortbildung (9)**

- 2.1.1. Warum überhaupt Wortbildung? (9)
- 2.1.2. Wortbildung verstanden als Derivation (10)
- 2.1.3. Die *-ung*-Suffigierung (11)
- 2.1.4. Ist das Suffix *-ung* eine Lexikoneinheit (12)
- 2.1.5. Wortbildungsmuster (13)
  - 2.1.5.1. Wortbildungsmuster (und semantische Muster) der *-ung*-Suffigierung (15)
  - 2.1.5.2. Aktivierung eines Lexikoneintrages oder produktive Wortbildung? (17)
- 2.1.6. Morphologische Produktivität (18)
- 2.1.7. Bildung, Beschränkungen und Produktivität der *-ung*-Suffigierungen (20)
  - 2.1.7.1. Aussagen über Bildung und mögliche Basen in den traditionellen Grammatiken (20)
  - 2.1.7.2. Blockierung (22)
  - 2.1.7.3. Pragmatische Bildungsrestriktionen (23)
  - 2.1.7.4. Die Produktivität der Wortbildungsmuster nach Motsch (1999) (24)
- 2.1.8. Messbarkeit von Produktivität (26)

### **2.2. Lernersprache (27)**

- 2.2.1. Der Weg zur Entwicklung des Konzeptes „Lernersprache“ (28)
- 2.2.2. Lernersprache – theoretische Annahmen (31)
  - 2.2.2.1. Regelhaftigkeit vs. Variabilität der Lernersprache (31)
  - 2.2.2.2. Sind Lernersprachen natürlich Sprachen? (33)
- 2.2.3. Der fortgeschrittene Lerner (34)
  - 2.2.3.1. Problem der Definition von Fortgeschrittenheit (35)
  - 2.2.3.2. Fortgeschrittenheit im Zusammenhang mit der Basisvarietät (36)
  - 2.2.3.3. Fortgeschrittenheit in dieser Arbeit (37)

## **2.3. Korpora, Lernerkorpora und das Falko-Korpus (38)**

2.3.1. Definition des Begriffes ‚Korpus‘ (38)

2.3.2. Lernerkorpora (39)

2.3.3. Falko – ein Lernerkorpus für DaF (41)

## **2.4. Fehleranalyse und Fehlerklassifikation (43)**

2.4.1. Was ist ein „Fehler“? (44)

2.4.2. Fehleridentifizierung, -beschreibung und -kategorisierung (46)

2.4.3. Entwicklung eines Fehlertagsets (48)

## **2.5 Zwischenfazit (48)**

# **3. Empirischer Teil (50)**

## **3.1. Quantitative Analyse – Produktivität der Wortbildungsmuster der -ung-Suffigierung (50)**

3.1.1. Auswahl, Kategorisierung und Zählung der Belege (50)

3.1.2. Auswertung der Belege und Berechnung der potential productivity (53)

3.1.3. Diskussion der Ergebnisse der quantitativen Analyse (55)

## **3.2. Qualitative Analyse – Untersuchung der Lernerfehler und Entwicklung des Fehlertagsets (56)**

3.2.1. Auswahl der Belege (56)

3.2.2. Fehler, Zielhypothese und Fehlerbeschreibung (59)

3.2.3. Kategorisierung der Fehler (70)

3.2.4. Ein mögliches Fehlertagset für Falko (72)

3.2.5. Diskussion der Ergebnisse der qualitativen Analyse (73)

# **4. Zusammenfassung und Ausblick (76)**

# **5. Quellenverzeichnis (78)**

## **5.1. Literaturverzeichnis (78)**

5.2. **Untersuchte DaF-Lehrwerke (83)**

5.3. **Internet-Links (84)**

## **6. Anhang (85)**

**6.1. Abstract (85)**

**6.2. Eigenständigkeitserklärung (86)**

**6.3. Lebenslauf (87)**

### **Abbildungsverzeichnis**

**Abb. 1:** Wortbildungsmuster der *-ung*-Suffigierung nach Motsch (1999) (17)

**Abb. 2:** Wortbildungsmuster der *-ung*-Suffigierung mit Produktivitätsangaben nach Motsch (26)

**Abb. 3:** Beispiel für die Metainformation zu den Lernertexten im Falko-Subkorpus „Essays L2“ (43)

**Abb. 4:** Beispiel für die Metainformation zu den Textautoren im Vergleichskorpus „Essays L1“ (43)

**Abb. 5:** Auszählung der Belege für die *-ung*-Suffigierung und Berechnung der potential productivity bei den Lernern (54)

**Abb. 6:** Auszählung der Belege für die *-ung*-Suffigierung und Berechnung der potential productivity bei den Muttersprachlern (54)

**Abb. 7:** Auflistung der Lerner-Fehler mit Quellenangabe, Zielhypothese und Fehlerbeschreibung (59-70)

**Abb. 8:** Fehlertagset für die Fehler der *-ung*-Suffigierung (73)

### **Verzeichnis der Abkürzungen**

WBK = Wortbildungskonstrukte

DaF = Deutsch als Fremdsprache

SM = Semantische Muster

EA = Error Analysis

L1 = Muttersprache eines Sprechers

L2 = Fremdsprache

BV = Basisvarietät

## 1. Einleitung

In seinem Einführungswerk für Studenten und Lehrer des Faches Deutsch als Fremdsprache schreibt Günther Storch über die Wortbildung Folgendes: „Obwohl dieser Bereich der Sprache eine wichtige Möglichkeit darstellt, den Wortschatz und somit das Ausdrucksvermögen der Lernenden auf eine sehr ökonomische Art zu erweitern, werden die damit verbundenen Möglichkeiten in den meisten DaF-Lehrwerken nicht ausgeschöpft“.<sup>1</sup> Damit scheint der Nagel auf den Kopf getroffen.

Wortbildung, gemeint ist hier der Bereich der Derivation im Gegensatz zur Flexion - also produktive Muster, die der Produktion und Rezeption von morphologisch komplexen Wörtern dienen - kommt selbstverständlich in fast allen DaF-Lehrwerken vor. Dennoch ist die Darstellung dieses Bereichs der Grammatik, sofern er überhaupt Erwähnung findet<sup>2</sup>, oft stark vereinfacht und verkürzt, in der Regel wird Wortbildung erst ab der Mittelstufe explizit behandelt, obwohl laut Storch „das Wortbildungspotenzial des Wortschatzes von der ersten Lektion eines Deutschkurses an sehr groß ist – und somit auch die Möglichkeit, ab der ersten Unterrichtsstunde Wortbildung zu betreiben“.<sup>3</sup> Dieser Eindruck hat sich mir bei einer kurzen, keineswegs den Anspruch auf Repräsentativität oder Vollständigkeit erhebenden Durchsicht einiger gängiger Lehrwerke wie „Delfin“, „em“, „Deutsch aktiv Neu“, „Wege“, „Themen neu“ oder „Berliner Platz“<sup>4</sup> ebenfalls bestätigt. Als kurzes Beispiel sei hier die Darstellung aus dem Lehrbuch „em. Hauptkurs. Deutsch als Fremdsprache für die Mittelstufe“ erwähnt. Darin wird in einem Grammatik-Kapitel über das Nomen lediglich erläutert, dass es den Prozess der Nominalisierung gibt, der „eine andere Wortart zum Nomen“ macht, und darauf folgt eine Übersicht über „Endungen deutscher Nominalisierungen“, in der unter anderem die Endung „-ung“ aufgelistet ist mit dem Beispiel „*Bedeutung (bedeuten)*“.<sup>5</sup> Weitere Erklärungen oder explizite Regeln finden sich hier nicht. Ganz ähnlich sieht es auch in den anderen Lehrwerken für die Grund- und Mittelstufe aus.

Diese Feststellung soll nun keine harsche Kritik an den Lehrwerken bedeuten; der Umfang, in welchem Wortbildung dort behandelt wird, mag didaktisch sehr wohl überlegt sein. Auch Storch räumt ein, dass „(u)nter Sprachlehr- und -lernaspekten (...) Wortbildung ein schwieriger Bereich (ist), da es viele Irregularitäten und Polyfunktionalitäten gibt“.<sup>6</sup> Und weiter: „Das Regelhafte kann gelernt und geübt werden; für das Irreguläre müssen die Lernenden im produktiven Bereich nach

---

<sup>1</sup> Storch 2001: 91

<sup>2</sup> Selbst die sehr umfassende „Deutsche Grammatik“ von Helbig und Buscha hat diesen Bereich ausgeklammert.

<sup>3</sup> Storch 2001: 91

<sup>4</sup> Vgl. Literaturliste im Anhang

<sup>5</sup> Perlmann-Balme 2000: 82

<sup>6</sup> Storch 2001: 92

und nach ein Sprachgefühl entwickeln“. Deshalb, so Storch, „sollten die Lernziele im Bereich der Wortbildung nicht zu hoch angesetzt werden“. Für Anfänger gilt, dass „die Lernziele (...) vor allem im rezeptiven und reproduktiven Bereich liegen“.<sup>7</sup>

Doch wie sieht es bei fortgeschrittenen Lernern aus? Die Fähigkeit, Wortbildungsregularitäten zu erkennen und für die Rezeption und Produktion von komplexen Wörtern zu nutzen, gehört nach Wolfgang Motsch zur Sprachkompetenz eines Muttersprachlers. Wortbildungsmuster sind ihm zu Folge „Bestandteil der Lexikonkomponente der Grammatik“.<sup>8</sup> Von weit fortgeschrittenen Lernern wäre deshalb eventuell zu erwarten, dass sie diese Kompetenz ebenfalls ausgebildet haben. Darüber hinaus sollten sie über ein Lexikon verfügen, in dem sie die nicht mehr transparenten Lexeme mit idiosynkratischen Eigenschaften, welche also gerade nicht durch die Regeln der Wortbildung zu erklären sind, gespeichert haben.

Es gibt nach meinem Erkenntnisstand nur wenige Untersuchungen über den Erwerb von derivationsmorphologischen Wortbildungsregularitäten bei Lernern des Deutschen als Fremdsprache (DaF-Lerner). Dies mag unter anderem damit zusammenhängen, dass in der Zweitspracherwerbsforschung traditionell der frühe Zweitspracherwerb im Mittelpunkt des Interesses stand, auch wenn diese Fokussierung seit den 1990er Jahren nachgelassen hat.<sup>9</sup> Zu untersuchen, wie fortgeschrittene Lerner mit den produktiven Möglichkeiten der Wortbildung umgehen, erscheint mir daher gleichsam spannend wie lohnend.

Nähert sich die Sprachkompetenz der Lerner in diesem Bereich an die Kompetenz der Muttersprachler an? Gehen die Regeln und Beschränkungen, die für die produktiven Muster der Wortbildung gelten, tatsächlich irgendwann ins Sprachgefühl über? Oder bleibt der Bereich der produktiven Wortbildung bis zuletzt schwierig und stark anfällig für Fehler? Welche Art von Fehler machen die Lerner, wenn sie selbst Texte produzieren? Weit fortgeschrittene Lerner produzieren oft nur wenige falsch gebildete Wörter oder ungrammatische Konstruktionen – trotzdem unterscheidet sich ihr Sprachgebrauch nachweislich von dem der Muttersprachler, indem sie etwa bestimmte Konstruktionen signifikant häufiger oder seltener benutzen. Wie ist der Sprachgebrauch fortgeschrittener Lerner in Bezug auf produktive Wortbildungsmuster?

Mit diesen Fragen beschäftigt sich die vorliegende Arbeit. Um den Rahmen nicht zu sprengen, beschränkt sich die Untersuchung auf jene produktiven Wortbildungsmuster, bei denen Verben durch Anhängen des Suffixes „-ung“ substantiviert werden. Dafür werden Daten aus dem fehlerannotierten Lernerkorpus Falko analysiert. Es handelt sich dabei um frei verfasste Essays zu vier kontroversen Themen, die von fortgeschrittenen DaF-Lernern geschrieben wurden. Diesem

---

<sup>7</sup> Storch 2001: 93

<sup>8</sup> Motsch 1999: 1

<sup>9</sup> Vgl. Walter 2008: 3

Subkorpus von Falko ist ein L1-Korpus zugeordnet, das Texte von Berliner und Brandenburger Gymnasiasten enthält und als muttersprachliches Vergleichskorpus dient.<sup>10</sup>

Der Ansatz der Korpusanalyse ist sowohl quantitativ als auch qualitativ, wobei der qualitativen Analyse deutlich mehr Raum zukommen wird. Quantitativ ist die Untersuchung der Produktivität. Es wird verglichen, ob die Lerner von den zu untersuchenden Wortbildungsmustern mehr oder weniger Gebrauch machen als Muttersprachler, bzw. sie gleich häufig verwenden. Die qualitative Analyse der Lernerdaten beschäftigt sich mit den fehlerhaften Wortbildungen der Lerner. Im Anschluss an die Fehleranalyse wird eine Fehlerklassifikation vorgenommen und ein mögliches Fehlertagset für den zukünftigen Gebrauch in Falko erarbeitet.

Im zweiten Kapitel dieser Arbeit wird der theoretische Hintergrund der Untersuchung dargelegt. Dabei soll anhand von traditionellen Grammatiken und Motsch (1999) zunächst aufgezeigt werden, wie die Wortbildung nach einem produktiven Muster abläuft, welchen Beschränkungen das Muster unterliegt und welche syntaktischen und morphosyntaktischen Eigenschaften den Wortbildungsprodukten zukommen. Zudem werden alle wichtigen Begriffe und Konzepte aus den Bereichen der Wortbildung, der Zweitspracherwerbsforschung und der Korpuslinguistik vorgestellt und diskutiert, die für die anschließende Korpusanalyse von Bedeutung sind. Des Weiteren wird näher ausgeführt, welche Möglichkeiten die Verwendung von Sprachkorpora und im Besonderen von Lernerkorpora bietet, was genau das Lernerkorpus Falko ist und warum ich mich in dieser Arbeit für eine Analyse von Falko-Daten entschieden habe.

Kapitel drei ist der Kern dieser Arbeit, die Korpusanalyse. Zunächst wird erläutert, wie die Datengrundlage für die Analyse zustande kam, d.h. welche Suchabfrage an Falko gestellt wurde und welche Belege als für die Untersuchung relevant eingestuft wurden. Diese Datengrundlage dient in einem ersten Schritt der quantitativen Analyse, der Messung der Produktivität der Wortbildungsmuster der *-ung*-Suffigierung im Gebrauch von deutschen Muttersprachlern und fortgeschrittenen DaF-Lernern mittels der Formel zur Messung der „potential productivity“ von Rolf Baayen.<sup>11</sup> Darauf folgt die qualitative Analyse, in der es darum geht zu ermitteln, welche Fehler die fortgeschrittenen Lerner bei der Bildung von Wörtern nach einem produktiven Muster machen. Auf der zuvor beschriebenen theoretischen Basis erfolgt die Auswertung der Falko-Daten, d.h. die Beschreibung der von den Lernern gemachten Fehler und die anschließende Kategorisierung der Fehler je nach Fehlerart. Diese Fehlerklassifikation dient als Grundlage für die Erstellung eines Fehlertagsets für Falko. Damit soll am Beispiel der *-ung*-Suffigierung ein Vorschlag gemacht werden, wie die Falkodaten unter morphologischen Kriterien für den Bereich

---

<sup>10</sup> Vgl. Lüdeling et al. 2008, sowie Abschnitt 2.3.3. dieser Arbeit

<sup>11</sup> Vgl. Baayen 2009

der Derivation getaggt werden könnten.

Eine abschließende Zusammenfassung der Ergebnisse dieser Arbeit sowie ein Ausblick auf sich anschließende Fragestellungen und Forschungsdesiderata bietet das Kapitel vier.

## 2. Theoretischer Hintergrund

In diesem Kapitel geht es um den theoretischen Hintergrund, vor welchem die anschließende Korpusanalyse stattfindet. Die linguistischen Teildisziplinen, welche in dieser Arbeit berührt werden, sind im Wesentlichen die Korpuslinguistik, die Zweitspracherwerbsforschung und ganz allgemein die Wortbildungslehre.<sup>12</sup> Ich möchte bei der theoretischen Darstellung der Wortbildung (Abschnitt 2.1.) nicht allzu weit ausholen<sup>13</sup>, sondern in erster Linie auf die *-ung*-Nominalisierung eingehen. Des Weiteren werden zentrale Begriffe und Konzepte dieser Arbeit wie „Wortbildungsmuster“ und „(morphologische) Produktivität“ vorgestellt. Die folgenden Abschnitte liefern eine Diskussion der Begriffe „Lernersprache“ und „fortgeschrittener Lerner“ (Abschnitt 2.2.), „Korpora“ und im Besonderen „Lernerkorpora“ (Abschnitt 2.3.) sowie „Fehleranalyse und Fehlerklassifikation“ (Abschnitt 2.4.). Alle diese Begriffe werden in der linguistischen Literatur mit zum Teil recht unterschiedlichen Bedeutungen oder Konnotationen verwendet und es soll deutlich werden, wie sie für die vorliegende Arbeit zu verstehen sind und welche theoretischen Implikationen damit verknüpft sind. In Abschnitt 2.5. werden in einem kurzen Zwischenfazit die wichtigsten Erkenntnisse aus diesem Kapitel zusammengefasst.

### 2.1. Zur Theorie der Wortbildung

#### 2.1.1. Warum überhaupt Wortbildung?

Die Frage, ob es für die Sprecher einer Sprachgemeinschaft notwendig ist, immer wieder neue Wörter zu erschaffen, lässt sich ohne zu zögern mit „ja“ beantworten. Es besteht aus vielerlei Gründen ein stetiger, hoher Bedarf an Neuschöpfungen. Erben (2000) nennt zunächst einmal die Tatsache, dass

„die geistig-sprachliche Auseinandersetzung mit der vielseitigen und veränderlichen Wirklichkeit weitergeht, und nach wie vor die kommunikative Notwendigkeit besteht, alles, was man kennen lernt oder lehrt, auch **nennen** zu müssen – nicht nur das, was man vorfindet oder im Laufe der Forschung herausfindet, sondern auch das, was man empfindet, erfindet oder entwirft, was man ahnt und was man (...) plant“.<sup>14</sup>

Neben dieser Benennungsfunktion der Wortbildungskonstrukte (WBK) gibt es noch eine Reihe

---

<sup>12</sup> „Wortbildungslehre ist derjenige Teil der Grammatik, der die Wortbildung, die Bildung neuer Wörter unter wissenschaftlichen oder praktischen Gesichtspunkten darstellt und dadurch sowohl angemessene Urteile über Wortbildungsprozesse und ihre Bedingungen ermöglicht, als auch über Wortbildungsergebnisse, die Struktur und Funktion vorhandener und möglicher Wörter.“ Erben 2000: 16

<sup>13</sup> Zur Forschungsgeschichte der Wortbildungslehre vgl. Kapitel I bei Erben (2000).

<sup>14</sup> Ebd. S. 19f.

weiterer Gründe, warum neue Wörter – mit bereits vorhandenem Wortmaterial – gebildet werden.<sup>15</sup> Dazu gehören subjektive Ausdrucksnotwendigkeiten wie das Bedürfnis, eine Sache mit einem neuen Ausdruck zu benennen, etwa weil der alte Ausdruck dem Sprecher zu abgenutzt oder nicht expressiv genug erscheint, oder ähnliches. Vor allem aber sind es syntaktische Notwendigkeiten, die „Einpassung lexikalischer Einheiten in verschiedene Kontexte und Konstruktionen“, die die Bildung neuer Wörter erforderlich machen:

„Wortbildung ermöglicht nämlich u. a. den Wechsel der Wortklasse (Transposition) (...), sie vermehrt also die syntaktischen Einsatzmöglichkeiten der Grundmorpheme. (...) (D)enn Wortbildung hilft mit, die grammatische Umstrukturierung (Transformation) einer Aussage herbeizuführen, (...) die situationell oder kontextuell angemessener ist“.<sup>16</sup>

Dies trifft für die *-ung*-Suffigierung sicher in hohem Maße zu, wo die Wortbildung beispielsweise die Alternative schafft zwischen Formulierungen wie *Das Flugzeug landete um 16 Uhr* oder *Die Landung des Flugzeugs erfolgte um 16 Uhr*.

### 2.1.2. Wortbildung verstanden als Derivation

Wenn in dieser Arbeit von „Wortbildung“ die Rede ist, dann meint das immer die Derivation im Gegensatz zur Flexion. Fleischer unterscheidet die „Bildung von 'Wortstämmen' wie *Edel-stein*, *stein-ig*, *ver-stein-er(n)*, und 'Wortformen-Bildung', wie *Stein*, *Stein-(e)s*, *Steine*“.<sup>17</sup> Zwar gibt es bestimmte Phänomene wie etwa die Komparation der Adjektive, wo „die Grenze zwischen Flexion und Wortbildung (im engeren Sinn: Derivation) fließend ist“<sup>18</sup>, aber das in dieser Arbeit verhandelte Phänomen, die Substantivierung von Verben durch *-ung*-Suffigierung, kann eindeutig als Bildung eines Wortstammes verstanden werden. Mit Fleischer ist Wortbildung „die Produktion von Wörtern (Wortstämmen) auf der Grundlage und mit Hilfe vorhandenen Sprachmaterials“ und damit klar abgegrenzt von der Wortschöpfung, welche darin besteht, „daß Wörter aus Lautkomplexen 'geschaffen' werden, die in der Sprache (noch) nicht als bedeutungstragende Elemente (Zeichen) vorhanden sind“.<sup>19</sup>

Nach Fleischer gibt es drei verschiedene Verfahren der Wortbildung, das semantische, das syntaktische und das morphologisch-strukturelle.<sup>20</sup> Letzteres „greift in die Wortstruktur ein – durch Veränderung der Formativstruktur (Erweiterung, Reduktion, sonstige Veränderung) in Verbindung mit semantischen und syntaktischen Prozessen. Dieses Verfahren wird gewöhnlich als die

<sup>15</sup> Vgl. Erben 2000: 19-24

<sup>16</sup> Ebd. S. 22

<sup>17</sup> Fleischer 1992: 3

<sup>18</sup> Ebd. S. 4

<sup>19</sup> Ebd. S. 5

<sup>20</sup> Vgl. ebd. S. 6f.

Wortbildung im eigentlichen oder engeren Sinne betrachtet“.<sup>21</sup> Zum morphologisch-strukturellen Verfahren zählt auch die Substantivierung von Verben durch Suffigierung mit dem Suffix *-ung*.

### 2.1.3. Die *-ung*-Suffigierung

Mit Erben (2000) verstehen wir „komplexe, mehrgliedrige Wörter als Morphemgefüge“.<sup>22</sup> Morpheme sind hiernach die „kleinsten Zeichen des Sprachsystems, die nicht nur einen Lautwert haben, sondern eine semantische Funktion“.<sup>23</sup> Nach – vor allem – Unterschieden in der Semantik lassen sich die Morpheme in zwei Gruppen einteilen: zum einen die 'Grundmorpheme', die einen hohen Inhaltswert haben und entweder als freie Morpheme, also selbständige Lexeme (Einheiten des Lexikons), vorkommen oder als 'Kernmorphem' „die Basis satellitenhafter Affixe bilden“<sup>24</sup> wie etwa die Verbstämme in *Turn-er* oder in *Be-werb-ung*. Zum anderen gibt es die Gruppe der *Formantien*, also Affixe, die nicht wort- oder basisfähig sind und „mehr grammatischen als semantischen Signalwert“<sup>25</sup> haben, wenngleich auch inhaltliche Merkmale angezeigt werden können wie zum Beispiel die Bezeichnungsklasse „Handlung“ in *Erober-ung* oder *Beweg-ung*.

Die *-ung*-Suffigierung bezeichnet ein morphologisch-strukturelles Verfahren der Wortbildung, genauer gesagt der Transposition, bei dem aus Verben Substantive gebildet werden. Diese Nomen haben immer das Genus Femininum. Die Bildung erfolgt, indem das Suffix *-ung* an einen Verbstamm angehängt wird, die Verbalstämme können prinzipiell einfach oder komplex sein. Jedoch kann nicht von jedem Verb ein entsprechendes Derivat gebildet werden; die Beschränkungen und Blockierungen, die hier wirksam sind, werden weiter unten noch ausführlicher besprochen.

Draeger (1996) weist darüber hinaus darauf hin, dass die Basis in seltenen Fällen auch ein Partizip sein kann, das eine andere Bedeutung hat als die Grundbedeutung des ihm zugrundeliegenden Verbs. Als Beispiel führt sie das suffigierte Substantiv *Veranlagung* (im Sinne von „charakterlichem Merkmal“) an, das aufgrund seiner Semantik von dem Partizip *veranlagt* abgeleitet sein muss und nicht von dem Verb *veranlag* (in seiner Bedeutung „Steuern erheben“).<sup>26</sup>

### 2.1.4. Ist das Suffix *-ung* eine Lexikoneinheit?

---

<sup>21</sup> Fleischer 1992: 7

<sup>22</sup> Erben 2000: 26

<sup>23</sup> Ebd. S. 25

<sup>24</sup> Ebd. S. 26

<sup>25</sup> Ebd. S. 26

<sup>26</sup> Vgl. Draeger 1996: 88

Sind Morpheme – also auch Affixe – oder nur Wörter Grundeinheiten des Lexikons? Zu dieser Frage gibt es in der Literatur verschiedene Auffassungen. Olsen (1986) etwa versteht Affixe als Lexikoneinheiten:

„Wie die 'freien' lexikalischen Einheiten der Sprache werden auch Affixe mit einem vollspezifizierten Lexikoneintrag im Lexikon verzeichnet. Eine affixale Eintragung umfasst die phonologische Charakterisierung (PC) des Affixes, die Charakterisierung seiner Kategorie (KC) und seiner Semantik (SC)“.<sup>27</sup>

Eine solche Annahme impliziert, dass Affixe als phonologische Form und semantische Muster als semantische Repräsentation einer gebundenen Lexikoneinheit betrachtet werden. Dem gegenüber steht die Argumentation von Motsch (1999), der ausführt, dass diese Annahme problematisch ist in den Fällen, wo die morphonologische Indizierung leer bleibt (z.B. bei der Konversion *laufen* --> *das Laufen*). In diesen Fällen müsste man annehmen, dass „elementare sprachliche Zeichen lediglich aus einer semantischen Repräsentation bestehen könnten, d.h. keine Lautform haben müssen“.<sup>28</sup> Ich nehme deshalb mit Motsch an, dass Derivationsaffixe keine Einheiten des primären Lexikons sind.<sup>29</sup> Sie haben zwar eine phonologische Form und Suffixe bestimmen als Heads von Wortbildungskonstruktionen auch deren Wortart, aber sie haben keine Bedeutung im eigentlichen Sinne.

„Um den Unterschied zwischen Lexikoneinheiten und Derivationsaffixen zu verdeutlichen, sagen wir nicht, ein Affix habe eine Bedeutung, sondern **ein Affix indiziert ein semantisches Muster**. Ein Affix ist natürlich umso stärker mit einem semantischen Muster assoziiert, je weniger semantische Muster es indiziert. Und es ist umso weniger mit Bedeutungen assoziierbar, je mehr semantische Muster es indiziert“.<sup>30</sup>

Zur Veranschaulichung sei hier nur ein kurzes Beispiel angeführt: Das Suffix *-tel*, das an Kardinalzahlen angehängt werden kann (z.B. *Viertel*, *Fünftel*, usw.) indiziert beispielsweise nur ein einziges Muster, es ist produktiv in der Bildung von Bruchzahlen. Hier könnte man, wie Fleischer das tut, davon sprechen, dass „Bruchzahl“ die „Wortbildungsbedeutung“ ist.<sup>31</sup> Das Suffix *-ung* dagegen indiziert eine Vielzahl semantischer Muster, es seien hier erstmal nur zwei davon angeführt: Die Bildung von „nomina actionis“ versteht Motsch als reine Nominalisierung, „d.h. nur die syntaktische Kategorie wird verändert, die semantische Repräsentation des Verbs (...) bleibt erhalten“<sup>32</sup>, wie zum Beispiel bei der Nominalisierung des Verbs *erziehen* durch *-ung*-Suffixierung zu *Erziehung*. Der semantische Gehalt des Verbs, nämlich einen Vorgang zu beschreiben, wohnt auch dem Nomen inne. Das semantische Muster nimmt also lediglich eine Umkategorisierung vor.

<sup>27</sup> Olsen 1986: 75

<sup>28</sup> Motsch 1999: 16

<sup>29</sup> Eine ähnliche Position vertritt auch Draeger (1996).

<sup>30</sup> Motsch 1999: 16

<sup>31</sup> Vgl. Fleischer 1992: 170

<sup>32</sup> Motsch 1999: 321

Anders sieht es etwa bei dem derivierten Nomen *Verblüffung* aus. Hier drückt das Nomen das Resultat eines Vorgangs aus, der durch das Verb *verblüffen* bezeichnet wird. Es ist also ein anderes semantisches Muster anzunehmen. Weitere semantische Muster können durch das Suffix *-ung* indiziert sein, eine ausführliche Beschäftigung mit den semantischen Mustern folgt weiter unten.

### 2.1.5. Wortbildungsmuster

Bei den folgenden Ausführungen beziehe ich mich im Wesentlichen auf Motsch (1999), wo Wortbildungsmuster<sup>33</sup> als „Paare von semantisch-syntaktischen und phonologisch-morphologischen Beschreibungen“<sup>34</sup> dargestellt werden. „Wortbildungsmuster sind somit Muster für die Analyse komplexer Wortzeichen mit je spezifischen Laut- und Bedeutungseigenschaften“. Motsch nimmt an, dass sie psychische Realität<sup>35</sup> besitzen und dem Sprecher zumindest unbewusst zur Verfügung stehen:

„Wortbildungsmuster gehören nicht zum selbständigen Sprachwissen, wie syntaktische und phonologische, z.T. auch flexionsmorphologische Regeln, sondern es handelt sich um Verallgemeinerungen von Eigenschaften zum Lexikon gehörender Wörter, die die Sprecher einer Sprache vornehmen können. Zum Lexikon gehörende Wörter sind aus psychologischer Sicht voll spezifizierte Einheiten. Werden sie in aktuellen Redesituationen verwendet, müssen keine Wortbildungsmuster herangezogen werden. Von Wortbildungsmustern muss aber Gebrauch gemacht werden, wenn ein Sprecher ein neues Wort bildet oder ein Hörer sich ein neues Wort verständlich machen will. In diesen Fällen werden mit dem neuen Wort vergleichbare Lexikoneinheiten herangezogen und das ihnen zugrundeliegende Wortbildungsmuster aktiviert.“<sup>36</sup>

Hier ist schon ein wichtiger Punkt angesprochen, auf den es später noch einmal zurückzukommen gilt, nämlich die Potenz von Wortbildungsmustern, für die Produktion und Rezeption neuer Wörter nutzbar zu sein. Doch zunächst soll etwas anderes näher ausgeführt werden: die zentrale Rolle des Lexikons. Wortbildungsmuster werden gleichsam aus dem Lexikon gewonnen und sind dort aktiv. Ich nehme mit Plag (1999: 9-11) die Konzeption eines Lexikons an, in dem nicht alle komplexen Wörter bei jeder Aktualisierung neu gebildet werden müssen.<sup>37</sup> Zum großen Teil sind sie als fertig

<sup>33</sup> Fleischer (1992) verwendet den Begriff „Wortbildungsmodell“

<sup>34</sup> Motsch 1999: 1

<sup>35</sup> Vgl. auch Olsen (1986), wo Wortbildungsmuster im Rahmen einer generativen Grammatik untersucht werden

<sup>36</sup> Motsch 1999: 22

<sup>37</sup> Plag (1999: 9-11) liefert eine kurze Zusammenfassung der theoretischen Diskussion um die Konzeption des Lexikons im Hinblick auf die Aktualisierung komplexer Wörter. Kurz gesagt gibt es zwei Extrempositionen: die eine nimmt an, dass alle Wörter (einfache und komplexe) als einzelne Lexeme im Lexikon gespeichert sind („storage“) und bei jeder Aktualisierung als Ganzes abgerufen werden; die andere Position nimmt an, dass nur die kleinsten morphologischen Einheiten im Lexikon gespeichert werden und komplexe Wörter bei jeder Aktualisierung neu zusammengesetzt werden müssen („parsing“). Vertreter beider Positionen argumentieren damit, dass der jeweilige Prozess ökonomischer sei, also weniger Leistung des Gehirns beanspruche. Plag liefert Argumente gegen beide Extrempositionen und plädiert für eine Lexikonkonzeption, die einen Mittelweg darstellt: „(T)he conflict of the two models can be resolved by assuming a morphological race model, in which both ways of accessing a

abrufbare Einheiten im Lexikon gespeichert. Diese Einheiten haben sowohl systematische (und damit vorhersagbare), als auch idiosynkratische (also von der individuellen Existenz eines Wortes abhängige) Eigenschaften. Aus den systematischen Eigenschaften, die den Wortbildungen gemein sind, lassen sich Wortbildungsmuster extrahieren, die gleichsam wiederum für die Analyse der WBK im Lexikon genutzt werden. Es lässt sich nur spekulieren, welche Konsequenzen diese Dualität der Mechanismen zur Produktion von Wörtern (Aktivierung eines Lexikoneintrages vs. Bildung mittels produktivem Wortbildungsmuster) für die Bewertung von Lerneräußerungen hat. Ich will dazu einige Überlegungen in Abschnitt 2.1.5.2. anschließen.

Das Lexikon ist „die zentrale Schaltstelle der Grammatik, eines Mechanismus, der die Abbildung von Bedeutungsstrukturen auf Lautstrukturen regelt“.<sup>38</sup> Dazu ist nach Motsch die Annahme einer eigenen Wortsyntax zur Bildung komplexer Wörter hinfällig, es genügen hierfür allein die Wortbildungsmuster mit ihren jeweiligen Beschränkungen.

„Wortbildungsmuster sind (...) Bestandteile der Lexikonkomponente der Grammatik (...): 1. Sie beschreiben komplexe Lexikoneinheiten. 2. Sie verlangen keine syntaktischen Regeln, die die Kombination elementarer Einheiten vermitteln, sondern nehmen die Abbildung von Bedeutungsstrukturen auf Lautstrukturen direkt vor.“<sup>39</sup>

Ausgangspunkt eines Wortbildungsprozesses sind also immer die Bedeutungsstrukturen, d.h. die semantischen Muster. (Es sei hier noch auf eine wichtige terminologische Unterscheidung hingewiesen: das 'semantische Muster' ist nicht dasselbe wie die 'semantischen Repräsentation' einer Wortbildung. Vielmehr gehört zur semantischen Repräsentation sowohl das „semantische Muster für Wortbildungen als auch die semantischen Repräsentationen der Lexikoneinheiten, die in das Muster eingehen“).<sup>40</sup> Doch obwohl die Rolle des semantisch-syntaktischen Teils hier im Vordergrund steht, gilt es zu beachten, dass der phonologisch-morphologische Teil von Derivationen durchaus auch auf das semantische Muster zurückwirken kann, indem er bestimmte Bildungen, die das semantische Muster vorhersagt, blockiert. Ein beliebtes Beispiel dafür ist die Distribution der Suffixe *-heit*, *-keit* und *-igkeit*.<sup>41</sup> Die Beschränkungen, die für die semantischen Muster gelten, welche durch das Suffix *-ung* indiziert werden, werden in Abschnitt 2.1.7. noch eingehend behandelt.

### **2.1.5.1. Wortbildungsmuster (und semantische Muster) der *-ung*-Suffigierung**

complex form are in competition. The winner of the race is determined mainly by the frequency and the phonological and semantic transparency of the word to be accessed“ (Plag 1999: 10). Danach werden transparente, aber wenig frequente Wörter am ehesten durch parsing zusammengesetzt, für stark frequente Wortformen oder solche, die opak sind, ist es am wahrscheinlichsten, dass sie direkt als Ganzes aus dem Lexikon abgerufen werden.

<sup>38</sup> Motsch 1999: 3

<sup>39</sup> Ebd. S. 1

<sup>40</sup> Motsch 1999: 3f.

<sup>41</sup> Vgl. Fleischer 1999: Kapitel 2.3.2.7.

In Abschnitt 2.1.4 wurde bereits angedeutet, dass das Suffix *-ung* mehrere verschiedene semantische Muster indiziert. Wie viele man letztendlich annimmt, ist theorieabhängig und variiert stark von Autor zu Autor. Wellmann (1975) unterteilt die *-ung*-Suffigierungen in sieben Gruppen, Draeger (1996) sogar in 14, wobei viele dieser Gruppen nur sehr wenige lexikalisierte Mitglieder haben. Motsch (1999) dagegen nimmt nur vier grundlegende semantische Muster an, was eine größere Abstraktion bedeutet, und es ist klar, dass sich nicht alle belegten *-ung*-Derivate einer der vier Gruppen zuordnen lassen. Der überwiegende Teil des Wortschatzes ist jedoch abgedeckt, die wenigen Lexeme, die keiner der Gruppen zuzuordnen sind, sind vernachlässigbar. Ich schließe mich hier der Aufteilung von Motsch an, die mir sowohl sinnvoll als auch praktikabel für mein weiteres Vorgehen erscheint, denn im Analyseteil dieser Arbeit wird für jedes dieser semantischen Muster getrennt ermittelt, wie produktiv es für die Bildung von *-ung*-Suffigierungen jeweils von fortgeschrittenen Lernern und deutschen Muttersprachlern genutzt wird.

Die Wortbildungsmuster für *-ung*-Suffigierungen haben bei Motsch (1999) die gemeinsame Form **[SM; PF<sub>v</sub>-ung, Femininum]**, welche so zu verstehen ist, dass mit Zugrundelegung eines semantischen Musters (SM) aus der phonologischen Form eines Verbs (Pf<sub>v</sub>) durch Anhängen des Suffixes *-ung* ein Nomen mit dem Genus Femininum gebildet wird.

Nach Motsch existieren also vier Wortbildungsmuster für *-ung*-Suffigierungen, die zwar alle morphologisch identisch sind, aber auf verschiedenen semantischen Mustern basieren<sup>42</sup>, die im Folgenden aufgeführt sind.

#### 1. Reine Nominalisierung (Bildung von „nomina actionis“):

Hier wird ein semantisches Muster vorausgesetzt, das lediglich eine Umkategorisierung vornimmt. Nur die syntaktische Kategorie wird verändert (aus einem Verb wird ein Nomen), die semantische Repräsentation des Verbs bleibt erhalten, d.h. die Nominalisierungen drücken ein 'Geschehen' aus und können dieses je nach Kontext „als Tatsache, als Verlauf und als generelles Geschehen“<sup>43</sup> charakterisieren. Beispiele hierfür sind nomina actionis wie *Erstürmung*, *Beseitigung*, *Ausgrabung*.

#### 2. Resultierender Zustand:

Hier liegt ein semantisches Muster zugrunde, welches die WBK semantisch festlegt als Zustände, die Resultat eines Geschehens sind, das durch das Basisverb ausgedrückt wird. Diese Bildungen bezeichnen besonders psychische und physische Zustände von Personen, wie zum Beispiel

---

<sup>42</sup> Vgl. Motsch 1999: Kapitel 4: Nomenbildung

<sup>43</sup> Ebd. S. 327

*Verzückung, Ernüchterung, Begeisterung.*

3. Thema eines Geschehens (Bildung von „nomina acti“):

Diese Bildungen bezeichnen physikalische oder abstrakte und ideelle Gegenstände in einem vom Verb bezeichneten Geschehen, zum Beispiel *Lieferung, Abbildung, Erfindung*. Da viele dieser Bildungen auch als nomina actionis vorkommen, kann nur mit Hilfe des Kontextes interpretiert werden, welches semantische Muster zugrunde liegt.

4. Institutionen oder Ort eines Geschehens (nomina loci):

Die Bildungen nach diesem semantischen Muster sind „Institutionen (...), deren Zweck es ist, die Tätigkeit des Basisverbs auszuüben“<sup>44</sup>, z.B. *Regierung*, oder Orte, an denen die Tätigkeit des Basisverbs ausgeübt wird, z.B. *Wohnung, Siedlung, Mündung*.

Die folgende Tabelle stellt diese vier Wortbildungsmuster noch einmal übersichtlich nebeneinander und verwendet für die Beschreibung der semantischen Muster die Terminologie und Formulierungen aus Motsch (1999).

<b>Wortbildungsmuster</b>	<b>Semantisches Muster</b>	<b>Beispiele</b>
1) Nomina actionis	Reine Umkategorisierung: [V (x <sup>1</sup> , x <sup>2</sup> ..., s)] <sub>N</sub> (x <sup>1</sup> , x <sup>2</sup> ..., r)	<i>Erstürmung, Beseitigung, Ausgrabung</i>
2) Resultierender Zustand	[ZUSTAND & RESULTAT VON (ZUSTAND, V (x <sub>thema</sub> , s))] ((x <sub>thema</sub> ,) r) 'Referenten sind Zustände, die Resultat eines Geschehens V	<i>Verzückung, Ernüchterung, Begeisterung</i>

<sup>44</sup> Motsch 1999: 348

	sind, von dem ein Aktant x betroffen ist'	
3) Nomina acti	[GEGENSTAND & V (x <sub>agens</sub> , GEGENSTAND <sub>thema</sub> )] ((x <sub>agens</sub> )r) 'Referenten, die Gegenstände und Thema in einem vom Verb bezeichneten Geschehen sind'	<i>Lieferung, Abbildung, Erfindung</i>
4) Nomina loci	[INST & ZWECK (V (AGENS, THEMA))] (r) 'Referenten, die Institutionen sind, deren Zweck es ist, eine Tätigkeit V auszuüben' (Das Wortbildungsmuster charakterisiert neben Institutionen auch lediglich Orte, für die eine Tätigkeit typisch ist)	<i>Regierung, Wohnung, Siedlung, Mündung</i>

Abb. 1: Wortbildungsmuster der -ung-Suffigierung nach Motsch (1999)

### 2.1.5.2. Aktivierung eines Lexikoneintrages oder produktive Wortbildung?

Weiter oben in Abschnitt 2.1.5. wurde bereits auf die beiden unterschiedlichen Möglichkeiten, die Sprecher haben, wenn sie eine Äußerung produzieren, hingewiesen: sie können entweder im Lexikon gespeicherte Lexeme aktivieren oder komplexe Worte mit Hilfe eines produktiven Wortbildungsmusters bilden. Ein Problem hierbei ist, dass man prinzipiell nicht unterscheiden kann, auf welchem Wege ein Sprecher in einer konkreten Äußerung zur Aktualisierung eines Wortes gelangt ist. Für Muttersprachler lässt sich annehmen, dass opake und hochfrequente Wörter aus dem Lexikon abgerufen werden.<sup>45</sup> Doch bei weniger frequenten Wörtern ist es schon höchst unklar, zumal auch nicht mit Eindeutigkeit gesagt werden kann, welche Wörter denn eigentlich zum Lexikon einer Sprachgemeinschaft gehören und zwar in dem Sinne, dass jeder Sprecher sie abgespeichert hat.<sup>46</sup> Was in Bezug auf native Sprecher einer Sprache schon schwierig zu bewerten ist, scheint mir noch unmöglicher in Bezug auf die Lerner einer Sprache. Um plausibel erklären zu können, ob ein Lerner ein komplexes Wort aus dem Lexikon abgerufen oder produktiv gebildet hat, müssten folgende Fragen beantwortet werden können: welche Wörter umfasst das Lexikon der (fortgeschrittenen) Lerner? Wie frequent sind die einzelnen komplexen Wörter im Input der Lerner? Sind die hochfrequenten Wörter im Input der Lerner und im Input der Muttersprachler dieselben? Was ist mit Wörtern in Lerneräußerungen, die zwar korrekt gebildet sind, vom Lerner aber mit einer

<sup>45</sup> Vgl. Plag 1999: 9-11

<sup>46</sup> Zu dieser Problematik vgl. auch Abschnitt 2.1.7.4. dieser Arbeit

falschen Bedeutung verwendet werden? Wurden diese Wörter mit Hilfe der Wortbildungsmuster produziert oder aus dem Lexikon abgerufen, wo sie eventuell bereits mit einer falschen Bedeutung gespeichert waren? Anders gefragt: welche idiosynkratischen Eigenschaften von komplexen Wörtern, die ein Muttersprachler im Lexikon gespeichert hat, hat auch der Lerner in seinem Lexikon gespeichert, bzw. wann analysiert der Lerner ein komplexes Wort, dessen Bedeutungseigenschaften gar nicht mehr aus der Bedeutung der einzelnen Wortbausteine zusammengesetzt werden kann?

Da diese Fragen offen bleiben, lautet die Konsequenz für diese Studie, dass bei der Berechnung der Produktivität der Wortbildungsmuster, also bei der Zählung der *-ung*-Derivate, sowie bei der Fehleranalyse nicht versucht wird zu unterscheiden, wie diese Wörter produziert wurden. Es steht allerdings außer Frage, dass es natürlich zu den Lernaufgaben eines Lerners gehört, sowohl die Regeln für die produktive Wortbildung zu erwerben als auch ein Lexikon mit den Eintragungen jener komplexen Wörter, die idiosynkratische Eigenschaften haben.

### **2.1.6. Morphologische Produktivität**

Die Frage nach der Produktivität des morphologischen Prozesses der *-ung*-Suffigierung ist eine der zentralen Fragen dieser Arbeit. Deshalb ist eine Definition des Begriffs „morphologische Produktivität“ unumgänglich, wenngleich keineswegs eine triviale Aufgabe. Nicht ohne Grund hat Laurie Bauer dem ersten Kapitel seiner Monographie „Morphological Productivity“<sup>47</sup> folgendes Zitat von Mayerthaler vorangestellt: „'Produktivität' zählt zu den unklarsten Begriffen der Linguistik“. Auch wenn dieses Zitat aus dem Jahre 1981 stammt und sich seither viel getan hat in der Forschung über morphologische Produktivität, so ist nicht von der Hand zu weisen, dass eine klare, allgemein anerkannte Definition des Begriffes noch aussteht. Eine solche Definition kann von dieser Arbeit nicht geleistet werden und ist auch nicht ihr Ziel. Vielmehr soll im folgenden Abschnitt umrissen werden, mit welchen grundlegenden Vorstellungen über Produktivität hier gearbeitet wird. Bauer bietet folgende Definition an:

„'Productivity' deals with the number of new words that can be coined using a particular morphological process, and is ambiguous between the sense 'availability' and the sense 'profitability'. The availability of a morphological process is its potential for repetitive rule-governed morphological coining, either in general or in a particular well-defined environment or domain. Availability is determined by the language system, and any process is either available or unavailable, with no middle ground. It creates psychologically real distinctions between available ('living') and unavailable ('dead') processes, which can be tested in a number of ways. The profitability of a

---

<sup>47</sup> Bauer 2001

morphological process reflects the extent to which its availability is exploited in language use, and may be subject unpredictably to extra-systematic factors. Where a single morphological process has easily distinguishable meanings or sub-uses, these may be assessed independently for both availability and profitability.<sup>48</sup>

Übernehmen möchte ich hieraus zunächst die prinzipielle Feststellung, dass Produktivität etwas damit zu tun hat, wie viele neue Wörter mit einem bestimmten Wortbildungsmuster gebildet werden können. Bauers Unterscheidung in „Verfügbarkeit“ (availability) und „Ergiebigkeit“ (profitability) eines morphologischen Prozesses erscheint hinfällig mit Baayen (2009), der erläutert, dass es keine gänzlich unproduktiven Muster gibt, die überhaupt nicht mehr für neue Wortbildungen genutzt werden könnten.<sup>49</sup> Theoretisch könnte also von den Sprechern einer Sprache jedes Wortbildungsmuster zur Bildung neuer Wörter genutzt werden, es bleibt letzten Endes nur die Frage, ob und in welchem Maße dies geschieht. Dafür sind zum einen interne Faktoren des Sprachsystems verantwortlich, wie etwa die Zahl der verfügbaren Basen für einen morphologischen Prozess oder die semantischen und morphologischen Beschränkungen, die wirksam sind. Für Bauer gibt es daneben noch Faktoren, die außerhalb des Sprachsystems liegen und nicht vorhersagbar sind, wie etwa das individuelle Bedürfnis eines Sprechers, für seine kommunikativen Zwecke ein neues Wort zu bilden. Gegen Bauers Formulierung „may be subject unpredictably to extra-systematic factors“ gibt es offenbar berechtigten Widerspruch von Baayen, der zumindest andeutet, wie diese sehr komplexen, extra-systemischen Faktoren untersucht werden können und erste interessante Ergebnisse mit Hilfe von Korpusanalysen liefert.<sup>50</sup> Darauf näher einzugehen, würde an dieser Stelle allerdings zu weit führen.

Produktivität kommt in unserem Zusammenhang also Bauers Begriff der „profitability“ am nächsten, der nach meinem Verständnis allerdings zwei Aspekte enthält: ein Wortbildungsmuster kann erstens als produktiv bezeichnet werden, wenn nach diesem Muster bereits viele Worte gebildet wurden oder zweitens, wenn das Potential besteht, nach diesem Muster weitere neue, für die Sprachgemeinschaft akzeptable Wörter zu bilden.<sup>51</sup> Vor allem in letzterer Auslegung wird der Begriff der Produktivität für diese Arbeit interessant: nur unter der Annahme, dass die Wortbildungsmuster der *-ung*-Suffigierung noch Potential für Neuschöpfungen bieten, macht es

---

<sup>48</sup> Bauer 2001: 211

<sup>49</sup> „(A) distinction between totally unproductive rules or schemas on the one hand, and productive rules or schemas on the other hand, would imply that the stored exemplars of unproductive schemas would not be available for generalization, while the stored exemplars of productive rules would allow generalization. This is not only implausible, but also contradicts the well-documented finding that the schemas of irregular verb classes can serve as attractors in both synchrony and diachrony“. Baayen 2009: 911

<sup>50</sup> Vgl. Baayen 2009: Abschnitt 4.

<sup>51</sup> Vgl. Bauer 2001: 41 und auch Plag (1999) kommt zu diesem Schluss: „Having scrutinized the different criteria put forward in standard definitions of productivity, it can be stated that this notion boils down to the property of a given word formation process or affix to be used to derive a new word in a systematic fashion“ (S.22).

Sinn zu vergleichen, ob Muttersprachler und fortgeschrittene Lerner des Deutschen dieses Potential in unterschiedlichem Maße ausschöpfen. Diese Annahme wird gestützt durch die Literatur zur Wortbildung und so wird im folgenden Abschnitt zunächst einiges zusammengetragen, was in den Standardwerken zur Wortbildung über die Produktivität der *-ung*-Suffigierung zu lesen ist. Anschließend folgt die Einschätzung von Motsch, die, wie von Bauer gefordert, getrennt eingeht auf die vier verschiedenen Wortbildungsmuster mit ihren jeweiligen semantischen Mustern, wie sie im vorangehenden Abschnitt erläutert wurden.

## **2.1.7. Bildung, Beschränkungen und Produktivität der *-ung*-Suffigierungen**

### **2.1.7.1. Aussagen über Bildung und mögliche Basen in den traditionellen Grammatiken**

Nach einer Durchsicht der großen Werke zur Wortbildung – Henzen (1947), Erben (1975), Wellmann (1975) und Fleischer/Barz (1992) – lässt sich zunächst einmal festhalten, dass alle Autoren die *-ung*-Suffigierung als eines der produktivsten Wortbildungsmuster der deutschen Sprache bezeichnen.<sup>52</sup> Es lassen sich damit Nomen mit dem Genus Femininum bilden. Als Grundregel für die Bildung der Nominalisierungen geben alle an, dass das Suffix *-ung* an einen Verbstamm angehängt wird – welche Verben allerdings als Basis für die *-ung*-Suffigierung in Frage kommen, darüber herrscht Uneinigkeit und es scheint, als ob kein Autor diese Frage abschließend beantworten kann. Einigkeit besteht nur bei der Feststellung, dass *-ung*-Suffigierungen eher von Präfix- und Partikelverben als von Simplizia möglich seien. Bei Henzen (1947) lesen wir, dass unter den intransitiven Verben eine Ableitung bei den Inchoativa möglich sei, und dass reflexive Verben mögliche Basen seien. Fleischer/Barz schließen modale Hilfsverben und Komposita aus zwei Verbstämmen (z.B. *sitzenbleiben*) als Basen aus. Wellmann (1975) merkt an, dass sich nahezu alle morphologischen Gruppen von Verben mit *-ung* kombinieren lassen, außer Verbkomposita mit dem Präfixoid *umher-*, die eine abgeschlossene Handlung bezeichnen. Außerdem, so Wellmann, werden mit *-zer* präfigierte Intransitiva (*zerschmelzen*, *zerfließen*) nur sehr selten für die *-ung*-Nominalisierung genutzt. Neben diesen Einzelbeobachtungen erkennt Wellmann einen größeren Zusammenhang zwischen Verbsemantik und *-ung*-Suffigierung, der immerhin für eine größere Gruppe von Verben eine Aussage macht: „Insgesamt ist die *-ung*-Bildung also vor allem bei Verben eingeschränkt, die einen andauernden Zustand, einen zeitlich nicht begrenzten Ablauf, einen wiederholten Vorgang oder eine mehrfache Tätigkeit bezeichnen, also das Merkmal der

---

<sup>52</sup> Allerdings wird Produktivität bei Wellmann und Fleischer/Barz als „Häufigkeit der Regelanwendung verstanden und an der Zahl der belegten Lexeme gemessen“. (Mikolajczyk 2002: 12)

Unabgeschlossenheit enthalten“.<sup>53</sup>

Zusammenfassend lässt sich festhalten, dass die einzelnen Autoren mehr oder weniger umfangreiche Informationen über die Bildung der *-ung*-Nominalisierungen zusammentragen, ohne zu einem vollständigen Regelwerk zu gelangen.

Zu dem Schluss, dass die Bildungsrestriktionen, die für die *-ung*-Suffigierung gelten, bisher nicht ausreichend beschrieben sind, kommt auch Monika Mikolajczyk in ihrer Magisterarbeit. Sie fordert einen neuen Ansatz, um „zuverlässige Regeln zur Bildungsweise (...) herausarbeiten zu können“, da traditionelle Ansätze bisher versagt hätten und „leider in langen Aufzählungen der möglichen Basisverben mit unzähligen Ausnahmefällen enden“.<sup>54</sup> Meines Wissens ist dieses Desiderat bislang noch unerfüllt geblieben. Es muss aber letztendlich Ziel einer Theorie der Wortbildung sein, nicht nur die Einheiten des Lexikons und weitere in Texten belegte Wörter zu beschreiben und zu interpretieren, sondern auch Aussagen darüber zu machen, was „mögliche“ oder „potentielle Wörter“<sup>55</sup> einer Sprache sind. Dazu braucht man jedoch genaue Regeln über die Bildungsweise potentieller Wörter.

Motsch (1999) weist auf eine Besonderheit bei der Pluralbildung von *nomina actionis* mit dem Suffix *-ung* hin. Generell gilt, dass nur von zählbaren Nomen ein Plural gebildet werden kann und „Zählbarkeit setzt voraus, dass das Nomen als Bezeichnung einer Klasse, einer Gattung verwendet werden kann, der Exemplare oder Instanzen zugeordnet sind“.<sup>56</sup> *Nomina actionis* können ein durch das Basisverb ausgedrücktes Geschehen als Tatsache, als generelles Geschehen und als Verlauf darstellen. Für die ersten beiden Bedeutungen sind keine Gattungsbezeichnungen mit *-ung* möglich (*die Erforschung der Sonne*, *\*die Erforschungen der Sonne*; *die Beseitigung des Ärgernisses*, *\*die Beseitigungen des Ärgernisses*), für den Verlauf dagegen schon. Hier sind Geschehen denkbar, die in mehrere Einzelgeschehen unterteilt werden können (*Der Flug nach Neuseeland wird durch zwei Zwischenlandungen unterbrochen*<sup>57</sup>).

Es sei noch auf eine Regelhaftigkeit hingewiesen, die für die Bildung von Komposita gilt: wenn das Erstglied des Kompositums ein *-ung*-Derivat ist, steht zwischen Erst- und Zweitglied regulär eine *-s*-Fuge wie in *Wohnungsinhaber*.<sup>58</sup> Abweichend von dieser Regel gibt es wenige Ausnahmen wie etwa die Zusammenbildung „Stellungnahme“.

---

<sup>53</sup> Wellmann 1975: 214

<sup>54</sup> Mikolajczyk 2002: 16f

<sup>55</sup> „A possible, or potential, word can be defined as a word, existing or non-existing, whose morphological and phonological structure is in accordance with the rules of the language. It is obvious that before one can assign the status of 'possible word' to a given form, these rules need to be stated as clearly as possible“. (Plag 1999: 7)

<sup>56</sup> Motsch 1999: 327

<sup>57</sup> Die Beispiele sind Motsch (1999: 327f) entnommen.

<sup>58</sup> Vgl. Erben 1975: 59f

### 2.1.7.2. Blockierung

Die Bildungsrestriktionen für *-ung*-Suffigierungen umfassen nicht nur Beschränkungen in Bezug auf mögliche Basen für die Derivation. Auch für die Ableitungsprodukte gibt es Beschränkungen, die als „Blockierung“ („blocking“) bezeichnet werden können. Darunter versteht Plag (1999: 51) vor allem „token-blocking“, definiert als „the blocking of potential forms by already existing synonymous words (...)“.

Neben dem token-blocking gibt es auch den Begriff des „type-blocking“ zur Bezeichnung der „interaction of more or less regular rival morphological processes (...)“<sup>59</sup>, die synonyme komplexe Wörter bilden. Es ist jedoch nicht klar, welche Prozesse damit im Deutschen konkret gemeint sein können. Wellmann (1975) beobachtet zwar, dass bei den Verben auf *-ieren* (*formulieren, isolieren, provozieren, usw.*) das Suffix *-ung* (bzw. die allomorphe Variante *-ierung*) mit den Suffixen *-(at)ion*, *-(at)ur* und *-ment* konkurriert und *-ierung* dabei verhältnismäßig selten zur Bildung von Derivaten genutzt wird. Allerdings hat Wellmann keine Regeln dafür formuliert, welche Verben auf *-ieren* mit welchem Suffix abgeleitet werden müssen. Zudem scheint es eher so zu sein, dass die Konkurrenz dieser Suffixe nicht in einer gegenseitigen Blockierung resultiert. Der Duden gibt für die konkurrierenden Suffixe *-(at)ion* und *-(ier)ung* an, dass sie „oft ohne Bedeutungsunterschied nebeneinander (stehen)“ und sich höchstes Bedeutungsnuancen zeigen: *-(at)ion* bezeichnet demnach „vorrangig das Ergebnis einer Handlung oder Tätigkeit, seltener das Geschehen selbst“ und *-ierung* bildet Substantive, die „eine Handlung oder Tätigkeit, seltener deren Ergebnis“ bezeichnen.<sup>60</sup>

Type-blocking ist also ein weniger klares Konzept und sollte nach Plag (1999) ganz aufgegeben werden<sup>61</sup>, der Begriff „Blockierung“ wird meistens für das häufiger vorkommende token-blocking verwendet und von Rainer (1993) wie folgt definiert: „Gewöhnlich versteht man darunter die Erscheinung, daß die Bildung eines morphologisch komplexen Wortes durch ein usuelles Wort behindert oder ganz blockiert werden kann. Ist dieses usuelle Wort ein Synonym, so spricht man von *synonymischer Blockierung*, ist es ein Homonym, von *homonymischer Blockierung*“.<sup>62</sup> Rainers Definition ist allgemeiner und bezieht sich natürlich nicht nur auf die *-ung*-Suffigierung, für die mir kein Beispiel für eine homonymische Blockierung bekannt ist. Für die synonymische Blockierung sei ein Beispiel aus den Korpusdaten dieser Arbeit angeführt: ein DaF-Lerner schreibt in seinem Essay das Wort *Schwangerschaftsabbruchung* (Essays L2, 1.0, fk017\_2006\_07), welches zwar

---

<sup>59</sup> Plag 1999: 51

<sup>60</sup> Duden Bd. 5: 100

<sup>61</sup> Vgl. die Argumentation gegen type-blocking, S. 51-54

<sup>62</sup> Rainer 1993: 114

eigentlich nach dem Wortbildungsmuster für *-ung*-Suffigierungen regelkonform gebildet ist, jedoch durch das bereits als Lexem im Lexikon des Muttersprachlers gespeicherte Wort *Schwangerschaftsabbruch* blockiert ist. Mehrere solcher Fälle werden später noch im Analyseteil dieser Arbeit aufgeführt sein. Diese Art von Lernerfehler ist in gewissem Maße erwartbar, da bei fortgeschrittenen Lernern angenommen werden kann, dass sie die regelhafte Bildung von *-ung*-Derivaten verinnerlicht haben, aber noch nicht denselben Wortschatz wie Muttersprachler erworben haben, sodass die Blockierung in vielen Fällen nicht wirksam werden kann.

Das Phänomen, dass die Blockierung versagt, ist allerdings nach Rainer nicht nur bei Lernern des Deutschen zu beobachten: „In Versprechern kommen solche normalerweise blockierten Bildungen manchmal an die Oberfläche (...). Besonders häufig sind solche Bildungen in der Kindersprache (...), da das Kind das blockierende Wort oft noch gar nicht kennt, noch nicht fest genug gespeichert hat oder aber nicht als Synonym des zu blockierenden Wortes erkennt“.<sup>63</sup>

Ergänzend sei hier eine Feststellung von Plag (1999) hinzugefügt, nach welcher token-blocking generell kein absolutes Phänomen ist und gelegentlich sogar synonyme Dubletten auftreten können. Plags Erklärung dafür sind Lexem-Speicherung und Frequenzphänomene: je seltener ein usuelles Wort von einem Sprecher aktiviert wird, desto geringer ist die Wahrscheinlichkeit, dass es ein neu gebildetes Wort blockiert.<sup>64</sup>

### 2.1.7.3. Pragmatische Bildungsrestriktionen

Die bisher genannten Bildungsrestriktionen für *-ung*-Suffigierungen bezogen sich entweder auf das Sprachsystem (Auswahl der möglichen Basen) oder die Sprachnorm (Blockierung durch Wörter, die in der Sprachgemeinschaft den Status von usuellen Wörtern haben). Daneben gibt es das, was von verschiedenen Autoren „pragmatic factors“ (Plag 1999), „a need to form words“ (Bauer 2001), „Bedarf an Neubildungen“ (Rainer 1993) oder „pragmatische Aspekte“ (Motsch 1999) genannt wird. Die Idee ist dabei immer, dass der Grad, zu welchem das Wortbildungspotential eines morphologischen Prozesses ausgeschöpft wird, davon abhängt, ob die Sprecher zur Befriedigung ihrer kommunikativen Bedürfnisse die Notwendigkeit sehen, mittels Wortbildungsmustern neue Wörter zu erschaffen. Dabei kann die Bildung neuer Wörter ganz verschiedene Funktionen erfüllen, wie bereits in Abschnitt 2.1.1. aufgeführt ist: z.B. Benennung neuer Gegenstände, Konzepte oder Ideen, Umkategorisierung zur Einpassung lexikalischer Einheiten in verschiedene Kontexte und Konstruktionen oder stilistische Variation, um einen persönlichen Schreibstil zu prägen oder die

---

<sup>63</sup> Ebd. S. 115

<sup>64</sup> Vgl. Plag 1999: 52

Sprech- bzw. Schreibweise einem bestimmten Genre anzupassen, und vieles mehr. Wie produktiv ein Wortbildungsmuster gebraucht wird, hängt nach Motsch „von der Funktion der nach dem Muster gebildeten Wörter ab, d.h. die Funktion hat Einfluss auf die Häufigkeit der Verwendung eines Musters für Neubildungen“.<sup>65</sup> So ist mit Motsch weiter davon auszugehen, dass Muster für die Bildung neuer Lexikoneinheiten, die konkrete oder abstrakte Dinge benennen, nicht so häufig gebraucht werden, Muster, die Wörter zur stilistischen Variation schaffen, dagegen viel häufiger. Am häufigsten, da für die Textgestaltung nahezu unerlässlich, werden neue Wörter nach den Mustern für die Umkategorisierung gebildet. Im Falle der *-ung*-Suffigierung ist also mit den meisten Neuschöpfungen in der Kategorie 'nomina actionis' zu rechnen.

Das Bedürfnis, ein neues Wort zu bilden, hat mitunter eine höchst subjektive Komponente, wie Bauer (2001) anmerkt: „For coinage to occur, there has to be a need (...), a real or perceived gap in the speaker's lexicon“. Bauer gebraucht den Ausdruck 'perceived gap', denn „the need may arise from temporary memory failure or from ignorance of an established word, as well as from the absence of any appropriate lexeme in the norm of the language being used“.<sup>66</sup> Was Bauer hier für den erwachsenen Muttersprachler formuliert, trifft natürlich in noch höherem Maße für den Lerner einer Sprache zu.

#### **2.1.7.4. Die Produktivität der Wortbildungsmuster nach Motsch (1999)**

In dieser Arbeit wurde Produktivität definiert als das Potential zur Bildung neuer Wörter. Um nicht zu weit ausholen zu müssen und den Argumentationsfluss nicht zu stören, habe ich bisher auf eine Definition des Begriffs „neues Wort“ und damit auf die Unterscheidung zwischen „Neologismus“ und „Textwort“ verzichtet, will dies aber nun nachholen. Neologismen sind neue Lexeme, die bereits in das Lexikon der Sprachgemeinschaft aufgenommen wurden – zunächst einmal unabhängig davon, ob sie auch schon in einem Wörterbuch eingetragen sind. Textwörter sind mit Motsch (1999) Wörter, die „ad hoc für die Zwecke eines Textes gebildet sind“<sup>67</sup>, unter Text wird hier auch die gesprochene Rede verstanden. Ohne ins Detail zu gehen, deutet Motsch an, dass die Unterscheidung zwischen Textwort und Neologismus, also die Frage, ob eine Neubildung nur für einige Sprecher neu wirkt oder ob sie bereits zum Lexikon einer Sprachgemeinschaft gehören könnte, im konkreten Fall eine nicht triviale soziolinguistische oder lexikographische Frage darstellt. So theoretisch sinnvoll die Unterscheidung sein mag, so schwierig ist sie in der Praxis zu treffen, wenn es darum geht, eine konkrete Neubildung als Textwort oder als Neologismus –

---

<sup>65</sup> Motsch 1999: 22

<sup>66</sup> Bauer 2001: 41

<sup>67</sup> Motsch 1999: 19

natürlich nur, sofern dieser noch nicht in einem Wörterbuch eingetragen ist – einzustufen. Für Untersuchungen zur Produktivität eines Wortbildungsmusters wäre die Frage aber durchaus relevant. Schließlich gehört zur Definition eines produktiven Musters, dass es dazu dient, für die Produktion und Rezeption von Textwörtern genutzt zu werden. Für Neologismen wird das Muster streng genommen nicht mehr benötigt, da es sich hier bereits um Einheiten des Lexikons handelt, die als solche abgerufen werden können und nicht mithilfe eines Musters gebildet oder analysiert werden müssen.

Für die folgenden Angaben zur Produktivität der Wortbildungsmuster der *-ung*-Suffigierung, die Motsch (1999) entnommen sind, gilt es also sich bewusst zu machen, dass die strenge Unterscheidung zwischen Textwort und Neologismus nicht vorgenommen wurde. Motsch selbst schreibt, dass die Angaben „mangels genauerer Kriterien, subjektive Einschätzungen“<sup>68</sup> wiedergeben<sup>69</sup>. Des Weiteren gibt es eine terminologische Besonderheit zu beachten. Motsch spricht von der „Aktivität“ von Wortbildungsmustern und meint damit das Potential, für die Bildung neuer Textwörter genutzt zu werden. „Ein Wortbildungsmuster wird als **inaktiv** eingeschätzt, wenn es keine oder nur auf einzelne Autoren beschränkte, willkürlich wirkende Textwörter gibt, als **schwach aktiv**, wenn wiederholt Textwörter und Neologismen auftreten und es gilt als **stark aktiv**, wenn relativ häufig Textwörter und Neologismen zu diesem Muster zu finden sind“ (Hervorhebungen wie im Originaltext).<sup>70</sup> Nach meinem Verständnis meint „Aktivität“ bei Motsch genau das, was in dieser Arbeit mit dem Begriff „Produktivität“ benannt wird und statt „inaktiv“, „schwach aktiv“ und „stark aktiv“ werde ich im folgenden die Formulierungen „unproduktiv“, „wenig produktiv“ und „sehr produktiv“ verwenden, sofern es sich nicht um direkte Zitate aus Motsch handelt.

Für die einzelnen Muster der *-ung*-Suffigierung, wie sie in Abschnitt 2.1.5.1. aufgelistet sind, macht Motsch also die folgenden Angaben zur Produktivität: als sehr produktiv wird der Bildungstyp eins mit dem semantischen Muster der Umkategorisierung eingestuft. Der Bildungstyp zwei „Resultierender Zustand“ wird als produktiv eingestuft mit Verben, die psychische oder physische Zustände bezeichnen. Die Bildung von *nomina acti* wird ebenfalls als sehr produktiv eingeschätzt, wogegen die Bildung von *nomina loci* als unproduktiv bewertet wird (vgl. Abb. 2).

<b>Wortbildungsmuster</b>	<b>Produktivität (nach Motsch 1999)</b>
1) <i>Nomina actionis</i>	sehr produktiv
2) Resultierender Zustand	Produktiv
3) <i>Nomina acti</i>	Produktiv

<sup>68</sup> Ebd. S. 19

<sup>69</sup> Zur Datengrundlage von Motschs Untersuchung vgl. ebd. S. 26f

<sup>70</sup> Ebd. S. 19

4) Nomina loci	unproduktiv
----------------	-------------

**Abb. 2:** Wortbildungsmuster der *-ung*-Suffigierung mit Produktivitätsangaben nach Motsch (1999)

Vor diesem Hintergrund werden im Hauptteil dieser Arbeit die Ergebnisse der Korpusanalyse zu betrachten sein.<sup>71</sup> Zumindest das Vergleichskorpus mit den Texten von Muttersprachlern sollte ähnliche Bewertungen liefern. Das Hauptaugenmerk wird darauf liegen, wie produktiv die fortgeschrittenen Lerner die jeweiligen Wortbildungsmuster in ihrer lernersprachlichen Produktion anwenden. Das setzt voraus, dass Produktivität quantitativ messbar ist. Der folgende Abschnitt beschäftigt sich mit der Messbarkeit von Produktivität und erklärt das Produktivitätsmaß, welches für die Korpusanalyse in dieser Arbeit verwendet wird.

### 2.1.8. Messbarkeit von Produktivität

Die Frage nach der Messbarkeit von Produktivität setzt zunächst eine klare Vorstellung von Produktivität voraus. Wie in Abschnitt 2.1.6. gezeigt wurde, ist Produktivität alles andere als ein klar definierter Begriff in der Linguistik. Trotzdem wurden in dieser Arbeit einige Grundannahmen getroffen bezüglich dessen, was unter Produktivität zu verstehen sei, allen voran die Vorstellung, dass es sich bei der Produktivität eines morphologischen Prozesses (bzw. Wortbildungsmusters) um das Potential dieses Prozesses zur Bildung neuer Wörter handelt. Baayen (2009: 902) nennt dies die „potential productivity“<sup>72</sup> einer Wortbildungsregel und erklärt, wie man das Produktivitätsmaß  $P$  dieser potential productivity berechnen kann. Dazu benötigt man die Zahl der Hapax Legomena  $V(1,C,N)$ , also die Zahl der Wörter einer morphologischen Kategorie  $C$ , die nur einmal vorkommen bei einer Größe von  $N$  Tokens dieser Kategorie in dem Korpus. Dahinter steht die Idee, dass usuelle Wörter von vielen Sprechern verwendet werden und in einem Korpus häufiger vorkommen, während Neubildungen eher unter den Wörtern mit einer sehr geringen Vorkommenshäufigkeit zu finden sind. Wenn in einem Korpus also viele Hapax Legomena einer bestimmten morphologischen Kategorie vorkommen, ist nach Scherer (2006: 37) „die Wahrscheinlichkeit hoch, dass das untersuchte sprachliche Muster von den Sprechern bzw. Schreibern produktiv eingesetzt wird und dass nach seinem Vorbild neue Bildungen vorgenommen werden. Allgemein kann der Anteil der Hapax Legomena an der Zahl der Tokens dazu verwendet werden, die Produktivität eines sprachlichen Musters zu bestimmen. Je höher der Anteil der Einmalbelege, desto höher ist die Wahrscheinlichkeit, dass das Muster Neubildungen hervorbringt“. Die Zahl der Hapax Legomena

<sup>71</sup> Vgl. Abschnitt 3.1.3.

<sup>72</sup> Für andere Konzepte von Produktivität und deren Messbarkeit vgl. Baayen (2009: 901f) und Plag (1999: Kap. 2.3)

$V(1,C,N)$  in Kategorie  $C$  muss also geteilt werden durch die Zahl  $N(C)$  aller Tokens dieser Kategorie im Korpus und das Ergebnis ist die potential productivity  $P$  dieser Kategorie. Die Formel zur Berechnung des Produktivitätsmaßes lautet nach Baayen (2009: 902):

$$P = V(1,C,N) / N(C)^{73}$$

Baayen betont, dass dies keine Gleichsetzung von Hapax Legomena und Neologismen bedeutet.<sup>74</sup> Idealerweise wären es natürlich die Neologismen, die man zählen sollte, doch mit dem Begriff des Neologismus sind große theoretische Schwierigkeiten verbunden<sup>75</sup> und korpusbasierte Zählungen von Hapax Legomena bedeuten einen praktikablen, indirekten Weg. Außerdem kann die Berechnung der potential productivity auf Basis von Hapax Legomena auch in sehr kleinen Korpora wie dem Falko-Korpus durchgeführt werden, in denen unter Umständen gar keine echten Textwörter (also Neubildungen, die nicht im Wörterbuch verzeichnet sind) vorkommen. In kleinen Korpora ist nach Baayen die Zahl der Neologismen unter den Hapax Legomena sehr gering, sie wird größer, sobald die Korpusgröße zunimmt.

## 2.2. Lernersprache

In dieser Arbeit werden schriftliche Produktionen von fortgeschrittenen Lernern des Deutschen als Fremdsprache untersucht, um Erkenntnisse über die „Lernersprache“ in diesem Erwerbsstadium in Bezug auf den produktiven Umgang mit der *-ung*-Suffigierung zu gewinnen. Lernersprache meint hier ein theoretisches Konstrukt, das im Wesentlichen auf Selinkers Konzept der „Interlanguage“ (vgl. Selinker 1972) zurückgeht, verstanden als ein eigenständiges System, das nicht mit der Muttersprache (L1) oder der zu erlernenden Fremdsprache (L2) eines Lerners identisch ist, sondern seine eigenen Regeln besitzt, die den Äußerungen von Lernern zugrunde liegen.

Einem kurzen historischen Abriss, der die Entwicklung des Lernersprachen-Konzeptes aus den unbefriedigenden Ergebnissen der „kontrastiven Analyse“ sowie der „Fehleranalyse“ heraus nachzeichnet, folgt eine detaillierte Beschreibung der theoretischen Annahmen, die in dieser Arbeit gemacht werden, damit Lernersprache zu einem geeigneten Untersuchungsgegenstand für eine Korpusanalyse wird.

---

<sup>73</sup> „This ratio, known as the category-conditioned degree of productivity (Baayen 1993), estimates the growth rate of vocabulary of the morphological category itself“. (Baayen 2009: 902)

<sup>74</sup> Vgl. Baayen 2009: 905f

<sup>75</sup> Vgl. ebd. 2009: Abschnitt 4.1.2.

### 2.2.1. Der Weg zur Entwicklung des Konzeptes „Lernersprache“

Lernersprache ist heute das zentrale Konzept in der Zweitsprachenerwerbsforschung. An dessen Anfang steht jedoch die kontrastive Analyse, die vor allem auf Fries (1945) und Lado (1957) zurückgeht und in den 60er und frühen 70er Jahren des vergangenen Jahrhunderts große Popularität genoss.<sup>76</sup> Die kontrastive Analyse entstand auf der Grundlage von behavioristischen Modellen der Sprache und des Spracherwerbs. Dazu gehört die Annahme, dass der „Gebrauch sprachlicher Strukturen (...) auf (halb)automatischen Stimulus-/Responseverbindungen („habits“) (beruht), die der bewußten Kontrolle nicht oder nur schwer zugänglich sind“ und dass die „muttersprachlich ausgebildeten habits (...) auf die fremde Sprache übertragen, d.h. transferiert (werden)“.<sup>77</sup> Die Grundidee der kontrastiven Analyse war es also, dass Sprachtransfer aus der Muttersprache die treibende Kraft hinter dem Erwerb und Gebrauch einer Fremdsprache sei. Damit wurde angenommen, dass der Kontrast zwischen Mutter- und Fremdsprache dafür verantwortlich sei, welche Schwierigkeiten beim Erwerb der Fremdsprache auftreten: strukturelle Ähnlichkeiten zwischen zwei Sprachen sollten zu Lernerleichterungen führen, Unähnlichkeiten dagegen zu Problemen und Fehlerquellen. Dies ging soweit, dass man glaubte, durch den linguistischen Vergleich zweier Sprachen vorhersagen zu können, welche Fehler Lerner beim Erwerb der Fremdsprache machen würden, und basierend auf diesen Annahmen Konzepte für den Unterricht entwickeln zu können. Diesem ehrgeizigen Anspruch konnte die kontrastive Analyse nicht gerecht werden, wie empirische Untersuchungen, nämlich die Analyse von Fehlern in den Produktionen von Lernern, zeigten: zum einen führten nicht alle von der kontrastiven Analyse vorhergesagten Lernschwierigkeiten zu Fehlern, zum anderen waren bei weitem nicht alle Fehler auf einen Einfluss der Muttersprache, also auf irgendeine Form von Transfer, zurückzuführen. „Zahlreiche, scheinbar kontrastive Fehler ließen sich auch für Sprecher strukturell verschiedener Muttersprachen nachweisen oder waren sogar aus dem Erstsprachenerwerb bekannt.“<sup>78</sup>

Trotz dieser offensichtlichen Misserfolge der kontrastiven Analyse hatte die Beschäftigung mit echten Lernerdaten den Weg geebnet in die Fehleranalyse, die eng mit dem Namen Pit Corder und seinen Veröffentlichungen in den 60er und 70er Jahren verbunden ist. Corder erkannte, dass die Fehler der Lerner nicht nur zur Überprüfung kontrastiver Vorhersagen dienen, sondern noch viel mehr über das Sprachwissen der Lerner verraten, wie Rod Ellis (1995) würdigend festhält:

„In an early, seminal article, Corder (1967) noted that errors could be significant in three ways: (1) they provided the teacher with information about how much the learner had

---

<sup>76</sup> Vgl. Kohn 1990: Abschnitt 1.1.

<sup>77</sup> Ebd. S. 3

<sup>78</sup> Kohn 1990: 6

lernt, (2) they provided the researcher with evidence of how language was learnt, (3) they served as devices by which the learner discovered the rules of the target language.“<sup>79</sup>

Mit Corders Erkenntnissen geht gleichsam eine „Aufwertung“ der Fehler einher. Die Nutzbarmachung der Fehler zur Gewinnung neuer Erkenntnisse über den Spracherwerbsprozess tritt in den Vordergrund, denn die Fehler „weisen über sich hinaus auf die Prozesse und Strategien hin, die den Erwerb ermöglichen und den Erwerbsverlauf bestimmen. Und diese Prozesse und Strategien wurden das eigentliche Ziel der Analyse“.<sup>80</sup> Damit verbunden ist eine Art Paradigmen-Wechsel in der Zweitsprachenerwerbsforschung: weg vom traditionellen Schwerpunkt des Lehrens hin zum Lernen, d.h. der Lerner rückt in den Mittelpunkt, es wird angenommen, dass sowohl muttersprachlicher als auch Fremdsprachenerwerb auf einer „aktiven und kreativen Erwerbsleistung des Lerners beruhen“.<sup>81</sup> Dem zugrunde liegt eine neue Erwerbstheorie, nämlich die mentalistische, welche vor allem auf Chomskys (1957, 1965) Arbeiten zur generativen Grammatik zurückgeht und annimmt, dass ein Kind im Verlauf des Spracherwerbs mithilfe einer angeborenen Disposition für Sprache aus dem sprachlichen Input, den es bekommt, seine mentale Grammatik erstellt. Diese Erwerbstheorie wurde zunächst als nur für den Erstspracherwerb relevant angesehen und erst nach und nach setzte sich die Ansicht durch, dass auch beim Zweitsprachenerwerb von einer ähnlichen – wenngleich nicht identischen – kreativen Leistung des Lerners auszugehen ist.

Die drei bei Corder genannten Punkte finden sich auch im Konzept der Lernersprache wieder, das allerdings noch darüber hinaus geht. Denn auch eine so verstandene und genutzte Fehleranalyse weist noch deutliche Mängel auf. Erstens verstellt das Problem der Vermeidung den Blick auf das Wissen des Lerners: eine reine Fehleranalyse deckt nicht die sprachlichen Strukturen auf, die dem Lerner schwer fallen und anfällig für Fehler sind, wenn er sie in seinen Produktionen umgeht. Zweitens muss man, wenn man wissen möchte, was ein Lerner über die Zielsprache weiß, nicht nur seine fehlerhaften Produktionen berücksichtigen, sondern vor allem auch anschauen, was er richtig macht und drittens ist die Definition dessen, was ein Fehler ist, keineswegs trivial und in der Literatur uneinheitlich. Ich komme auf diese Problematik der Fehlerdefinition und Fehlerklassifikation im Abschnitt 2.4. noch mal zurück.

Kontrastive Analyse und Fehleranalyse haben sich also als unzureichende Methoden erwiesen, um zu einem tieferen und umfassenden Verständnis dessen zu gelangen, wie ein Lerner eine Fremdsprache lernt, was ein Lerner über die zu erlernende Fremdsprache an einem bestimmten Punkt seiner Entwicklung weiß, und wie er dieses Wissen in Rezeption und Produktion einsetzt.

---

<sup>79</sup> Ellis 1995: 48

<sup>80</sup> Kohn 1990: 7

<sup>81</sup> Ebd. S. 9f

Gleichwohl haben sie wichtige Vorarbeit geleistet für die „Entdeckung“ der Lernersprache, die laut Selinker spätestens Mitte der 60er Jahre begann, als verschiedene Forscher (Selinker nennt als Beispiel Corder, Nemser, James und sich selbst) unabhängig voneinander die Vorstellung entwickelten, dass beim Spracherwerb nicht nur zwei linguistische Systeme – Muttersprache und Zielsprache – involviert sind, sondern drei, wobei für das dritte lange Zeit keine Bezeichnung zur Verfügung stand.<sup>82</sup> Der Terminus „Interlanguage“, der sich für die Bezeichnung dieses dritten Systems durchsetzen sollte, erschien zum ersten Mal gedruckt 1969 in einem Artikel von Selinker in der Zeitschrift „General Linguistics“:

„An 'interlanguage' may be linguistically described using as data the observable output resulting from a speaker's attempt to produce a foreign norm, i.e. both his errors and non-errors. It is assumed that such behavior is highly structured. In comprehensive language transfer work, it seems to me that recognition of the existence of an interlanguage cannot be avoided and that it must be dealt with as system, not as an isolated collection of errors“.<sup>83</sup>

Hier ist bereits der Kern dessen festgehalten, was später das theoretische Konstrukt „Interlanguage“ ausmachen sollte. Ich möchte im Rahmen dieser Arbeit nun auf eine ausführliche Beschreibung der Entwicklung des Lernersprachenkonzeptes von diesen Anfängen bis zum heutigen Stand, der gekennzeichnet ist durch eine Vielzahl verschiedener und teils auch konkurrierender Erklärungsmodelle für die Lernersprache, verzichten<sup>84</sup> und direkt auf jene theoretischen Implikationen zu sprechen kommen, die heute mit der Lernersprache verbunden sind, und die die Lernersprache zu einem Untersuchungsgegenstand für diese Arbeit machen.

### **2.2.2. Lernersprache – theoretische Annahmen**

Ich möchte diesem Abschnitt zwei Zitate voranstellen, die die theoretischen Annahmen, die ich mit dem Konstrukt Lernersprache verbinde, recht gut zusammenfassen. Klaus Vogel (1990) schreibt in seiner Abhandlung über die Lernersprache:

„Mit Lernersprache bezeichnen wir das Sprachgebilde, das sich in einem Fremdsprachenlerner infolge der Konfrontation mit zielsprachlichen Daten herausbildet, ohne dabei jedoch völlig mit der jeweiligen Zielsprache identisch zu sein. Für die Entstehung der Lernersprache sind die Muttersprache, gegebenenfalls weitere vorgängige Fremdsprachen und die Zielsprache konstitutiv; für den Grad ihrer Ausprägung, ihren Entwicklungsstand, ihre Idiosynkrasien sind insbesondere individuelle, soziale lernsituationsabhängige und dabei gegebenenfalls auch

<sup>82</sup> Vgl. Selinker 1992: 222

<sup>83</sup> Zitiert nach Selinker 1992: 231

<sup>84</sup> Einen detaillierten Überblick über die Herausbildung des Lernersprachenkonzeptes und verschiedene Erklärungsmodelle der Lernersprache leistet das erste Kapitel von Kohn (1990), weitere Ergänzungen bietet das achte Kapitel von Selinker (1992).

methodisch-didaktische Variablen ausschlaggebend“.<sup>85</sup>

Ergänzend dazu lesen wir bei Kurt Kohn (1990):

„Dem Lerner wird also eine Sprache unterstellt, die, wie jede andere Sprache auch, in sich regelhaft und systematisch ist und die den universellen Bedingungen folgt, die für eine natürliche menschliche Sprache gelten (...). Sie unterscheidet sich allerdings von anderen natürlichen Sprachen – weniger der Art als dem Grade nach – in der Stabilität, Variabilität und Robustheit ihre Systems“.<sup>86</sup>

Was in diesen beiden Zitaten sehr prägnant über die Lernaltersprache gesagt wird, möchte ich im Folgenden näher ausführen.

### **2.2.2.1. Regelhaftigkeit vs. Variabilität der Lernaltersprache**

Zentral für das Konstrukt Lernaltersprache ist im Prinzip ein Gegensatz: der Lernaltersprache wird unterstellt, regeldeterminiert, systematisch und relativ eigenständig zu sein, während sie gleichzeitig als variabel, instabil und dynamisch konzipiert wird. Die wichtige Erkenntnis, dass Lernaltersprache ihren eigenen Regeln folgt und nicht willkürlich ist, ist unter anderem ein Ergebnis der Arbeiten im Bereich der Fehleranalyse. Fehler können nämlich oftmals als systematisch und regelhaft interpretiert werden – wobei es sich hier nicht um Regeln der Mutter- bzw. der Zielsprache des Lerners handeln muss (das Regelsystem der Lernaltersprache kann also relativ autonom sein in Bezug auf Mutter- und Zielsprache), sondern um Regeln, die der Lerner im Verlauf des Spracherwerbs in dem ihm zur Verfügung stehenden Input bewusst oder unbewusst extrahiert hat. Diese Regeln haben oft den Charakter von Hypothesen, die der Lerner bildet und die er in seinen Produktionen testet und überprüft, bei negativem Feedback etwa von Lehrern idealerweise auch revidiert. Trotz dieser erwerbsbedingten Instabilität der Regeln macht „erst die grammatische Regelhaftigkeit der Lernaltersprache die Untersuchung des Erwerbs zu einem sinnvollen Unterfangen (...)“.<sup>87</sup> Ein typisches Beispiel für eine lernaltersprachliche Regel, die zu systematischen Fehlern führen kann, ist laut Vogel (1990) die Übergeneralisierung zielsprachlicher Strukturen. Dies passiert, wenn ein Lerner eine zielsprachliche Struktur als in der Zielsprache regelhaft interpretiert, ohne dabei die Ausnahmen mitzulernen und in der Folge jene zielsprachliche Struktur auf Kontexte anwendet, in denen sie nicht normgerecht ist. Ein Beispiel dafür aus dieser Arbeit ist das Phänomen der lexikalischen Blockierung, auf welches in Abschnitt 2.1.7.2. bereits näher eingegangen wurde.

---

<sup>85</sup> Vogel 1990: 13

<sup>86</sup> Kohn 1990: 11

<sup>87</sup> Kohn 1990: 30

Vogel nennt weitere Argumente für die relative Autonomie und Systemhaftigkeit der Lerner-  
sprache<sup>88</sup>: empirische Untersuchungen haben nahegelegt, dass es für den Fremdspracherwerb  
Erwerbsuniversalien zu geben scheint, die – ähnlich dem Erstspracherwerb – eine gewisse  
Erwerbssequenz von sprachlichen Merkmalen festlegen, unabhängig von Muttersprache, Alter oder  
Sprachunterricht. Des Weiteren kann frühzeitig fossilisierte Lerner-  
sprache ähnlich wie Pidgin-  
Sprachen trotz ihrer Begrenztheit als Kommunikationsmittel verwendet werden. „Beständigkeit  
mancher Strukturen ebenso wie die Möglichkeit, fossilisierte Lerner-  
sprache als Kommunikations-  
mittel einzusetzen, lassen auf eine gewisse Autonomie lernersprachlicher Systeme schließen“.<sup>89</sup>  
Hinzu kommt, dass es möglich ist, trotz der Instabilität individueller Lerneräußerungen durch  
Lernergruppenbildung (z.B. nach Eigenschaften wie Muttersprache, Sprachniveau, Alter, etc.) einen  
idealisierten 'gruppentypischen Normal-Lerner' zu extrahieren, und dessen mentale Grammatik als  
relativ stabile Größe zu rekonstruieren. Eine solche Lernergruppenbildung liegt auch den Falko-  
Daten zugrunde, auf die in dieser Arbeit zurückgegriffen wird; nur so ist es möglich, sinnvoll über  
die Lerner-  
sprache einer bestimmten Population<sup>90</sup> zu sprechen, so wie in dieser Arbeit etwa Lerner  
verschiedener Muttersprachen als Gruppe 'fortgeschrittener Lerner' zusammengefasst werden.

Lerner-  
sprache wird also als regelbestimmt und systematisch gedacht, zugleich jedoch als variabel  
und instabil. Es ist typisch für lernersprachliche Produktionen, dass sie einen relativ hohen Grad an  
Abweichungen von dem, was die Regeln vorhersagen, aufweisen; und auch die lernersprachlichen  
Regeln selbst sind nicht so stabil und beständig wie die Normen einer natürlichen Sprache einer  
festen Sprachgemeinschaft. Für die Variabilität innerhalb der Lerner-  
sprache führt Vogel (1990)  
mehrere Gründe an. Zunächst entwickeln sich Lerner-  
sprachen „im Idealfall immer im Kontakt mit  
einer Zielsprache und kontinuierlich auf deren Norm zu“.<sup>91</sup> Das heißt, dass sich die Sprache eines  
Lerners vor allem im Verlauf der Zeit ändert. Daneben sind lernersprachliche Produktionen in viel  
höherem Maße als muttersprachliche von den Vorkommensbedingungen der Produktionssituation  
abhängig, Vogel nennt zum Beispiel die Gegensatzpaare „Situationen im Unterricht / Situationen  
außerhalb des Unterrichts“, „gesprochene Sprache / geschriebene Sprache“, „Testsituation /  
natürliche Kommunikationssituation“ und vieles mehr.<sup>92</sup> Außerdem gibt es eine Variabilität in  
Abhängigkeit vom Lerner und zwar eine interpersonelle und eine intrapersonelle. Für erstere steht  
beispielsweise die Beobachtung, dass innerhalb einer Lernergruppe, etwa einer Schulklasse, nach

---

<sup>88</sup> Vgl. Vogel 1990: 48-50

<sup>89</sup> Ebd. S. 50

<sup>90</sup> „Wenn wir von der Sprache einer Lernerpopulation reden, so ist dies immer eine Abstraktion und Hypostasierung  
auf der Basis einer Beschreibung der grammatischen Strukturen und strukturellen Beziehungen, die sich in  
lernersprachlichen Ausdrücken finden, mit denen reale Lerner aus dieser Population umgehen, d.h. die sie  
produzieren, verstehen oder beurteilen“. (Kohn 1990: 40)

<sup>91</sup> Vogel 1990: 52

<sup>92</sup> Vgl. ebd. Kapitel 2.2.2

einer bestimmten Zeit des Unterrichts, an dem alle gleichermaßen teilgenommen haben, große Unterschiede in der Realisierung von Lernzielen bestehen können. Letztere zeigt sich in der Tatsache, dass „Lerner schwankend in der Sicherheit ihrer Wissensbestände sind (...), ein- und dasselbe grammatische Phänomen wird im Sinne der Norm manchmal falsch, manchmal richtig verwendet, d.h. Lernersprache ist durch *Instabilität* gekennzeichnet“.<sup>93</sup> Grund für die Instabilität kann sein, dass „die Wissensbestände des Lerners nicht genügend automatisiert und gefestigt sind, oder daß der Lerner sie fälschlicherweise für korrekt hält, bzw. ihnen, indem er sie aktualisiert, den Status zu testender Hypothesen über die von ihm vermutete Struktur der zu erlernenden Fremdsprache verleiht“.<sup>94</sup>

#### **2.2.2.2. Sind Lernersprachen natürliche Sprachen?**

Die Frage, ob Lernersprachen natürliche Sprachen sind oder nicht, lässt sich nicht einem knappen 'ja' oder 'nein' beantworten. Vogel (1990) führt zunächst sechs Punkte an, in denen sich Lernersprachen von natürlichen Sprachen unterscheiden<sup>95</sup>:

- 1) zur Lernersprache gehören sowohl korrekte als auch grammatisch inkorrekte Äußerungen und nicht-zielsprachenkonforme Regeln.
- 2) Lernersprachen sind kein expliziter Lerngegenstand, sie kommen nicht im Input des Lerners vor.
- 3) Es gibt keine allgemeinverbindlichen Normen für eine Lernersprache.
- 4) Am Entstehen von Lernersprache sind mindestens 2 weitere Sprachen beteiligt: Mutter- und Zielsprache.
- 5) Lernersprache ist keine *langue*, d.h. sie ist nicht „die Sprache einer Sprachgemeinschaft, sie ist streng genommen ein individuelles und nicht ein soziales Phänomen“.<sup>96</sup>
- 6) Lernersprachen werden beschrieben, um Erkenntnisse zu gewinnen über die Planungs- und Verarbeitungsprozesse, die dem Zweitsprachenerwerb zugrunde liegen.

Trotz all dieser Unterschiede zu den natürlichen Sprachen sind die Lernersprachen diesen aber – wenn schon nicht gleich – zumindest sehr ähnlich. Argumente dafür wurden weiter oben bereits genannt. Es sind vor allem die Universalien in den Erwerbssequenzen (ähnlich dem Erstspracherwerb) und die Systematizität der Lernersprache, die dafür sorgen, dass „der Lernersprache trotz

---

<sup>93</sup> Ebd. S. 54

<sup>94</sup> Ebd. S. 55

<sup>95</sup> Vgl. Vogel 1990: 13-15

<sup>96</sup> Vogel 1990: 14

ihres Interimscharakters der epistemologische Status einer natürlichen Sprache zu(kommt)<sup>97</sup>. Das hat zur Konsequenz, dass Lernaltersprachen „linguistisch beschreibbar“<sup>98</sup> sind und mit linguistischen Methoden, wie etwa der korpuslinguistischen Analyse in dieser Arbeit, untersucht werden können. In diesem Zusammenhang möchte ich noch eine kurze Anmerkung zum Punkt 6) in der obigen Aufzählung machen. Die Zweitspracherwerbsforschung hat mit der Fokussierung auf die Planungs- und Verarbeitungsprozesse, die dem Zweitspracherwerb und den lernaltersprachlichen Produktionen zugrunde liegen, vor allem eine psycholinguistische Ausrichtung erfahren. Die Frage nach den Prozessen, die hinter einem bestimmten Phänomen in der Produktion von Lernern stehen, ist hochkomplex und erfordert vielseitige Untersuchungen, die weit über die Möglichkeiten einer Korpusanalyse hinausgehen und den Rahmen dieser Arbeit sprengen würden. Deshalb beschränke ich mich hier auf die linguistische Beschreibung der lernaltersprachlichen *-ung*-Suffigierungen, denn es ist klar, „(...) daß lernaltersprachliche Strukturen Beschreibungsstrukturen sind, ohne die ein Reden über Prozesse nicht möglich wäre. Nicht zuletzt aus diesem Grunde verdient die Beschreibung von Lernaltersprache als Struktur Beachtung“.<sup>99</sup>

### **2.2.3. Der fortgeschrittene Lerner**

Die Lernerdaten, die in dieser Arbeit untersucht werden, stammen allesamt von fortgeschrittenen Lernern. Dies ist typisch für die Arbeit mit Lernerkorpora: während sich die Zweitspracherwerbsforschung traditionell eher mit dem frühen Zweitspracherwerb und seinen kognitiven Prozessen beschäftigt hat, steht bei der Arbeit mit Lernerkorpora meist der fortgeschrittene Fremdsprachenlerner im Mittelpunkt des Interesses.<sup>100</sup> Dies hat nach Walter/Grommes (2008) zum einen den „pragmatischen Grund, (...) dass Genres, die als Basis für solche Untersuchungen dienen – wie beispielsweise das Essay (...) über ein kontroverses Thema –, nicht von Anfängern produziert werden können“.<sup>101</sup> Zum anderen ist eine quantitative Analyse, die untersucht, ob Lerner bestimmte sprachliche Strukturen signifikant häufiger oder seltener benutzen als Muttersprachler, erst dann aussagekräftig, wenn die Lerner bereits über ein breites sprachliches Inventar verfügen, das ihnen überhaupt ermöglicht, zwischen verschiedenen sprachlichen Alternativen zu wählen, um einen Gedanken zu verbalisieren.

#### **2.2.3.1. Problem der Definition von 'Fortgeschrittenheit'**

<sup>97</sup> Ebd. S. 47

<sup>98</sup> Ebd. S. 13

<sup>99</sup> Kohn 1990: 30

<sup>100</sup> Vgl. Walter/Grommes 2008

<sup>101</sup> Ebd. S. 16

Obwohl die meisten Arbeiten mit korpuslinguistischen Analysen von Lernersprache angeben, auf Daten von fortgeschrittenen Lernern zu basieren, besteht keineswegs eine einheitliche Definition dessen, was ein 'fortgeschrittener Lerner' ist. Die Fortgeschrittenheit wird zum Teil ganz unterschiedlich festgestellt, wie etwa beim Falko-Korpus<sup>102</sup>, dessen Subkorpora verschiedene Kriterien für die Fortgeschrittenheit der Lerner zugrunde legen: beim Zusammenfassungskorpus etwa werden nur Texte von Lernern aufgenommen, die die DSH-Prüfung bestanden und ihr Grundstudium in Deutschland absolviert haben; die Texte im Essaykorpus stammen von Lernern, die in einem C-Test mindestens die Punktzahl erreicht haben, die man benötigt, um auf das Niveau C1 eingestuft zu werden.<sup>103</sup> Andere Korpora definieren Fortgeschrittenheit durch die Länge des gesteuerten Spracherwerbs der Lerner oder durch eine Kombination aus diesem Kriterium und weiteren Kriterien wie zum Beispiel der Beherrschung bestimmter sprachlicher Konstruktionen.<sup>104</sup> Obwohl jeder Forscher und auch jeder Fremdsprachenlehrer eine bestimmte Vorstellung von Fortgeschrittenheit für seine Arbeit voraussetzen muss, gibt es keine einheitliche Definition. Die Anwendung ging, wie Walter und Grommes betonen, „einer theoretisch wie empirisch fundierten Grundlagenforschung voraus“.<sup>105</sup> Einfache Definitionen, die etwa nur die Länge des Spracherwerbs berücksichtigen, greifen – auch wenn sie zunächst intuitiv nachvollziehbar sind – zu kurz. Schulleistungstests zeigen sehr deutlich, dass die Unterrichtszeit in einer Zweitsprache nicht bei allen Schülern zu denselben sprachlichen Fortschritten führt.<sup>106</sup> Auch eine Definition, die auf die Beherrschung bestimmter sprachlicher Mittel durch den Lerner abstellt, ist schwierig. Welche Phänomene müssen beherrscht werden, damit ein Lerner als fortgeschritten gelten kann? Und wann genau kann man davon sprechen, dass ein Lerner ein Phänomen wirklich 'beherrscht'? Die Antwort auf diese Fragen ist noch offen.<sup>107</sup> Deshalb lässt sich nur darüber spekulieren, wo ein sprachliches Phänomen wie die produktive Verwendung von Wortbildungsmustern einzuordnen wäre. Auch die Angaben zu den einzelnen Niveaustufen des „Gemeinsamen Europäischen Referenzrahmens“<sup>108</sup> helfen hier übrigens nicht weiter. Die Einteilung in die einzelnen Niveaustufen soll „das Sprachkönnen in den Bereichen der kommunikativen Kompetenz, der aktiven und rezeptiven Fertigkeiten und der Sprachkompetenz für spezifische Zwecke“ widerspiegeln. Eine verbindliche Liste von

---

<sup>102</sup> Eine Definition des Begriffs 'Korpus' und eine ausführlichere Vorstellung des Falko-Korpus folgt in Abschnitt 2.3. dieser Arbeit

<sup>103</sup> Vgl. Lüdeling et al. 2008: 68

<sup>104</sup> Vgl. Walter/Grommes 2008: 20

<sup>105</sup> Ebd. S. 4

<sup>106</sup> Vgl. Walter/Grommes 2008: 8

<sup>107</sup> Ebd. S. 9

<sup>108</sup> Quelle der Niveaustufenbeschreibungen: Rat für kulturelle Zusammenarbeit / Bildungsausschuss "Sprachenlernen für europäische Bürger": "Gemeinsamer europäischer Referenzrahmen für Sprachen: Lernen, lehren, beurteilen", [www.goethe.de/z/50/commeuro](http://www.goethe.de/z/50/commeuro)

grammatischen Kompetenzen, in der festgelegt ist, welche Bereiche der Grammatik auf welcher Niveaustufe zu beherrschen sind, gibt es nicht.

### 2.2.3.2 Fortgeschrittenheit im Zusammenhang mit der Basisvarietät

Walter/Grommes (2008) schlagen eine Definition vor, die mit dem Begriff der 'Basisvarietät' (BV) arbeitet. Das Basisvarietätenmodell, welches hier nur extrem verkürzt umrissen werden kann, geht unter anderem auf Klein & Perdue (1997) zurück und nimmt als Basisvarietät eine Entwicklungsstufe an, auf welcher Lerner bereits sehr effektiv kommunizieren können, obwohl ihr mentales Lexikon hauptsächlich Nomen, unflektierte Verben und einige wenige Adverbien und Partikel enthält. Es werden kaum zielsprachliche Strukturen produziert, vielmehr folgen die Lerner beim Formulieren einer Äußerung einer geringen Zahl von „elementary principles of their innate human language capacity“.<sup>109</sup> Diese Prinzipien erzeugen universale lernersprachliche Strukturen, die weder mit der Muttersprache noch mit der Zielsprache des Lerners oder der Universalgrammatik im Chomsky'schen Sinne vereinbar sein müssen. Die BV ist eine Entwicklungsstufe, die die meisten Lerner erreichen; viele fossilieren auf dieser Stufe<sup>110</sup>, ein möglicher Grund ist, dass Lerner, die die BV beherrschen, bereits vielfältige kommunikative Bedürfnisse befriedigen können.

Als Vorstufe zur BV nehmen Klein & Perdue eine Prä-Basisvarietät an „mit einem eingeschränkten lexikalischen Inventar, das im Wesentlichen aus Nomina besteht, einige verbähnliche Wörter ohne Flexionsmerkmale enthält und neben wenigen Temporal- und Lokaladverbien die Konnektoren *und* und *aber* aufweist sowie Pronomina der ersten und zweiten Person Singular“.<sup>111</sup>

Angelika Becker (2008), die den Erwerb der Negation im ungesteuerten Erwerb dreier italienischer Deutschlerner untersucht, beschreibt zwei Entwicklungsstufen, die über die BV hinausgehen. Die „Post-basic learner variety I“ charakterisiert sie vor allem durch das Auftreten von flektierten Auxiliar- und Modalverben, entscheidend für die „Post-basic learner variety II“ ist das Auftreten von flektierten lexikalischen Verben (im Gegensatz zur BV, wo die Lerner nur Infinitive verwenden).<sup>112</sup> Der Erwerb der Verbflexion geht einher mit dem Erwerb zielsprachlicher Satzstrukturen, welche sich deutlich von der Basisvarietät unterscheiden. Weitere Entwicklungsstufen, auf denen der Lerner schrittweise die BV hinter sich lässt, sind denkbar.

Der Ansatz von Walter/Grommes (2008) arbeitet genau mit dieser Idee: „Fortgeschritten ist ein

---

<sup>109</sup> Klein/Perdue 1997: 343

<sup>110</sup> Hier sei noch angemerkt, dass sich die Untersuchung von Klein & Perdue auf Daten von erwachsenen Lernern stützt, die die Fremdsprache im ungesteuerten Erwerb erlernt haben. Eventuell müssen für den gesteuerten Erwerb andere Annahmen gemacht werden.

<sup>111</sup> Walter/Grommes 2008: 11

<sup>112</sup> Becker 2008: 292

Lerner, wenn er die BV hinter sich gelassen hat und zunehmend zielsprachliche Äußerungsstrukturen produziert. Steigende Grade von Fortgeschrittenheit zeichnen sich durch ein stufenweises Verschwinden universaler lernersprachlicher Strukturen aus.“<sup>113</sup>

### **2.2.3.3. Fortgeschrittenheit in dieser Arbeit**

Über den Zusammenhang von Fortgeschrittenheit mit der Verwendung von produktiven Wortbildungsmustern gibt es meines Wissens keine Untersuchungen. Die Frage, wie fortgeschritten ein Lerner sein muss, damit er in der Lage ist, nach einem bestimmten Wortbildungsmuster selbst neue Wörter zu bilden, ist völlig offen. Bei Klein/Perdue (1997) wird lediglich erwähnt, dass auf der Stufe der Basisvarietät „(f)ormation of new words is limited to noun-noun compounds“.<sup>114</sup> Damit lässt sich also höchstens sagen, dass ein Lerner für die produktive Verwendung eines Wortbildungsmusters wie der *-ung*-Suffigierung die BV hinter sich gelassen haben muss. Rein intuitiv scheint nachvollziehbar, dass dies nicht von Anfängern beherrscht werden kann, sondern eher von fortgeschrittenen Lernern. Das spiegelt auch der Umgang damit in DaF-Lehrwerken wider.<sup>115</sup>

Die Korpusanalyse, die für die vorliegende Arbeit durchgeführt worden ist, greift auf Daten aus dem Essaykorpus von Falko zurück. Wie weiter oben bereits erwähnt, ist für dieses Subkorpus die Fortgeschrittenheit der Lerner mit einem C-Test ermittelt worden.<sup>116</sup> Es wurden nur Texte von Lernern aufgenommen, die beim C-Test einen bestimmten Wert überschritten hatten. Das Niveau der Lerner ist als Niveaustufe C1 des Gemeinsamen Europäischen Referenzrahmens angegeben. Beide Angaben (C-Test-Ergebnis, C1-Niveau) machen zunächst, wie gesagt, keinerlei Aussage darüber, wie sich ein Lerner in Bezug auf morphologische Wortbildungsregularitäten in seiner eigenen Produktion verhält, bzw. verhalten sollte. Eine solche Aussage – natürlich unter dem Vorbehalt, dass sie sich vor allem auf die im Korpus untersuchte Lernerpopulation bezieht – könnte vielmehr das Ergebnis der vorliegenden Arbeit sein.

## **2.3. Korpora, Lernerkorpora und das Falko-Korpus**

In den vorangegangenen Abschnitten wurde schon viel über die Arbeit mit Korpora im Allgemeinen sowie über das Untersuchungsziel dieser Arbeit gesagt, an etlichen Stellen wurde bereits das Falko-

---

<sup>113</sup> Walter/Grommes 2008: 12

<sup>114</sup> Klein/Perdue 1997: 332

<sup>115</sup> Vgl. die Einleitung zu dieser Arbeit

<sup>116</sup> Nähere Informationen zum C-Test gibt es auf der Homepage [www.c-test.de](http://www.c-test.de)

Korpus erwähnt. Ich habe dabei zum Teil ein gewisses Vorwissen vorausgesetzt, möchte aber an dieser Stelle noch mal eine kurze Definition für die – in der Literatur nicht immer ganz eindeutig gebrauchten – Begriffe „Korpus“ und „Lernerkorpus“ liefern<sup>117</sup> und das Falko-Korpus ausführlicher vorstellen.

### 2.3.1. Definition des Begriffs 'Korpus'

Unter einem Korpus versteht man ganz allgemein zunächst einmal eine „geschlossen verfügbare Sammlung von Texten einer bestimmten Kategorie“.<sup>118</sup> Nach Lüdeling et al. (2008: 68) gilt in der Korpuslinguistik, dass diese Textmenge „für einen bestimmten Forschungszweck nach vorher definierten externen oder internen Kriterien zusammengestellt wurde“. Obwohl es nicht zwingend zur Definition des Begriffs 'Korpus' gehört, gilt für die großen Korpora, dass sie computerlesbar sind und mit bestimmten Tools am Computer bearbeitet werden können. Im Idealfall kommen zu den Texten des Korpus noch Metadaten hinzu (etwa Informationen darüber, wann, wo, von wem, oder in welcher kommunikativen Situation ein Text verfasst wurde) sowie linguistische Annotationen (z.B. Wortart-Tagging, Fehlertagging<sup>119</sup>), die den Textdaten zugeordnet sind. Während die Korpusgröße in der Regel kein entscheidendes Kriterium für die Qualität eines Korpus ist, ist es umso wichtiger, dass der Inhalt des Korpus möglichst repräsentativ ist für den Bereich der Sprache, über den mithilfe der Korpusanalyse Aussagen gemacht werden sollen.<sup>120</sup> Um diese Repräsentativität zu gewährleisten, muss vor allem beim Aufbau des Korpus sorgfältig darauf geachtet werden, dass man Texte auswählt, die einem zuvor festgelegten Korpusdesign entsprechen. Korpora werden untersucht, um empirische Aussagen über den Sprachgebrauch von Sprechern einer Sprache oder Varietät (z.B. Dialekt, Soziolekt, Lernersprache) machen zu können. Eine Voraussetzung dafür ist es, die Korpusdaten in keinem Fall gleichzusetzen mit den Äußerungen von Sprechern oder den Texten eines Autors. Sie sind immer interpretierte Repräsentationen davon und je genauer die einzelnen Interpretationsschritte dokumentiert sind, desto leichter können Untersuchungen von anderen Forschern wiederholt und Ergebnisse reproduziert und überprüft werden.<sup>121</sup> Man unterscheidet verschiedene Arten von Korpora im Hinblick auf ihren Inhalt:

---

<sup>117</sup> Weitergehende Ausführungen finden sich z.B. in den Einführungen in die Korpuslinguistik von Scherer (2006) und Lemnitzer/Zinsmeister (2006)

<sup>118</sup> Lehmann 2006: 17

<sup>119</sup> „Die Annotation auf Wortebene wird nach dem englischen *to tag* 'mit einem Etikett versehen' als **Tagging** bezeichnet.“ (Scherer 2006: 58)

<sup>120</sup> Allerdings wirft die Frage nach der Repräsentativität eines Korpus auch etliche theoretische wie praktische Probleme auf; vgl. Lemnitzer/Zinsmeister 2006: 50-54

<sup>121</sup> „Mindestens soviel muss klar sein, dass ein Textkorpus jedenfalls kein Ausschnitt aus dem Phänomenbereich, sondern eine Repräsentation davon ist. Es enthält einerseits Abstraktionen über den Rohdaten und andererseits Interpretationen des Kompilators. Es ist essentiell, dass diese methodisch kontrolliert sind. Sonst kann man sich die

Korpora der gesprochenen und der geschriebenen Sprache, Korpora der Gegenwartssprache und historische Korpora, Referenz- und Spezialkorpora sowie ein- und mehrsprachige Korpora.<sup>122</sup> Zu den Spezialkorpora gehören die Lernerkorpora, die im folgenden Abschnitt näher erläutert werden.

### 2.3.2. Lernerkorpora

Lernerkorpora sind Korpora mit Daten von Lernern. Diese zunächst triviale Feststellung gilt es, um einige Erläuterungen zu ergänzen. Bei den Lernern handelt es sich um Lerner einer Fremdsprache, nach Granger (2008) also um „speakers who learn a language which is neither their first language nor an institutionalized additional language in the country where they live“.<sup>123</sup> Diese Definition mag für viele Sprachen unproblematisch sein, für das Englische ist sie es nicht, wie Granger zeigt.<sup>124</sup>

Ein weiteres Problem betrifft die Authentizität der Lerneräußerungen, die in den Lernerdaten repräsentiert sind. Authentisches Material kommt nach Sinclair (1996) aus „genuine communications of people going about their normal business“<sup>125</sup> und entsteht nicht unter experimentellen oder anderweitig künstlich geschaffenen Bedingungen. Allerdings ist die Beschaffung von so definiertem authentischen Material ein großes praktisches Problem, denn natürliche Kommunikation in Situationen des alltäglichen Lebens aufzuzeichnen, ohne dass die Sprecher davon etwas merken, ist eigentlich kaum zu bewerkstelligen. Grangers (2008) Vorschlag lautet daher, nicht zu dogmatisch an die Frage des authentischen Materials heranzugehen, sondern vor allem dafür zu sorgen, dass die Lerner zusammenhängende Texte produzieren und nicht nur einzelne Wörter oder Sätze:

„While it is important not to adopt an overly dogmatic attitude to this issue, it is essential to bear in mind that the notion of 'continuous text' lies at the heart of corpushood. A series of decontextualized words or sentences produced by learners (...) will never qualify as learner corpus data. In addition, it is best to restrict the term 'learner corpus' to the most open-ended types of tasks, viz. those tasks that allow learners to choose their own wording rather than being requested to produce a particular word or structure“.<sup>126</sup>

Daraus folgt, dass nicht alle Aufgabenstellungen, mit denen Lernerdaten elizitiert werden können, gleichermaßen für die Korpuserstellung geeignete Lernerdaten liefern. Aufgaben, die dem Lerner eine eigenständige Textproduktion abverlangen, wie das Schreiben von Textzusammenfassungen

---

Arbeit mit Korpora gleich schenken; sie würde uns der Empirie keinen Schritt weiter bringen“ (Lehmann 2006: 17)

<sup>122</sup> Vgl. Scherer 2006: Kapitel 2

<sup>123</sup> Granger 2008: 260

<sup>124</sup> Vgl. ebd. S.260; Die Definition scheint zumindest in einem Fall auch problematisch zu sein für das Deutsche, nämlich wenn man an Lerner des Deutschen als Fremdsprache im französisch- oder italienisch-sprachigen Teil der Schweiz denkt. Doch dieses Problem soll hier nicht weiter erörtert werden.

<sup>125</sup> Zitiert nach Granger 2008: 260

<sup>126</sup> Granger 2008: 261

oder Essays zu vorgegebenen Themen, führen eher zu den gewünschten Daten als etwa Lückentexte.

Granger betont ebenfalls die Wichtigkeit einer gezielten Auswahl der Daten bei der Erstellung des Korpus entsprechend einem vorab festgelegten Korpusdesign. Lernaltersprache wird durch eine Vielzahl von Faktoren (linguistische, situationelle und psycholinguistische) beeinflusst und je mehr dieser Faktoren kontrolliert werden können, desto verlässlicher sind die Aussagen, die auf der Korpusanalyse basieren. In das Korpusdesign können Variablen aufgenommen werden, die sich direkt auf die Lerner beziehen, wie z.B. das Alter der Lerner, ihre Muttersprache, ihre Fortgeschrittenheit im Fremdspracherwerb, oder auch Variablen, die die Erhebung der Daten betreffen, beispielweise die Aufgabenstellung, unter welcher ein Lerner einen Text produziert hat.<sup>127</sup> Lernalterskorpora bieten besondere Möglichkeiten auf der Grundlage von tatsächlich belegten Lerneräußerungen zu Erkenntnissen über die Lernaltersprache zu gelangen und zwar durch „genaue Betrachtung von Fehlern bzw. Abweichungen (in einem Kontext) und andererseits vor allem auch (durch) die Ermittlung quantitativer Eigenschaften von Lernaltersprache (im Vergleich zur L1 oder im Vergleich mit anderen Lernervarietäten)“.<sup>128</sup> Die Verfahren dafür sind die „computergestützte Fehleranalyse“ (CEA: computer-aided error analysis<sup>129</sup>) und die „kontrastive Lernaltersprachenanalyse“ (CIA: contrastive interlanguage analysis<sup>130</sup>). Beide Verfahren kommen auch in dieser Arbeit zum Einsatz bei der Analyse von Daten, die Falko – welches im folgenden Abschnitt vorgestellt wird – entnommen sind.

### 2.3.3. Falko – ein Lernalterskorpus für DaF

Falko (fehlerannotiertes Lernalterskorpus) ist eines der wenigen elektronisch verfügbaren Lernalterskorpora für DaF. Es ist im Internet frei zugänglich<sup>131</sup> und kann über ein Web-Interface abgefragt werden. Falko setzt sich aus mehreren Subkorpora zusammen:

- einem Zusammenfassungskorpus mit Textzusammenfassungen, die von fortgeschrittenen Lernern des Deutschen geschrieben wurden. (Dazu gibt es ein Erweiterungskorpus mit Textzusammenfassungen, die von dänischen Lernern des Deutschen in Kopenhagen erstellt

---

<sup>127</sup> Für eine umfassendere Auflistung der möglichen Korpusdesign-Kriterien vgl. Granger 2008: 293f

<sup>128</sup> Lüdeling et al. 2008: 67

<sup>129</sup> „It differs from previous error analysis studies in some major respects, not least of which is the fact that errors are not isolated from the texts in which they originated, as was the case in traditional EA, but rather are studied in context alongside cases of correct use and over- and underuse“ (Granger 2008: 267)

<sup>130</sup> „These studies have all helped to bring to light the words, phrases, grammatical items and syntactic structures that are over- or underused by learners and which therefore contribute to the foreign-soundingness of perhaps otherwise error-free advanced interlanguage“ (Granger 2008: 267)

<sup>131</sup> (<http://www.linguistik.hu-berlin.de/institut/professoren/korpuslinguistik/forschung/falko>)

wurden, ein Vergleichskorpus mit Texten von Muttersprachlern und die Sammlung der Vorlagentexte.)

- einem Longitudinalkorpus mit Daten, die über mehrere Semester und Lernstände an der Georgetown-Universität in Washington erhoben wurden.
- einem Essaykorpus (Essays L2, Version 1.0) mit Essays von fortgeschrittenen Lernern<sup>132</sup> zu vier kontrovers zu diskutierenden Themen („Entlohnung“, „Kriminalität“, „Feminismus“ und „Studium“<sup>133</sup>). Diesem zugeordnet ist ein Vergleichskorpus (Essays L1, Version 0.5) mit Texten von Berliner und Brandenburger Gymnasiasten (12. und 13. Klasse).

Das Essaykorpus wird für die Korpusanalyse in dieser Arbeit verwendet. Es enthält Texte von ausländischen Germanistik-Studenten, die an Ferienkursen an der HU (Humboldt-Universität zu Berlin) und FU (Freie Universität Berlin) teilgenommen haben, sowie Daten, die an Partneruniversitäten in Dänemark, Kenia, Usbekistan, der Türkei und Südafrika erhoben wurden. Alle Texte in Falko sind automatisch nach Wortart<sup>134</sup> und Lemma getaggt und anschließend überprüft und teilweise manuell korrigiert worden. Darüber hinaus können und sollen Fehler, die von den Lernern in ihren Texten gemacht wurden, annotiert werden (das Zusammenfassungskorpus und Longitudinalkorpus sind mit syntaktischen Feldern und Fehlern annotiert). Dafür sind die Annotationen in Falko in einer Mehrebenenarchitektur angelegt, die es ermöglicht jederzeit beliebig viele voneinander unabhängige Annotationsebenen einzufügen, in denen dann beispielsweise eine oder mehrere Zielhypothesen und entsprechende Fehlertags annotiert werden können. Auf die Notwendigkeit einer Zielhypothese für die Fehleranalyse und die damit verbundenen Schwierigkeiten komme ich in Abschnitt 2.4.2. zurück.

Das Essaykorpus in Falko ist kein statisches Korpus, es wächst, d.h. es werden in Zukunft noch weitere Daten hinzugefügt. Folgende Korpusgröße war aktuell zum Zeitpunkt der Auswertung für diese Arbeit:

<b>Subkorpus</b>	<b>Texte</b>	<b>Tokens</b>	<b>Types</b>
Essays L2, Version 1.0	132	65387	7720
Essays L1, Version 0.5	39	33806	5831

Der Begriff „Token“ bezieht sich auf das konkrete Vorkommen einer sprachlichen Einheit (hier: einer Wortform) im Korpus. „Type“ dagegen ist die abstrakte sprachliche Einheit (hier: Lemma-

<sup>132</sup> Die Fortgeschrittenheit wurde mit einem C-Test ermittelt. Nur Texte von Lernern, die damit einem Sprachniveau von C1 des Gemeinsamen Europäischen Referenzrahmens zugeordnet werden konnten, wurden in das Korpus aufgenommen. (Vgl. auch Abschnitt 2.2.3.3. dieser Arbeit)

<sup>133</sup> Vgl. Metainformation zum Lerner in den Falko-Daten

<sup>134</sup> Dazu wird der TreeTagger verwendet mit dem Stuttgart-Tübingen-Tagset (<http://www.ims.uni-stuttgart.de/projekte/corplex/TreeTagger/>)

Type), die zusammengehörige Tokens zusammenfasst.

Im Übrigen sind jedem Text in Falko umfangreiche Metadaten hinzugefügt. Sie geben Auskunft über Geschlecht, Alter, Studienrichtung und Muttersprache des Lerners sowie darüber, welche Fremdsprachen und wie lange der Lerner deutsch gelernt hat. Schließlich ist im Essay-Korpus noch das Aufsatzthema angegeben. Das folgende Beispiel aus Falko veranschaulicht, wie die Metainformationen dort notiert sind:

<b>Metainformation zum Lerner:</b>	
<b>Id</b>	cbs001
<b>Muttersprache(n)</b>	da
<b>Geburtsjahr</b>	1982
<b>Geschlecht</b>	m
<b>Studienrichtung</b>	Deutsch
<b>Fremdsprache(n)</b>	de
<b>Erwerbsdauer der deutschen Fremdsprache</b>	10
<b>Aufsatzthema</b>	Studium

**Abb.3:** Beispiel für die Metainformation zu den Lernertexten im Falko-Subkorpus „Essays L2, 1.0“

Entsprechend sehen die Metadaten für die Schreiber der Texte im L1-Vergleichskorpus aus:

<b>Metainformation zum Textautor:</b>	
<b>Id</b>	dhw001
<b>Muttersprache(n)</b>	de
<b>Geburtsjahr</b>	1991
<b>Geschlecht</b>	f
<b>Studienrichtung</b>	None
<b>Fremdsprache(n)</b>	en
<b>Aufsatzthema</b>	Entlohnung

**Abb.4:** Beispiel für die Metainformation zu den Textautoren im Vergleichskorpus „Essays L1, 0.5“

Mit diesen Metadaten und den allgemeinen Informationen über die Art und das Zustandekommen der Lernertexte sind alle wesentlichen Variablen bedacht, die nach Granger (2008: 263f) als Teil des Korpusdesigns eines aussagekräftigen Korpus kontrolliert sein sollten.

## 2.4. Fehleranalyse und Fehlerklassifikation

In Abschnitt 2.1.1. wurde die Fehleranalyse im Rahmen der Entwicklung des wissenschaftlichen Konzeptes der Lernaltersprache bereits erwähnt. Populär geworden in den 1970er Jahren nach den Veröffentlichungen Pit Corders, wurde sie später als unzureichend für die Erforschung des Zweitsprachenerwerbs angesehen und verlor an Bedeutung. Gleichwohl gibt es immer noch eine Vielzahl an Untersuchungsgebieten und Fragestellungen, für die eine Fehleranalyse durchaus von großem Nutzen sein kann.<sup>135</sup> Vor allem für die Sprachpädagogik, so Ellis (2005: 70), bleibt die Untersuchung jener Bereiche einer Sprache, die einem Lerner offenbar Schwierigkeiten bereiten und zu Fehlern führen, von praktischer Bedeutung. Computergestützte Fehleranalysen, wie sie durch den Gebrauch von elektronischen Lernerkorpora möglich sind, haben außerdem einen entscheidenden Vorteil gegenüber den Fehlersammlungen vergangener Tage: Fehler werden nicht mehr isoliert von den Texten, in denen sie produziert wurden, betrachtet, sondern erscheinen im Kontext zusammen mit den Fällen von korrektem Gebrauch und einer quantitativen Auswertung von Über- und Mindergebrauch.<sup>136</sup> Dadurch, so Granger (2008: 266) ist die Fehlerannotation von Korpora sogar besonders wichtig für Lernaltersprachenstudien und erfreut sich wachsender Beliebtheit. Nichtsdestotrotz ist die Fehleranalyse mit einigen theoretischen und methodologischen Schwierigkeiten behaftet, über die es sich im Klaren zu sein gilt. Da ist zunächst die Problematik einer Definition des Begriffes „Fehler“, welche im folgenden Abschnitt (2.4.1.) erläutert wird. Dann geht es darum, Fehler zu beschreiben und zu kategorisieren (Abschnitt 2.4.2.). Zuletzt wird erklärt, was bei der Fehlerannotation und bei der Entwicklung eines Fehlertagsatzes zu beachten ist (Abschnitt 2.4.3.).

### 2.4.1. Was ist ein „Fehler“?

Es gibt nach wie vor das Desiderat einer Definition des Begriffes „Fehler“, welche in der Literatur häufig und kontrovers diskutiert worden ist, ohne dabei zu einer generellen Übereinstimmung zu kommen, was letztendlich als Fehler zu bezeichnen sei.<sup>137</sup> Corder (1971) fordert, eine Unterschei-

---

<sup>135</sup> „So darf auch nicht überraschen, wenn eine Lernaltersprachenanalyse, die als Fehleranalyse konzipiert und damit in ihrem Erhebungsinhalt an der zielsprachlichen Norm orientiert ist, über die Sprache des Lernalters nur bedingt Aufschluß geben kann. Dies bedeutet natürlich nicht, dass eine solche Erhebung wertlos ist. Man muß sich nur darüber im Klaren sein, daß die Erhebungsfragen, auf die sie eine Antwort erlaubt, eben das Vorkommen zielsprachlicher Eigenschaften in der Lernaltersprache betreffen.“ (Kohn 1990: 39)

<sup>136</sup> Vgl. Granger 2008: 268

<sup>137</sup> Für Literaturverweise zur Diskussion des Fehler-Begriffes siehe Ellis (2005) und Lüdeling (2007)

dung zwischen „errors“ und „lapses“ vorzunehmen. Errors sind definiert als Kompetenzfehler, die ein Lerner macht, weil er eine bestimmte Regel oder Norm nicht kennt; lapses<sup>138</sup> dagegen sind Performanzfehler, die ein Lerner macht, – obwohl er die Regeln und Normen, die er zu befolgen hat, kennt – weil er müde oder unkonzentriert ist. Wenngleich diese theoretische Unterscheidung sicher sinnvoll ist, so ist sie doch bei einer Korpusuntersuchung in der Regel wenig praktikabel, da errors und lapses an der Oberfläche nicht zu unterscheiden sind.<sup>139</sup> Aus diesem Grund werde ich in der vorliegenden Studie auf die Unterscheidung verzichten und durchgängig den Terminus „Fehler“ verwenden.

Ganz generell, so Ellis (2005: 56), gilt es bei der Definition des Fehler-Begriffes zu unterscheiden, ob Grammatikalität (in Bezug auf Wortbildung also Wohlgeformtheit) oder Akzeptabilität als Kriterium zugrunde gelegt werden soll. Wohlgeformtheit, wonach ein Fehler also der Verstoß gegen eine Regel des Sprachsystems wäre, scheint das klarere Kriterium zu sein – allerdings nur, wenn für den zu untersuchenden Bereich der Sprache klare Regeln existieren, bzw. bekannt sind. Für große Teile der Syntax mag dies zureffen, im Bereich der Wortbildung scheint es schon weniger klar zu sein. Für die *-ung*-Suffigierung fehlt bislang eine umfassende Beschreibung der Regeln zur Bildung von Derivaten und den wirksamen Beschränkungen (vgl. Abschnitt 2.1.7.). Zudem ist bei bestimmten Beschränkungen wie der Blockierung nicht ganz klar, ob es sich hier um Regeln des Sprachsystems oder um Normen des Sprachgebrauchs handelt. Die Tatsache, dass Blockierungen (vor allem das token-blocking) kein absolutes Phänomen sind und die Blockierungen oftmals nicht wirksam werden, legt nahe, dass es sich dabei um Normen des Gebrauchs handelt. Ein Fehler wäre demnach ein Verstoß gegen eine Norm, also nach dem Kriterium der Akzeptabilität definiert. Eine Fehler-Definition, die weitgehend ohne eine Unterscheidung von Grammatikalität und Akzeptabilität auskommt, bietet Lennon (1991: 182):

„A linguistic form of combination of forms which, in the same context and under similar conditions of production, would, in all likelihood, not be produced by the speakers' native counterparts“.<sup>140</sup>

Auch diese Definition wirft eine Menge Fragen auf (vgl. Lüdeling 2007: 35). Ellis (2005: 56) bezeichnet sie ebenfalls als „not unproblematic“, hält sie aber dennoch für eine „reasonable basis for performing an EA“. Auch in der vorliegenden Arbeit werden auf dieser Basis Entscheidungen darüber, was als Fehler zu werten ist, getroffen – zumindest in den Fällen, in denen nicht ein klarer

---

<sup>138</sup> Statt „lapses“ wird oft auch der Terminus „mistakes“ gebraucht, den Ellis (1995: 51) so definiert: „A mistake occurs when learners fail to perform their competence. (...) Mistakes, then, are performance phenomena and are, of course, regular features of native-speaker speech, reflecting processing failures that arise as a result of competing plans, memory limitations, and a lack of automaticity.“

<sup>139</sup> Vgl. Lüdeling 2007: 35

<sup>140</sup> Zitiert nach Ellis 2005: 56

Fall von Ungrammatikalität, der mit bekannten Regeln zu erklären ist, vorliegt. Zwei Beispiele aus dieser Arbeit mögen dies verdeutlichen. Das von einem Lerner gebildete Derivat *Ermorderung* (Falko Essays L2 1.0, fk007\_2006\_08) ist eindeutig falsch, weil hier nicht der korrekte Verbstamm *ermord-* für die Bildung des Derivats zugrunde gelegt wird, sondern ein Verbstamm *ermorder-*, den es im Deutschen nicht gibt. In dem Satz *obligatorische Praktika gibt es fast nicht an den französischen Universitäten, Treffungen zwischen den Studenten und den Arbeitgebern während der Studiumperiode auch nicht, usw.* (Falko Essays L2, 1.0, fk001\_2006\_08) ist das Substantiv *Treffungen* im Prinzip regelkonform gebildet (Verbstamm + Suffix *-ung*). Trotzdem, davon gehe ich aus, wird dieses Substantiv von deutschen Muttersprachlern so nicht gebildet, da die Bildung durch das bestehende Lexem *Treffen* blockiert ist. Dieser Entscheidung wohnt, wie in der Definition von Lennon angelegt, ein gewisses Maß an Subjektivität inne. Dies ist als Grundlage für eine Fehleranalyse sicher nicht ideal, aber meines Erachtens gibt es derzeit keine Lösung für das Problem, solange keine eindeutige Fehlerdefinition existiert. Die vielen offenen Fragen, die sich im Zusammenhang mit der Fehler-Diskussion stellen, können in dieser Arbeit nicht beantwortet werden.

Beachtung verdient nach Ellis (2005: 56) die Unterscheidung zwischen „sichtbaren“ („overt“) und „verdeckten“ („covert“) Fehlern. Danach ist ein Fehler sichtbar, wenn er bei der Untersuchung des einen Satzes, in dem er auftritt, entdeckt werden kann. Ein verdeckter Fehler kann nur erkannt werden, wenn ein längerer Abschnitt des Diskurses untersucht wird. Diese Unterscheidung bei Ellis gilt für die Syntax. Für das Wortbildungsphänomen der vorliegenden Studie kann gesagt werden, dass sichtbare Fehler an einem einzelnen Wort erkennbar sind, wie in dem Beispiel *Ermoderung*. Um verdeckte Fehler zu erkennen, muss ein ganzer Satz – manchmal auch mehrere Sätze – in Augenschein genommen werden, wie in dem folgenden Beispiel aus Falko: (...) *um sicher zu sein, dass die Studierenden für die Lesungen (Unterrichten) gehen* (Falko Essays L2, 1.0, kne08\_2006\_09). Das *-ung*-Derivat *Lesung* existiert selbstverständlich, aber mit der Bedeutung „Buchvorstellung eines Autors“ und nicht, wie vom Lerner hier intendiert, mit der Bedeutung „Vorlesung an der Universität“. Sichtbare Fehler werden sich also eher auf die Bildungs- und/oder Schreibweise eines Wortes beziehen, als verdeckte Fehler werden vor allem semantische Fehler auftreten.

#### **2.4.2. Fehleridentifizierung, -beschreibung und -kategorisierung**

Nach Ellis (2005: 57) nimmt Corder (1974) eine Unterteilung der Fehleranalyse in fünf Schritte vor:

1. Collection of a sample of learner language (Sammlung von Lernerdaten)
2. Identification of errors (Fehleridentifizierung)
3. Description of errors (Fehlerbeschreibung)
4. Explanation of errors (Erklärung der Fehler)
5. Error evaluation (Bewertung der Fehler)

Die Erklärung und Bewertung von Lernerfehlern ist ein schwieriges Unterfangen, das eine Vielzahl höchst komplexer linguistischer und psycholinguistischer Faktoren miteinbeziehen muss. Das kann und soll von der vorliegenden Studie nicht geleistet werden. Sie folgt Corder in den Schritten 1-3 und ergänzt einen vierten Schritt, der die Erstellung eines Fehlertagsets (Lernerfehler bei der Produktion von *-ung*-Nominalisierungen) für Falko beinhaltet. Der erste Schritt erübrigt sich insoweit, als keine eigene Lernerdatensammlung erstellt werden muss, sondern auf die Daten aus Falko zugegriffen werden kann. Allerdings müssen aus der Gesamtmenge der Falko-Daten noch die konkreten Belege für das zu untersuchende Wortbildungsphänomen herausgesucht werden. In diesen Belegen gilt es dann die Fehler zu identifizieren. In Übereinstimmung mit der in Abschnitt 2.4.1. gegebenen Fehlerdefinition erfordert die Identifizierung der Fehler einen Vergleich zwischen dem, was der Lerner produziert hat und dem, was ein Muttersprachler in demselben Kontext produzieren würde.<sup>141</sup> Es muss also eine möglichst plausible Hypothese darüber gebildet werden, welchen Inhalt der Lerner versprachlichen wollte und wie ein Muttersprachler dies in der Regel versprachlichen würde. Diese Hypothese nenne ich mit Lüdeling et al. (2008: 70) „Zielhypothese“; ein Fehler ist danach auch definiert als „als Abweichung von einer Zielhypothese, welche (...) auf der Basis der Lerneräußerung rekonstruiert wird“. Ausgehend von dieser Zielhypothese wird also entschieden, was an einer Lerneräußerung fehlerhaft ist, wie diese Abweichung von einer Norm beschrieben werden kann und, daran anschließend, welches Fehlertag zugewiesen werden muss. Die Zielhypothese ist also das zentrale Element in einer Fehleranalyse. Leider gilt für Zielhypothesen, dass sie „fast nie unumstritten (sind), meistens sind mehrere Zielhypothesen für eine Lerneräußerung möglich“.<sup>142</sup> Daraus leitet sich die Forderung ab, dass Zielhypothesen immer explizit angegeben werden sollten – sowohl in Korpusstudien als auch bei der Fehlerannotation im Korpus selbst.<sup>143</sup> Für diese Studie gilt sogar, dass Zielhypothesen bereits bei der Kategorisierung der Belege, also bei der Zuordnung der *-ung*-Derivate zu den Wortbildungsmustern der *-ung*-Suffigierung, gebraucht werden.<sup>144</sup>

Sind die Fehler mit Hilfe der Zielhypothesen identifiziert, folgt die Beschreibung und Kategori-

---

<sup>141</sup> Vgl. Ellis 2005: 58

<sup>142</sup> Lüdeling 2007: 36

<sup>143</sup> Vgl. ebd. S. 37

<sup>144</sup> Vgl. Abschnitt 3.1.1.

sierung der Fehler. Dafür gibt es kein bestehendes Inventar an Beschreibungskategorien, d.h. jeder Forscher muss für seine Studie eine eigene Fehlertaxonomie (System von Kategorien) erstellen. Diese sollte im besten Fall sowohl gut ausgearbeitet und in der Lage sein, feine Unterschiede zu benennen, als auch schlicht und selbsterklärend sein.<sup>145</sup> Lüdeling (2007: 39) benennt verschiedene Kriterien, nach denen eine Fehlertaxonomie erstellt werden kann:

„(...) (M)an kann zum Beispiel nach formaler Art eines Fehlers (Einfügung, Auslassung, Vertauschung) kategorisieren oder die Hypothese über die Fehlerentstehung (zum Beispiel Interferenz der Muttersprache des Lerners oder die mangelnde Beherrschung einer Regel) in das Fehlertag mit einfließen lassen. Man kann die Wortart des Fehlerexponenten<sup>146</sup> mit kodieren oder in bestimmten linguistischen Bereichen arbeiten.“

Wichtig ist dabei vor allem, dass die Kategorien auf der Grundlage der Lernerdaten entwickelt werden (anstatt sich mit einer vorab ausgearbeiteten Taxonomie den Daten zu nähern) und auf einer deskriptiven Grammatik der Zielsprache basieren.<sup>147</sup> Mithilfe dieser Beschreibungskategorien kann im Anschluss an die Fehlerklassifikation ein Fehlertagset für das Korpus, dem die Daten entnommen wurden, entwickelt werden.

### 2.4.3. Entwicklung eines Fehlertagsets

Die Annotation von Korpusdaten macht ein Korpus vielseitig nutzbar, weil sich konkrete Such- und Abfragemöglichkeiten bieten, die über die schlichte Suche nach einer Zeichenkette hinausgehen. Ein Korpus kann als eine Art „Beispielbank“ (etwa zur Suche nach authentischen Beispielen von Lernerfehlern bei der Wortbildung) verwendet werden, es kann für qualitative Untersuchungen dienen (z.B. der Frage, welche Wortbildungsfehler Lerner machen) und darüber hinaus für quantitative Analysen gebraucht werden (z.B. Zählungen eines bestimmten Fehlertyps).<sup>148</sup> Für alle diese Untersuchungszwecke, vor allem aber für die quantitative Analyse, ist es von besonderer Wichtigkeit, dass die Korpusdaten korrekt kategorisiert und annotiert sind und die Erstellung des Tagsets sowie die Annotation damit transparent dokumentiert wird.

Für die Annotation von Lernerfehlern in einem Lernerkorpus braucht man ein Fehlertagset, das unter Berücksichtigung der folgenden Kriterien (nach Lüdeling 2007: Abschnitt 4.3.) entwickelt werden sollte:

1. Fehlerexponent

---

<sup>145</sup> Vgl. Ellis 2005: 60

<sup>146</sup> Der „Fehlerexponent“ ist die Stelle, an der ein Fehlertag 'andockt'.

<sup>147</sup> Vgl. Ellis 2005: 60

<sup>148</sup> Vgl. Lüdeling 2007: 28

## 2. Granularität

## 3. Beschreibung des Fehlers

Zunächst muss also entschieden werden, an welches Wort oder welche Phrase ein Fehlertag angehängt wird. Dies kann je nach Fehlerart unterschiedlich problematisch sein, bei Wortstellungsfehlern ist die Entscheidung sicher schwieriger als etwa bei Wortbildungsfehlern, wo das Tag einfach an das falsch gebildete bzw. falsch gebrauchte Wort angehängt werden kann. Dann muss man sich entscheiden, wie fein bzw. wie grob das Tagset sein soll, also mit welcher Genauigkeit Fehlerarten getaggt werden sollen, wie viele Tags es dafür geben soll usw.<sup>149</sup> Die Frage nach der Granularität ist natürlich im Zusammenhang mit der Fehlerbeschreibung zu beantworten. Es muss geprüft werden, welche Informationen aus der Fehlerbeschreibung sich letztendlich im Tagset wiederfinden sollen und wie sie dort gebündelt werden.

### 2.5. Zwischenfazit

Kapitel 2 dieser Arbeit hat den theoretischen Hintergrund, vor welchem die Korpusanalyse dieser Studie stattfindet, dargelegt. An den Anfang wurden einige grundsätzliche Annahmen über die Wortbildung im Allgemeinen und die *-ung*-Suffigierung im Besonderen gestellt. Darauf folgte eine Definition des Begriffs „Wortbildungsmuster“ nach Motsch (1999), eine Übersicht über die verschiedenen Wortbildungsmuster der *-ung*-Suffigierung sowie einige Überlegungen zu der offenen Frage, wann eigentlich beim Sprechen von Wortbildungsmustern Gebrauch gemacht wird und wann gespeicherte Lexikoneinheiten abgerufen werden. Einer Annäherung an das schwer zu definierende Konzept der morphologischen Produktivität, ausgehend von einer Definition von Bauer (2001), folgten Aussagen aus diversen traditionellen Grammatiken zu Bildung, Bildungsbeschränkungen und Produktivität der *-ung*-Suffigierungen. Die abschließende Feststellung lautete, dass diese Informationen ein recht diffuses und noch unvollständiges Bild von den Regeln vermitteln, die der Bildung von *-ung*-Derivaten zugrunde liegen. Dies hat zur Folge, dass man als Linguist bei der Fehleridentifizierung in Lerneräußerungen teilweise introspektiv arbeiten und sich auf sein eigenes Sprachgefühl verlassen muss, wenn es um die Akzeptabilität von Wortbildungsprodukten der Lerner geht. Die Angaben zur Produktivität der Wortbildungsmuster der *-ung*-Suffigierung, die Motsch (1999) entnommen wurden, werden anhand der Produktivitätsmessung in dieser Studie überprüft. Dafür wurde ein quantitatives Produktivitätsmaß von Baayen vorgestellt. Ein großer Abschnitt widmete sich den theoretischen Annahmen über „Lernersprache“ und einer Erklärung, warum Lernersprache als Untersuchungsgegenstand einer linguistischen Korpusanalyse

---

<sup>149</sup> Zu den Schwierigkeiten, die damit verbunden sind, siehe Lüdeling (2007: 38f).

geeignet ist. In diesem Zusammenhang wurde auch auf die Probleme, die mit dem Konzept des „fortgeschrittenen Lerner“ bzw. einer Definition von „Fortgeschrittenheit“ verbunden sind, hingewiesen.

Darauf folgten einige grundlegende Angaben zur Arbeit mit Korpora und insbesondere Lernerkorpora, den Möglichkeiten, die eine Korpusanalyse bietet, sowie eine Vorstellung des Falko-Korpus, welches das Material für diese Studie liefert.

Der letzte Abschnitt brachte Hintergrundwissen für das Vorgehen im empirischen Teil dieser Arbeit. Eine Fehleranalyse ist ein methodologisch komplexes Unterfangen, das mit etlichen Schwierigkeiten wie der Uneinheitlichkeit einer Fehler-Definition oder dem Fehlen von klaren Vorgaben für die Kategorien der Fehlerbeschreibung und -klassifikation verbunden ist. Ein kurzer Ausblick auf das, was es bei der Entwicklung eines Fehlertagsets zu beachten gibt, bildete den Abschluss dieser theoretischen Vorüberlegungen.

### **3. Empirischer Teil**

Der empirische Teil dieser Studie enthält, wie angekündigt, zwei Untersuchungen: zum einen die quantitative Analyse der Produktivität der einzelnen Wortbildungsmuster der *-ung*-Suffigierung, wobei hier der Gebrauch von fortgeschrittenen DaF-Lernern mit dem Gebrauch von deutschen Muttersprachlern verglichen wird (Abschnitt 3.1.). Zum anderen werden die Fehler, die die Lerner machen, in einer qualitativen Analyse genauer betrachtet und kategorisiert. Anschließend wird für diese Kategorien ein Fehlertagset entwickelt, welches für die zukünftige Fehlerannotation in Falko verwendet werden könnte (Abschnitt 3.2.). In Abschnitt 3.3. werden die wichtigsten Ergebnisse der beiden Untersuchungen kurz zusammengefasst.

#### **3.1. Quantitative Analyse – Produktivität der Wortbildungsmuster der *-ung*-Suffigierung**

Die quantitative Analyse gliedert sich in drei Teile. Der erste Teil (Abschnitt 3.1.1.) erläutert detailliert die Auswahl der Belege, welche für die anschließende Berechnung der Produktivität verwendet werden. Dazu gehört zum einen die konkrete Korpusabfrage, zum anderen die Interpretation der *-ung*-Derivate, um sie den vier Wortbildungsmustern der *-ung*-Suffigierung zuordnen zu können. Abschnitt 3.1.2. enthält die Zahlen dieser Zuordnung und, basierend darauf,

die Berechnung der „potential productivity“<sup>150</sup> der einzelnen Wortbildungsmuster. Eine ausführliche Diskussion der Ergebnisse folgt in Abschnitt 3.1.3. Hier werden die Ergebnisse außerdem mit den Angaben von Motsch (1999) zur Produktivität der Wortbildungsmuster verglichen.

### 3.1.1. Auswahl, Kategorisierung und Zählung der Belege

Die Datengrundlage für die Auswahl der Belege, welche für die quantitative Analyse gezählt wurden, ist das Falko-Korpus, genauer gesagt das Subkorpus „Essays L2, Version 1.0“ sowie das dazugehörige Vergleichskorpus „Essays L1, Version 0.5“.<sup>151</sup> An beide Korpora wurde folgende Suchabfrage gestellt: [word=".\*ung.\*" & apos="NN"]. Diese Abfrage liefert als Ergebnis alle Substantive, in denen die Zeichenkette *ung* vorkommt. In manueller Korrektur wurden alle Wörter, bei denen es sich nicht um *-ung*-Derivate handelt, aussortiert, z.B. *Junge*, *Vorsprung*, *Hunger* usw. Die *-ung*-Derivate wurden anschließend den vier Wortbildungsmustern der *-ung*-Suffigierung (nomina actionis, resultierender Zustand, nomina acti und nomina loci)<sup>152</sup> zugeordnet und gezählt. Dies klingt zunächst nach einer unproblematischen Aufgabe, es stellen sich hierbei jedoch einige methodologische Fragen, die beantwortet werden müssen, wenn Ergebnisse der Zählung und letztendlich der ganzen quantitativen Analyse transparent und nachvollziehbar sein sollen. Auch Plag (1999: 28) betont diese Problematik im Zusammenhang mit der Messung der „potential productivity“:

„For the calculation of P (and other measures) it is necessary to count all words with a given affix in the corpus. This looks like a rather straightforward task, but in practice this may involve complex decisions, pivoting around two problems. The first is to determine which words can be considered to bear the affix in question, the second is to control for cohort effects in multiple affixation“.

Gerade die beiden hier von Plag genannten Schwierigkeiten spielen für die *-ung*-Suffigierung keine Rolle. Die Identifikation des Suffixes ist weitgehend unproblematisch, da es keine homonymen Affixe im Deutschen gibt. Mehrfachaffigierung tritt bei *-ung*-Derivaten ebenfalls nicht auf.<sup>153</sup> Dafür stellen sich im Zusammenhang mit der *-ung*-Suffigierung und speziell auch der Analyse von Lernerdaten andere Fragen, die ich im Folgenden erörtern will.

Zunächst einmal stellt die Zuordnung der *-ung*-Derivate zu den einzelnen Wortbildungsmustern keine leichte Aufgabe dar. Vor allem die Unterscheidung zwischen nomina actionis und nomina acti fällt oftmals schwer. Dies liegt daran, dass viele nomina acti durch den Prozess der Lexikalisierung

<sup>150</sup> Vgl. Abschnitt 2.1.8.

<sup>151</sup> Vgl. Abschnitt 2.3.3.

<sup>152</sup> Vgl. Abschnitt 2.1.5.1.

<sup>153</sup> Anders sieht dies im Fall des Nominalisierungssuffixes *-er* aus, vgl. Scherer (2005).

aus nomina actionis heraus entstanden sind<sup>154</sup> und etliche homonyme *-ung*-Derivate existieren, die beide Bedeutungen tragen können. Ein Beispiel aus dem Essays L2-Korpus in Falko ist das Nomen *Lösung*. In dem Satz *Meiner Ansicht nach, können härtere Gesetze bei der Lösung des Problems kaum helfen* (Essays L2, 1.0, fk019\_2006\_08) ist *Lösung* ein nomen actionis, es geht hier um den „Prozess des Lösens“ eines Problems. Im Gegensatz dazu ist *Lösung* in dem Satz *Eine Lösung wäre vielleicht, dass in jedem Bereich die Menschen dieselbe Entlohnung bekommen, seien sie Künstler oder Brückenbauer* (Essays L2, 1.0, hu003\_2006\_09) ein nomen acti, gemeint ist hier eine bestimmte, konkrete Lösung als Thema in dem von Verb bezeichneten Geschehen.

Neben diesem Problem der Interpretation und Zuordnung von homonymen *-ung*-Derivaten gibt es den Fall, dass ein *-ung*-Derivat keinem der Wortbildungsmuster zugeordnet werden kann, weil die vier von Motsch (1999) vorgeschlagenen Muster nicht alle Bildungen abdecken. Der Vorteil seiner Beschränkung auf die vier Muster liegt in der Übersichtlichkeit, andere Autoren, die den Anspruch haben, alle Bildungen erklären zu können, müssen weitaus mehr Wortbildungsmuster annehmen. So gibt es bei Motsch kein Muster, das dem Musters *-ung*<sup>5</sup> bei Wellmann (1975) entspricht. Dieses Muster führt zu *-ung*-Derivaten mit einer Agens-Bezeichnung wie z.B. *Bedienung* in der Bedeutung „derj., der jmdn. bedient“ oder *Erscheinung* als „etw., das erscheint“.<sup>155</sup> Zu dieser Gruppe würde wohl auch das Nomen *Bevölkerung* zählen, welches keinem der vier Wortbildungsmuster von Motsch zugeordnet werden kann. Ich bin in dieser Studie so verfahren, dass alles, was nicht sicher zugeordnet werden kann, von der Zählung ausgeschlossen wird.

Ein spezielles Problem stellt sich in Form der fehlerhaften Lerneräußerungen. Diese gilt es zunächst zu identifizieren<sup>156</sup> und dann ist zu überlegen, ob und wie die falsch gebildeten und/oder falsch gebrauchten *-ung*-Derivate den Wortbildungsmustern zugeordnet werden können. Wie soll man ein von einem Lerner gebildetes Derivat bewerten, das eigentlich gar nicht vorkommen dürfte, weil es als ungrammatisch oder unakzeptabel einzustufen ist? Und was ist mit im Prinzip korrekt gebildeten Derivaten, die aber mit einer offensichtlich falschen Bedeutung verwendet werden?

Als Beispiel für den ersteren Fall möge hier das Derivat *Treffungen* dienen aus dem Lerneratz *Obligatorische Praktika gibt es fast nicht an den französischen Universitäten, Treffungen zwischen*

---

<sup>154</sup> Vgl. Wellmann (1975: 422), der die nomina acti als „sekundäre Prägungen“ bezeichnet und dies wie folgt definiert: „Unter diesem Begriff werden die Funktionserweiterungen von primär als gramm. Abstrakta gebrauchten Ableitungen zusammengefaßt, hier also diejenigen Bildungen, die nicht nur als Prädikatsbez. fungieren, sondern auch dazu dienen, etwas als Obj. einer 'actio' zu kennzeichnen“.

<sup>155</sup> Vgl. Wellmann 1975: 94

<sup>156</sup> Da Wortbildungsfehler in Falko bisher noch nicht annotiert sind, bleibt nur die Möglichkeit einer „manuellen“ Fehleridentifikation, d.h. alle *-ung*-Derivate müssen in ihrem Kontext betrachtet werden, um vor allem den falschen Gebrauch aufzudecken, also korrekt gebildete Wörter, die in einer falschen Bedeutung verwendet werden. Dazu genügt in der Regel der Kontext von 20 Wörtern links und rechts neben dem zu untersuchenden Wort. Dieser Kontext kann in der Darstellung in Falko mitangezeigt werden. In seltenen Fällen erfolgt der Rückgriff auf den kompletten Lernertext, der ebenfalls in Falko abrufbar ist.

den Studenten und den Arbeitgebern während der Studiumperiode auch nicht, usw. (Essays L2, 1.0, fk001\_2006\_07). Dieses regelhaft gebildete nomen actionis wird von Muttersprachlern so nicht verwendet, weil es von dem usuellen Lexem *Treffen* blockiert wird. *Treffen* wäre in diesem Fall die Zielhypothese für die Korrektur des Fehlers.

Ein Beispiel für den Fall eines existierenden *-ung*-Derivates in falscher Bedeutung stellt der Gebrauch von *Rechnung* in dem Satz *Heute spielt den gesellschaftlichen Beitrag eines Menschen kaum eine Rolle bei der Rechnung eines Lohns* (Essays L2, 1.0, fk014\_2006\_07) dar. Hier hätte der Lerner, so meine Zielhypothese, das nomen actionis *Berechnung* verwenden müssen.

In beiden Fällen habe ich das jeweilige nomen actionis als Beleg in der entsprechenden Kategorie gezählt und zwar aufgrund der Tatsache, dass es eine Zielhypothese gibt, die eine plausible Interpretation der Lerneräußerung ermöglicht. Zudem besteht meines Erachtens die Gefahr, dass die Berechnung der potential productivity grob verfälscht wird, wenn man alle von den Lernern fehlerhaft produzierten *-ung*-Derivate nicht mitzählt. Es gibt jedoch vier *-ung*-Derivate, die ich von der Zählung ausgeschlossen habe, weil ich in diesen Fällen keine Zielhypothese für die Fehler in den Lerneräußerungen aufstellen konnte. Zu diesen Fällen gehört der Satz *Weil die Schüler oder die Studenten immer erst Mal Lehrersbenehmung achten* (Essays L2, 1.0, usb012\_2006\_10). Denkbar wäre eine Zielhypothese, die in etwa so lauten könnte: *Weil die Schüler oder die Studenten immer zuerst auf das Benehmen des Lehrers achten*. So ganz gesichert scheint mir diese Zielhypothese aber auch bei Betrachtung des ganzen Kontextes nicht und deshalb wurde dieser Beleg nicht in die Wertung aufgenommen.

Eine weitere Entscheidung, die im Zusammenhang mit der Kategorisierung der *-ung*-Derivate gefällt werden musste, betrifft die Behandlung von *-ung*-Derivaten, die als Kompositionsglied in einem Kompositum dienen. Ich bin hier so vorgegangen, dass ich all jene *-ung*-Derivate gezählt habe, die ich aufgrund einer plausiblen Interpretation einem Wortbildungsmuster zuordnen konnte. Zum Beispiel habe ich das Kompositum *Geldverschwendung* im Lernertext fk005\_2006\_07 als *Verschwendung von Geld* interpretiert und diese „Tatsache, dass jemand Geld verschwendet“, als nomen actionis gezählt. Genauso bin ich mit Fällen verfahren, wo das *-ung*-Derivat das Kompositionserstglied stellt: *Übersetzungsseminare* in fk015\_2006\_08 verstehe ich als *Seminare, in denen es um Übersetzung (von einer Sprache in eine andere) geht*, also um den „Vorgang des Übersetzens“, so dass hier ein Type dem Wortbildungsmuster nomen actionis zugeordnet wurde.

Wenn ein *-ung*-Derivat in zwei verschiedenen Komposita, aber mit der jeweils gleichen Bedeutung vorkam, so wurden zwei Belege für das *-ung*-Derivat gezählt und nicht jeweils ein Beleg für jedes Kompositum. Ein Beispiel: *Geldverschwendung* (Essays L2, 1.0, fk005\_2006\_07) und *Zeitverschwendung* (Essays L2, 1.0, fk006\_2006\_08) wurden nicht als Hapax Legomena der

jeweiligen Komposita gezählt, sondern als zwei Belege für das nomen actionis *Verschwendung*.

Die Grundlage für eine Zuordnung und Zählung der Belege bildet also immer die Tatsache, dass ein komplexes Wort – nach meinem Dafürhalten – eindeutig interpretiert werden kann, bzw. bei fehlerhaften Bildungen und Verwendungen eine plausible Zielhypothese aufgestellt werden kann.<sup>157</sup> Wo dies nicht der Fall ist, wurden die Belege von der Zählung ausgenommen.

### 3.1.2. Auswertung der Belege und Berechnung der potential productivity

Nach einer gründlichen Analyse der Belege und einer Zuordnung zu den vier Wortbildungsmustern der *-ung*-Suffigierung ergibt sich die folgende Verteilung:

L2-Daten („Essays L2, Version 1.0“ mit einer Größe von 65387 Tokens und 7720 Types)

Wortbildungs- muster	Types	Tokens	Hapax Legomena	potential productivity P
1) Nomina actionis	88	179	63	0,352
2) Resultierender Zustand	4	4	4	---
3) Nomina acti	85	659	36	0,055
4) Nomina loci	6	33	1	---

**Abb.5:** Auszählung der Belege für die *-ung*-Suffigierung und Berechnung der potential productivity bei den Lernern

Die Gesamtzahl der L2-Belege aller *-ung*-Suffigierungen beträgt 941 Tokens verteilt auf 191 Types. Aufgrund von Problemen, wie sie im vorangegangenen Abschnitt erläutert wurden, konnten diese neun Types nicht zugeordnet und gezählt werden: *Bewegung* (im Sinne einer politischen Denkrichtung, 33 Tokens), *Stellung* (15 Tokens), *Bevölkerung* (7 Tokens), *Richtung* (6 Tokens), *Benehmung* (2 Tokens), *Schreibung* (1 Token), *Abbestellung* (1 Token), *Strömung* (1 Token).

L1-Daten („Essays L1, Version 0.5“ mit einer Größe von 33806 Tokens und 5831 Types)

Wortbildungs- muster	Types	Tokens	Hapax Legomena	potential productivity P
1) Nomina	64	98	48	0,490

<sup>157</sup> Das Problem, dass hier ein gewisses Maß an Subjektivität mit im Spiel ist, lässt sich nur benennen, nicht umgehen. Lüdeling (2008) weist speziell auf die Schwierigkeit hin, dass für einen Lernerfehler meist verschiedene Zielhypothesen möglich sind.

actionis				
2) Resultierender Zustand	5	9	0	---
3) Nomina acti	73	269	35	0,130
4) Nomina loci	4	19	1	---

**Abb.6:** Auszählung der Belege für die *-ung*-Suffigierung und Berechnung der potential productivity bei den Muttersprachlern

Hier gibt es insgesamt 418 Tokens verteilt auf 151 Types. Folgende Types konnten nicht zugeordnet werden: *Bewegung* (im Sinne einer politischen Denkrichtung, 9 Tokens), *Stellung* (6 Tokens), *Bevölkerung* (5 Tokens), *Richtung* (2 Tokens), *Sitzung* (1 Token).

#### Zur Produktivitätsberechnung:

Die Abbildungen 5 und 6 zeigen die Ergebnisse der quantitativen Analyse in Zahlen. Eine Berechnung des Produktivitätsmaßes **P** der potential productivity wurde nur für die Wortbildungsmuster 1) und 3) vorgenommen, da für die Muster 2) und 4) zuwenig Belege vorhanden waren, um diese als Berechnungsgrundlage zu nehmen.<sup>158</sup> Berechnet wurde **P** mit der Formel von Baayen (2009), die in Abschnitt 2.1.8. bereits näher erläutert wurde:  $P = V(1,C,N) / N(C)$ . Der höchstmögliche Wert, den **P** erreichen kann, ist „1“, was dann der Fall wäre, wenn alle Belege im Korpus Hapax Legomena wären. Demnach gilt: je größer der Wert von P, desto größer die potential productivity der untersuchten Kategorie.

### 3.1.3. Diskussion der Ergebnisse der quantitativen Analyse

Die Berechnung der potential productivity in Abschnitt 3.1.2. liefert ein klares Ergebnis. Die Werte von **P** sind bei den Muttersprachlern signifikant höher als bei den fortgeschrittenen DaF-Lernern: beim Wortbildungsmuster 1) steht der Wert **P** = 0,490 bei den Muttersprachlern einem Wert **P** = 0,352 bei den DaF-Lernern gegenüber, beim Wortbildungsmuster 3) ist die potential productivity bei den Muttersprachlern mit **P** = 0,130 mehr als doppelt so hoch wie bei den DaF-Lernern mit einem Wert **P** = 0,055. Das bedeutet, dass beide Wortbildungsmuster – sowohl die *-ung*-Suffigierung zur Bildung von nomina actionis als auch die zur Bildung von nomina acti – von den Muttersprachlern sehr viel produktiver gebraucht werden als von den DaF-Lernern. Dies überrascht nicht, wenn man in Betracht zieht mit welchen Schwierigkeiten der Erwerb von morphologischen Wortbildungsregularitäten verbunden ist, angefangen bei einer nicht allzu

<sup>158</sup> „Problematisch für die Produktivität i.e.S. sind insbesondere sehr hohe und sehr niedrige Tokenfrequenzen, da die Berechnung von P in diesen Fällen keine zuverlässigen Ergebnisse liefert“. (Scherer 2005: 78)

umfangreichen Behandlung des Themas in DaF-Lehrwerken (vgl. die Einleitung zu dieser Arbeit), über die Problematik einer Dualität von Wortbildungsmustern und gespeicherten Lexemen bei der Aktualisierung von komplexen Wörtern (vgl. Abschnitt 2.1.5.1.), bis hin zur Unschärfe, bzw. noch nicht ausreichenden Formulierung der Regeln und Beschränkungen zur Bildung von *-ung*-Suffigierungen (vgl. Abschnitt 2.1.7.1.).

Demgegenüber scheint auch die Tatsache nicht ins Gewicht zu fallen, dass Lerner komplexe Neubildungen produzieren, die ungrammatisch sind, und die also von Muttersprachlern aller Wahrscheinlichkeit nach nicht produziert würden, was bedeutet, dass ein Lerner theoretisch mehr *-ung*-Suffigierungen bilden könnte. Den meisten der fehlerhaften Lernerprodukte konnte in der Analyse eine Zielhypothese zugewiesen und dadurch gewährleistet werden, dass nur ein sehr geringer Teil der Belege (insgesamt nur drei Hapax Legomena) aus dem Subkorpus „Essays L2, Version 1.0“ von der Zählung ausgenommen wurde. Dies nimmt also keinen signifikanten Einfluss auf die Zählung.

Ein Vergleich der hier errechneten P-Werte mit den Angaben zur Produktivität der Wortbildungsmuster bei Motsch (1999)<sup>159</sup> zeigt tendenziell Übereinstimmung an. Auch wenn konkrete Bezugswerte fehlen, die genauere Angaben darüber machen, wie hoch P sein muss, um das entsprechende Muster als „produktiv“ oder „sehr produktiv“ zu bezeichnen, lässt sich ohne weiteres erkennen, dass die Muster 1) und 3) produktiv sind, wobei 1) weitaus produktiver ist als 3) – und zwar, relativ gesehen, sowohl bei den Muttersprachlern als auch bei den DaF-Lernern. Für die Muster 2) und 4) können, wie gesagt, keine Aussagen gemacht werden, nur soviel: die sehr geringe Typenfrequenz von 4) spricht natürlich für die angenommene Unproduktivität des Musters.

### **3.2. Qualitative Analyse – Untersuchung der Lernerfehler und Entwicklung des Fehlertagsets**

Die qualitative Analyse, der Weg von der Fehleridentifikation zur Entwicklung des Fehlertagsets, gliedert sich in vier Schritte. Im ersten Schritt werden aus der Gesamtmenge der Vorkommen an *-ung*-Derivaten im Falko-Subkorpus „Essays L2, 1.0“ die fehlerhaften Belege herausgefiltert (Abschnitt 3.2.1.). In einem zweiten Schritt wird jedem Fehler eine Zielhypothese zugewiesen und es erfolgt die Beschreibung der Fehler durch einen Vergleich von Fehler und Zielhypothese. Hierzu wird es in Abschnitt 3.2.2. eine komplette tabellarische Auflistung von Fehler, Zielhypothese und Fehlerbeschreibung geben. Anschließend wird aufgrund dieser Fehlerbeschreibungen eine Fehlertaxonomie erarbeitet, welche Kategorien zur Einteilung der Fehler bereitstellt (Abschnitt 3.2.3.). Entsprechend dieser Fehlertaxonomie wird in einem letzten Schritt ein Fehlertagset

---

<sup>159</sup> Vgl. Abschnitt 2.1.7.4.

entwickelt, welches für einen zukünftigen Gebrauch in Falko geeignet sein soll (Abschnitt 3.2.4.). Eine kurze Zusammenfassung der Ergebnisse dieser qualitativen Analyse bietet Abschnitt 3.2.5.

### 3.2.1. Auswahl der Belege

Die Auswahl der Belege für die qualitative Analyse wurde im Prinzip schon mit der Auswahl der Belege für die quantitative Analyse vorweggenommen. Wie in Abschnitt 3.1.1. ausgeführt ist, mussten bereits bei der Zuordnung der *-ung*-Derivate zu den Wortbildungsmustern der *-ung*-Suffigierung eine Identifikation der falsch gebildeten bzw. falsch gebrauchten Derivate<sup>160</sup> sowie eine Interpretation aufgrund einer konkreten Zielhypothese durchgeführt werden. Die Fehleridentifikation verlief dabei „manuell“, d.h. es konnte auf keine bestehende Fehlerannotation in Falko zurückgegriffen werden. Stattdessen musste zu jedem *-ung*-Derivat der unmittelbare Kontext in die Betrachtung miteinbezogen werden (im Übrigen nicht nur für die Fehleridentifikation, in vielen Fällen auch für die Kategorisierung der korrekten *-ung*-Derivate). Als Kontext reichten meist die 20 Wörter links und rechts von dem zu untersuchenden Wort aus, sie können in Falko gemeinsam mit dem eigentlichen Beleg angezeigt werden. In wenigen Fällen musste auch der komplette Lernertext, der ebenfalls in Falko abgerufen werden kann, als Kontext herangezogen werden. Es liegt auf der Hand, dass diese Art der Fehleridentifikation sowie die Zuweisung einer Zielhypothese ein gewisses Maß an Subjektivität beinhalten und bei verschiedenen Forschern zu mehr oder weniger unterschiedlichen Ergebnissen führen können.<sup>161</sup> Eine vollständige Auflistung der Fehler und der zugewiesenen Zielthesen liefert der nächste Abschnitt.

Es seien hier aber noch einige weitere wichtige Anmerkungen in Bezug auf die Fehleridentifikation gemacht. Da die vorliegende Studie den Gebrauch der *-ung*-Suffigierung in der Lernerproduktion untersucht, zielt die Fehleridentifikation darauf ab, nur Fehler zu behandeln, die in direktem Zusammenhang stehen mit den Wortbildungsregeln der *-ung*-Suffigierung, wie sie in den Abschnitten 2.1.3., 2.1.5., und 2.1.7. Erwähnung finden. Dazu gehören beispielsweise die Bildung blockierter Formen, die Bildung von *-ung*-Derivaten, die nicht wohlgeformt sind, oder der semantisch falsche Gebrauch von *-ung*-Derivaten. Diese Identifikation mag trivial erscheinen, sie bietet aber etliche Schwierigkeiten, die im Rahmen der Fehlerbeschreibung deutlich werden.<sup>162</sup> Weitere problematische Fälle sind der Umgang mit vermeintlichen Rechtschreibfehlern oder Genusfehlern.

---

<sup>160</sup> Beispiele für den Unterschied zwischen falsch gebildeten Derivaten und Derivaten, die im Prinzip korrekt gebildet sind, aber mit einer falschen Bedeutung verwendet werden, finden sich in Abschnitt 3.1.1.

<sup>161</sup> Zu dieser Problematik vgl. Abschnitt 2.4.2 und Lüdelling (2007).

<sup>162</sup> Vgl. auch die Diskussion der Ergebnisse in Abschnitt 3.2.5.

Bloße Rechtschreibfehler in den *-ung*-Derivaten sollten eigentlich keine explizite Beachtung finden. Allerdings ist es nicht immer so einfach zu bestimmen, was „nur“ ein Rechtschreibfehler ist. Das Wort *Kwalifisierung* kann im Kontext des Lernertextes sa002\_2006\_09 eindeutig als eine falsch geschriebene Form von *Qualifizierung* identifiziert werden und somit liegt hier kein Fehler vor, der direkt etwas mit der Wortbildung zu tun hat. Wie soll aber beispielsweise mit der Groß- und Kleinschreibung verfahren werden? Zu den Regeln der *-ung*-Suffigierung gehört ja gerade, dass mit diesem Wortbildungsprozess aus Verben Nomen gebildet werden, also müsste die Kleinschreibung von *-ung*-Derivaten eigentlich als Fehler behandelt werden. Ich habe mich jedoch entschieden, darauf zu verzichten, wenn das betreffende Wort als Nomen getaggt ist und auch für mich aus dem Kontext klar ersichtlich ist, dass der Lerner es wie ein Nomen verwendet. Meist kommt in diesen Lernertexten hinzu, dass generell Fehler bei der Groß- und Kleinschreibung gemacht werden und dies nicht im Zusammenhang mit der Wortbildung zu sehen ist. Ebenfalls unklar ist, wie Fälle wie der folgende zu behandeln sind: im Lernertext sa005\_2006\_09 kommt das Wort *Trainung* vor. Hier sind mindestens zwei Versionen denkbar. Entweder handelt es sich um einen Wortbildungsfehler, weil der Lerner für die Nomenbildung einen Verbstamm *train-* angenommen hat, den es im Deutschen so nicht gibt, oder es handelt sich um einen Flüchtigkeitsfehler, weil der Lerner die korrekte Schreibung *Training* kennt, aber sich hier „verschrieben“ hat.<sup>163</sup> Hier kommt das bereits angesprochene Problem der Unterscheidung von „errors“ und „lapses“ ins Spiel, eine Unterscheidung, die an der Oberfläche eigentlich nicht zu treffen ist<sup>164</sup> – es sei denn, es gibt im Lernertext einen sehr plausiblen Hinweis, der für die eine oder die andere Interpretation spricht. Dies ist in Lernertext sa005\_2006\_09 der Fall: in demselben Satz, in dem das falsch geschriebene Wort *Trainung* vorkommt, findet sich auch die korrekte Form *Training*. Aus diesem Grund würde ich den Beleg *Trainung* nicht als Wortbildungsfehler, sondern als Schreib- bzw. Tippfehler (unabhängig ob von Seiten des Lerners oder der Texteingabe ins Korpus) interpretieren. In vergleichbaren Fällen, in denen der Kontext weniger deutliche „Interpretationshilfen“ gibt, muss das falsch geschriebene Wort als Wortbildungsfehler interpretiert werden. In der Fehlerbeschreibung im folgenden Abschnitt wird – wo dies der Fall ist – allerdings in Klammern auf die Möglichkeit, dass auch ein einfacher Rechtschreibfehler plausibel erscheint, verwiesen.

Ein ähnliches Problem ist zum Beispiel die Auslassung von Buchstaben wie das Fehlen eines *l* in *Vorstelung* (Essays L2, 1.0, fk018\_2996\_08). Während hier noch von einem Rechtschreibfehler, der nichts mit der Wortbildung als solcher zu tun hat, ausgegangen werden kann, muss das Fehlen einer *-s*-Fuge als Verstoß gegen die Regeln der Kompositumbildung mit *-ung*-Derivaten gewertet werden.

---

<sup>163</sup> Denkbar wäre außerdem noch, dass der handschriftliche Lernertext an der Stelle nicht eindeutig zu entziffern war oder sogar ein Tippfehler bei der Texteingabe ins Korpus gemacht wurde.

<sup>164</sup> Vgl. Abschnitt 2.4.1.

Zu diskutieren ist auch der Umgang mit Genusfehlern. Die *-ung*-Suffigierung bildet regelhaft und ausnahmslos Nomen mit dem Genus Femininum, die Verwendung eines *-ung*-Derivates mit einem Artikel im Maskulinum oder Neutrum müsste demnach als Genusfehler im Zusammenhang mit der Wortbildung betrachtet werden. Da aber das Flexionsparadigma der Artikel eine Kombination der Kategorien Genus, Kasus und Numerus darstellt und generell für Lerner große Schwierigkeiten birgt<sup>165</sup>, ist oft nicht eindeutig zu klären, ob der falsche Gebrauch eines Artikels einen Genus-, Kasus- oder Numerusfehler bedeutet. Noch schwieriger ist die Lage bei der Adjektivdeklination. Aus diesem Grund wird hier auf den Versuch einer Identifikation von Genusfehlern verzichtet; außerdem kann ich hinzufügen, dass mir bei einer Durchsicht der *-ung*-Derivate in den Lernertexten nicht ein einziger Fehler beim Gebrauch der Artikel aufgefallen ist.

### 3.2.2. Fehler, Zielhypothese und Fehlerbeschreibung

Die folgende Auflistung beinhaltet alle fehlerhaften *-ung*-Derivate<sup>166</sup>, die im Falko-Subkorpus (Essays L2, 1.0) gefunden wurden, sowie eine Zielhypothese und eine Fehlerbeschreibung. Bei einigen Belegen war es nicht möglich zu einer klaren Zielhypothese zu gelangen, dies ist dann mit „(!?)“ kommentiert und wird auch in der Fehlerbeschreibung diskutiert. Es gibt verschiedene Gründe hierfür, unter anderem das Problem, dass man manchmal eine Lerneräußerung als eindeutig fehlerhaft oder unverständlich identifizieren kann, jedoch selbst bei Betrachtung des kompletten Lernertextes nicht eindeutig zu erschließen vermag, was der Lerner eigentlich ausdrücken wollte. Ebenso gibt es den Fall, dass man eine Lerneräußerung als Verstoß gegen die Gebrauchsnorm empfindet, d.h. man hat das subjektive Gefühl, dass ein Muttersprachler die betreffende Wortform oder Formulierung niemals verwenden würde, aber man kann nicht genauer erklären, worin der Fehler eigentlich liegt, also welche Regeln oder Normen verletzt wurden.

Bei der qualitativen Analyse wurde auf eine getrennte Untersuchung der einzelnen Wortbildungsmuster verzichtet, die Fehler sind hier zusammengefasst als Fehler bei der Bildung und dem Gebrauch von *-ung*-Derivaten.

1.

Lernertext	cbs002_2006_09
Fehler im Kontext	In dieser <b>Verbindung</b> soll erwähnt werden (...)
Zielhypothese	In diesem <b>Zusammenhang</b> (...)
Fehlerbeschreibung	Falscher Gebrauch: obwohl die Lerneräußerung mit dem Wort <i>Verbindung</i>

<sup>165</sup> Vgl. Diehl 2000

<sup>166</sup> Mit Ausnahme der „reinen Rechtschreibfehler“, sofern ein Fehler als solcher identifiziert werden konnte.

	wohl ohne Probleme zu verstehen ist, gibt die zielsprachliche Gebrauchsnorm die Verwendung anderer Lexeme wie z.B. <i>Zusammenhang</i> vor.
--	---

2.

Lernertext	cbs005_2006_09
Fehler im Kontext	Praktikas und Studienjobs sind gut. Nicht nur für das Lebenslauf, sondern auch für das Verständnis und die <b>Motivierung</b> des Studiums.
Zielhypothese	(...) und die <b>Motivation</b> für das Studium.
Fehlerbeschreibung	Falsche Semantik: Motivierung bedeutet (inhaltliche) „Begründung“, bei der Motivation handelt es sich um die „Beweggründe, die das Handeln eines Menschen bestimmen“ <sup>167</sup>

3.

Lernertext	cbs007_2006_09
Fehler im Kontext	Zunächst muss man dann überlegen, was man nach dem Universitätsabschluss genau will? In einem Unternehmen arbeiten? Selbständig sein? oder vielleicht sogar selbst an einer Universität Unterrichten? Die <b>Profilierung</b> muss hier natürlich auch berücksichtigt werden.
Zielhypothese	(...) Die <b>Eignung</b> muss hier natürlich auch berücksichtigt werden.
Fehlerbeschreibung	Falsche Semantik: Profilierung bedeutet „Entwicklung der Fähigkeiten“ <sup>168</sup> und kommt nur als nomen actionis vor, nicht als nomen acti wie hier gebraucht.

4.

Lernertext	cbs008_2006_09
Fehler im Kontext	In Dänemark praktiziert man es oft ökonomische <b>Verbrechungen</b> härter zu bestrafen als die körperlichen (...)
Zielhypothese	(...) ökonomische <b>Verbrechen</b> (...)
Fehlerbeschreibung	Blockierung des <i>-ung</i> -Derivates durch usuelles Synonym

5.

Lernertext	cbs012_2006_09
Fehler im Kontext	(...) wollte ich eigentlich versuchen eine Arbeit zu finden. Als ich die <b>Stellungsanzeige</b> gelesen habe, (...)
Zielhypothese	(...) die <b>Stellenanzeige</b> (...)
Fehlerbeschreibung	Falsche Semantik von <i>Stellung</i> , kann nicht „Arbeitsstelle“ bedeuten

6.

<sup>167</sup> Duden Bd. 1: 667

<sup>168</sup> Duden Bd. 5: 808

Lernertext	cbs013_2006_09
Fehler im Kontext	(...) obwohl die Arbeit sonst schön und mit vielen <b>Überlegungen</b> geschrieben ist.
Zielhypothese	(...) mit viel <b>Überlegung</b> geschrieben ist.
Fehlerbeschreibung	Plural ist falsch: <i>mit Überlegung</i> in der Bedeutung von „mit Bedacht“ bzw. „mit Nachdenken“ ist ein nomen actionis und kommt nur im Singular vor.

7.

Lernertext	fk001_2006_08
Fehler im Kontext	(...) obligatorische Praktika gibt es fast nicht an den französischen Universitäten, <b>Treffungen</b> zwischen den Studenten und den Arbeitgebern während der Studiumperiode auch nicht, usw.
Zielhypothese	(...) <b>Treffen</b> zwischen (...)
Fehlerbeschreibung	Blockierung des <i>-ung</i> -Derivates durch ein usuelles Synonym

8.

Lernertext	fk004_2006_07
Fehler im Kontext	Die <b>Entwicklung</b> der Frauen ist durch die Ausbildung möglich gewesen.
Zielhypothese	Die <b>Entfaltung</b> der Frauen (...) (!?)
Fehlerbeschreibung	Hier ist es schwer zu sagen, ob mit <i>Entwicklung</i> wirklich ein falscher Gebrauch des Wortes vorliegt. In meiner subjektiven Einschätzung klingt <i>Entfaltung</i> sehr viel besser, aber eventuell handelt es sich hier eher um eine Stilfrage, als um einen echten Fehler.

9.

Lernertext	fk006_2006_07
Fehler im Kontext	(...) ob alle ihre eigen <b>Verantwortungen</b> verstehen und übernehmen können.
Zielhypothese	(...) ihre eigene <b>Verantwortung</b> (...)
Fehlerbeschreibung	Keine Pluralbildung möglich. Allerdings scheint dies nicht am <i>-ung</i> -Derivat selbst zu liegen, welches hier ein nomen acti ist und von dem prinzipiell ein Plural möglich ist. Es ist wohl eher die Kollokation <i>Verantwortung übernehmen</i> , in der nur der Gebrauch des Singulars der Norm entspricht.

10.

Lernertext	fk007_2006_07
Fehler im Kontext	In der Gesellschaft und besonders in <b>Werbungen</b> , im Fernsehen oder im Zeitschriften werden die Frauen als Objekte dargestellt.
Zielhypothese	(...) in der <b>Werbung</b> , im Fernsehen (...)
Fehlerbeschreibung	Keine Pluralbildung möglich. <i>Werbung</i> in der Bedeutung „Gesamtheit aller werbenden Maßnahmen“ kommt nur im Singular vor. <sup>169</sup>

11.

Lernertext	fk007_2006_08
Fehler im Kontext	Kriminell ist nicht nur eine <b>Ermordung</b> , sondern auch einen Diebstahl oder die Ausrottung eines ganzen Volks.
Zielhypothese	(...) eine <b>Ermordung</b> , (...)
Fehlerbeschreibung	Der Wortbildung wurde ein nicht existenter Verbstamm (* <i>ermorder-</i> ) zugrunde gelegt.

12.

Lernertext	fk009_2006_07
Fehler im Kontext	Die Probleme liegen heute nicht so viel in Gesetzen und <b>Kvotierungen</b> , sondern in den Erwartungen, die die Gesellschaft auf Frauen und Männer hat.
Zielhypothese	(...) in Gesetzen und <b>Quoten</b> , sondern (...) (!?)
Fehlerbeschreibung	Es ist unklar, welche Äußerung genau vom Lerner intendiert war. Denkbar wäre vielleicht etwas wie <i>Quote</i> oder <i>Frauenquote</i> , da es in dem Essay um Feminismus und die Gleichberechtigung der Frau geht. <i>Quotierung</i> jedenfalls ist ein Fachwort aus der Wirtschaft und bedeutet „das Angeben eines Preises oder Kurses an der Börse“ <sup>170</sup> und ist in dem vorliegenden Kontext eindeutig falsch gebraucht.

13.

Lernertext	fk009_2006_07
Fehler im Kontext	Die meisten Menschen wollen ja keine Veränderungen, und sie wollen nicht, dass ihre <b>Lebenshaltung</b> in Frage gestellt wird.
Zielhypothese	(...) dass ihre <b>(Lebens-)Einstellung</b> in Frage (...) (!?)
Fehlerbeschreibung	Falsche Semantik: <i>Lebenshaltung</i> als „Gestaltung des Lebens in bezug auf die privaten wirtschaftlichen Verhältnisse“ <sup>171</sup> passt nicht in den Kontext des

<sup>169</sup> Vgl. Duden Bd. 10: 750

<sup>170</sup> Vgl. Duden Bd. 5: 836

	Textes, der auf Meinungen und Einstellungen der Menschen abhebt, weniger auf ökonomische Umstände. Eine eindeutige Zielhypothese fällt allerdings schwer, vielleicht wäre die Lerneräußerung am ehesten durch ( <i>Lebens-</i> ) <i>Einstellung</i> zu ersetzen.
--	--

14.

Lernertext	fk009_2006_07
Fehler im Kontext	(...) dass Frauen nicht mehr in <b>Werbungen</b> und Tabloiden als seelenlose Puppen für die Blicke der Männer dargestellt werden.
Zielhypothese	(...) nicht mehr <b>in der Werbung</b> und in (...)
Fehlerbeschreibung	Keine Pluralbildung möglich. <i>Werbung</i> in der Bedeutung „Gesamtheit aller werbenden Maßnahmen“ kommt nur im Singular vor. (Vgl. Nr. 10)

15.

Lernertext	fk011_2006_07
Fehler im Kontext	Der Unterschied von Perspektiven scheint in der <b>Schätzung</b> des Verhältnisses zwischen dem Wert des Menschen und seiner ökonomischer Leistung zu bestehen.
Zielhypothese	(...) in der <b>Einschätzung</b> des Verhältnisses (...)
Fehlerbeschreibung	Falsche Semantik: der Wert einer Sache kann <i>geschätzt</i> werden, ein Verhältnis, eine Relation wird <i>eingeschätzt</i> .

16.

Lernertext	fk012_2006_07
Fehler im Kontext	Die Bewegung des Feminismus und seine <b>Nutzung</b> in unserer Gesellschaft sind aber nicht zu vergessen.
Zielhypothese	(...) und sein <b>Nutzen</b> in unserer Gesellschaft (...).
Fehlerbeschreibung	Falsche Semantik: Es scheint plausibel, das hier das nomen acti <i>Nutzen</i> vom Lerner intendiert war und nicht das nomen actionis <i>Nutzung</i> . Der <i>Nutzen des Feminismus</i> bedeutet, dass der Feminismus jemandem nutzt, die <i>Nutzung des Feminismus</i> , dass der Feminismus genutzt wird, um etwas zu tun.

17.

Lernertext	fk012_2006_07
Fehler im Kontext	Dieses Geschmack zum Skandal und zur Provozierung scheint ein Kennzeichen des Feminismus bis heute sein, (...).
Zielhypothese	Dieser Geschmack (besser: diese Neigung) zum Skandal und zur <b>Provokation</b> scheint (...). (!?)
Fehlerbeschreibung	Diese Zielhypothese gibt in erster Linie das subjektive Gefühl wieder, dass <i>Provokation</i> in diesem Kontext besser ist. Der Gebrauch des nomen actionis <i>Provozierung</i> scheint mir inkorrekt zu sein, zumindest so lange ein Genitivattribut wie zum Beispiel in <i>die Provozierung einer Auseinander-</i>

<sup>171</sup> Vgl. Duden Bd. 10: 410

	<i>setzung</i> fehlt. Ich konnte diese Vermutung allerdings nicht in der Literatur zur Wortbildung bestätigt finden.
--	--

18.

Lernertext	fk013_2006_07
Fehler im Kontext	Ich denke, zum Beispiel, an die Millionen und Millionen von Leuten, die in den großen Firmas (...) weiter arbeiteten, an ihre winzige <b>Lohnungen</b> , (...)
Zielhypothese	(...) an ihre extrem niedrigen <b>Löhne</b> , (...)
Fehlerbeschreibung	Blockierung des <i>-ung</i> -Derivates durch usuelles Synonym; zudem ist nicht klar, wie durch die regelhafte Nominalisierung des Verbs <i>lohn</i> ein Substantiv mit der Bedeutung „Lohn“ gebildet werden soll

19.

Lernertext	fk014_2006_07
Fehler im Kontext	Heute spielt den gesellschaftlichen Beitrag eines Menschen kaum eine Rolle bei der <b>Rechnung</b> eines Lohns.
Zielhypothese	(...) bei der <b>Berechnung</b> eines Lohns.
Fehlerbeschreibung	Falsche Semantik: Von dem intransitiven Verbs <i>rechnen</i> kann kein nomen actionis gebildet werden, die <i>-ung</i> -Suffigierung liefert hier nur das idiosynkratische Lexem <i>Rechnung</i> . Im vorliegenden Kontext kann nur das nomen actionis des transitiven Verbs <i>berechnen</i> gebraucht werden.

20.

Lernertext	fk016_2006_07
Fehler im Kontext	Ich glaube, dass ohne eine starke (...) <b>Stellungnahme</b> der Frauen
Zielhypothese	(...) <b>Stellungnahme</b> der Frauen
Fehlerbeschreibung	<i>Stellungnahme</i> weicht durch das Fehlen der <i>-s</i> -Fuge von der regelhaften Bildung von Komposita mit <i>-ung</i> -Derivaten ab. <sup>172</sup>

21.

Lernertext	fk016_2006_08
Fehler im Kontext	Die Studenten beschaeftigen sich nicht mit der <b>Fassung</b> von Vertragen oder von Akten.
Zielhypothese	(...) mit der <b>Abfassung</b> von Verträgen (...)
Fehlerbeschreibung	Falsche Semantik: das Schreiben von Schriftstücken kann nur mit den präfigierten Verben <i>abfassen</i> oder <i>verfassen</i> , nicht mit dem Simplex <i>fassen</i> ausgedrückt werden.

22.

<sup>172</sup> Vgl. Abschnitt 2.1.7.1

Lernertext	fk017_2006_07
Fehler im Kontext	Sie haben die Wahl eine Schwangerschaft <b>abbrechung</b> (...)
Zielhypothese	(...) Schwangerschafts <b>abbruch</b> (...)
Fehlerbeschreibung	Blockierung des <i>-ung</i> -Derivates durch ein usuelles Synonym.

23.

Lernertext	fk017_2006_07
Fehler im Kontext	(...) manchmal bekommen Frauen nicht eine <b>Leitungsstelle</b> , nur die niedriger Arbeitsstelle.
Zielhypothese	(...) nicht eine <b>leitende Position</b> , nur (...)
Fehlerbeschreibung	Falscher Gebrauch: obwohl die Lerneräußerung mit dem Wort <i>Leitungsstelle</i> wohl ohne Probleme zu verstehen ist, gibt die zielsprachliche Gebrauchsnorm die Verwendung anderer Lexeme wie z.B. <i>leitende Position</i> vor.

24.

Lernertext	fk018_2006_08
Fehler im Kontext	Meiner Meinung nach lohnt sich auf jeden Fall die <b>Bemühung</b> .
Zielhypothese	(...) lohnt sich die <b>Mühe</b> auf jeden Fall.
Fehlerbeschreibung	Falscher Gebrauch: <i>Die Mühe lohnt sich</i> ist eine feste Kollokation, in der die Gebrauchsnorm <i>Mühe</i> vorgibt und das sinnverwandte <i>Bemühung</i> ausschließt.

25.

Lernertext	fk022_2006_08
Fehler im Kontext	(...) sollte ich mich viel mehr mit Bauhandlungen, <b>Vermassung</b> , Bauökonomie und Präsentationen von Projekten beschäftigen.
Zielhypothese	(...) Bauhandlungen, <b>Vermessung</b> , Bauökonomie (...)
Fehlerbeschreibung	Falsche Semantik: das Verb <i>vermassen</i> existiert in der Bedeutung „etwas zur Massenware zu machen“, was in dem gegebenen Kontext nicht vom Lerner intendiert gewesen sein kann (plausibel wäre allerdings auch, einfach einen Rechtschreibfehler anzunehmen).

26.

Lernertext	hu004_2006_09
Fehler im Kontext	Weg mit Trustfonds--jeder muss arbeiten, und ihre Entlohnung sollten dem Beitrag entsprechen, den er für die Gesellschaft geleistet hat. Mit dieser <b>Stellung</b> bin ich nicht ganz einverstanden, (...)
Zielhypothese	(...) Mit dieser <b>Position</b> bin ich (...)
Fehlerbeschreibung	Falsche Semantik: <i>Stellung</i> kann nicht die Position in einem Diskurs bezeichnen.

27.

Lernertext	hu006_2006_09
------------	---------------

Fehler im Kontext	(...) neues Auto, neues Haus und in einem kapitalistischen System gibt es kein Platz an sozialen <b>Verantwortungen</b> nachzudenken.
Zielhypothese	(...) über soziale <b>Verantwortung</b> nachzudenken. (!?)
Fehlerbeschreibung	Der Plural scheint mir in diesem Kontext falsch gebraucht zu sein. Allerdings kann ich nicht begründen, warum, und es handelt sich hier womöglich ebenfalls um eine Stilfrage und nicht um einen wirklichen Fehler (vgl. Nr. 8).

28.

Lernertext	hu008_2006_09
Fehler im Kontext	(...) wenn wir diese berühmten Menschen nicht so sehr wie Götter behandeln würden dann wären sie auch nicht reich weil sie dann nicht für <b>Bewerbung</b> benutzt werden.
Zielhypothese	(...) dann nicht für die <b>Werbung</b> (...)
Fehlerbeschreibung	Falscher Gebrauch. „Die Gesamtheit aller werbenden Maßnahmen“ wird durch das Lexem <i>Werbung</i> ausgedrückt (vgl. Nr. 10). Der Gebrauch von <i>Bewerbung</i> wäre nur denkbar mit einer Erweiterung wie in <i>die Bewerbung von Produkten</i> .

29.

Lernertext	hu008_2006_10
Fehler im Kontext	(...) ist es wahrscheinlich, dass das Gefängnis zur Mentale <b>Folterung</b> wird.
Zielhypothese	(...) zu einer mentalen <b>Folter</b> wird.
Fehlerbeschreibung	<i>Folterung</i> in der Verwendung als nomen acti ist durch <i>Folter</i> blockiert.

30.

Lernertext	hu009_2006_09
Fehler im Kontext	In der DDR Zeit gab es eigentlich nicht so viel Unterschied zwischen Löhne aber trotzdem gab es kein richtiges System basiert auf eine <b>Massung</b> der Leistung fuer die Gesellschaft.
Zielhypothese	(...) auf einer <b>Messung</b> der Leistung (...)
Fehlerbeschreibung	Falsches Verb zugrunde gelegt: das Verb <i>*massen</i> existiert nicht (plausibel wäre allerdings auch, einfach einen Rechtschreibfehler anzunehmen). (Vgl. Nr. 25)

31.

Lernertext	hu011_2006_09
------------	---------------

Fehler im Kontext	(...) sie können heiraten oder nicht und sie können auch schneiden lassen, ohne negative <b>Verfolgungen</b> für sie und ihre Familien.
Zielhypothese	(...) ohne negative <b>Folgen</b> für (...)
Fehlerbeschreibung	Falsche Semantik: <i>Verfolgungen</i> , kann nicht in der Bedeutung von „Konsequenzen“ gebraucht werden.

32.

Lernertext	kne01_2006_07
Fehler im Kontext	Alle wissen die negative <b>Wirkungen</b> , die Kriminalität verursacht, durch Mord und Diebstahl beispielweise. Wegen dieser negativen <b>Wirkungen</b> unterstützt keiner Kriminalität.
Zielhypothese	(...) negative <b>Auswirkungen</b> (...). (...) dieser negativen <b>Auswirkungen</b> (...)
Fehlerbeschreibung	Falsche Semantik: wie bei den Beispielen weiter oben gibt es nur einen kleinen, aber für den Gebrauch wichtigen Bedeutungsunterschied zwischen dem präfigierten Verb und dem Simplex, bzw. den daraus gebildeten <i>-ung</i> -Derivaten.

33.

Lernertext	kne08_2006_09
Fehler im Kontext	In die meisten Universitäten gibt kein gutes System, um sicher zu sein, dass die Studierenden für die <b>Lesungen</b> (Unterrichten) gehen. Man darf nicht weniger als 70 % von den <b>Lesungen</b> abwesend sein. (...) viele Studierende sind für eine kleine Teil die <b>Lesungen</b> anwesend.
Zielhypothese	(...) in die <b>Vorlesungen</b> (...) gehen. (...) von den <b>Vorlesungen</b> abwesend sein. (...) Teil der <b>Vorlesungen</b> anwesend.
Fehlerbeschreibung	Falsche Semantik: <i>Lesungen</i> sind Buchvorstellungen eines Autors, <i>Vorlesungen</i> sind Veranstaltungen der Universitäten.

34.

Lernertext	kne13_2006_07
Fehler im Kontext	Wenn man die Arbeit der Hebamme als ein Beispiel nimmt oder die medizinische <b>Versorgungen</b> generel; (...)
Zielhypothese	(...) die medizinische <b>Versorgung</b> (...)
Fehlerbeschreibung	Keine Pluralbildung möglich, als nomen actionis drückt Versorgung hier die

	„Tatsache, dass jemand versorgt wird“ aus.
--	--

35.

Lernertext	sa001_2006_09
Fehler im Kontext	Deswegen stimme ich überein mit der <b>Stellung</b> dass die meisten Universitätsabschlüsse nicht praxisorientiert sein.
Zielhypothese	(...) mit der <b>Meinung</b> , dass (...)
Fehlerbeschreibung	Falsche Semantik: <i>Stellung</i> kann nicht die Position in einem Diskurs bezeichnen.

36.

Lernertext	sa005_2006_09
Fehler im Kontext	Es gibt zwei <b>Stellungen</b> in dieser Frage; eins - der Universitätsabschluss ist nicht praxisorientiert; zweitens - deshalb sind sie von geringem Wert. Ich stimme ein mit der ersten <b>Stellung</b> , aber gar nicht mit der zweiten <b>Stellung</b> .
Zielhypothese	(...) zwei <b>Positionen</b> in (...) (...) ersten <b>Position</b> , aber (...) zweiten <b>Position</b> .
Fehlerbeschreibung	Falsche Semantik: <i>Stellung</i> kann nicht die Position in einem Diskurs bezeichnen.

37.

Lernertext	sa008_2006_09
Fehler im Kontext	Ich denke diese Haltung <b>im beziehung</b> auf Kriminalität ist total falsch. (...) dass Südafrika sein Gesetz ändern müssen <b>in beziehung</b> auf die Todesstrafe.
Zielhypothese	(...) <b>in Bezug</b> auf Kriminalität (...). (...) müssen <b>in Bezug</b> auf die Todesstrafe.
Fehlerbeschreibung	Obwohl <i>Beziehung</i> und <i>Bezug</i> prinzipiell sinnverwandt sind, handelt es sich in dem Kontext hier um eine feste Kollokation als Diskursmittel. Die Gebrauchsnorm gibt nur <i>in Bezug auf</i> als korrekte Verwendung vor.

38.

Lernertext	sa010_2006_09
Fehler im Kontext	Die gesellschaftliche <b>Neugung</b> zum kriminale Aktivitäten hat ein großer Einfluss (...). Die größte Faktor ist die <b>Neugung</b> des mans zur Mag und Geld und (...). Die Kriminalität ist ein Symptom von einer größer Problem - die von Armut, Krankheit und die <b>Neugung</b> des mans zur Mag und Geld.

Zielhypothese	Die gesellschaftliche <b>Neigung</b> zu (...). Der größte Faktor ist die <b>Neigung</b> des Mannes (...). (...) und der <b>Neigung</b> des Mannes (...).
Fehlerbeschreibung	Der Wortbildung wurde ein nicht existenter Verbstamm <i>*neug-</i> zugrunde gelegt (plausibel wäre allerdings auch, einfach einen Rechtschreibfehler anzunehmen).

39.

Lernertext	usb004_2006_10
Fehler im Kontext	Sie wissen nie, dass ihre <b>Verübungen</b> der Menschlichkeit viel Unglück bringen, (...).
Zielhypothese	(...) dass ihre <b>Verbrechen</b> den Menschen (...).
Fehlerbeschreibung	Wortbildungsfehler: <i>Verüben</i> ist ein Funktionsverb, das aufgrund seines geringen semantischen Gehalts nicht als nomen acti verwendet werden kann. Möglich ist lediglich die Bildung eines nomen actionis mit angeschlossenem Genitivattribut wie in <i>die Verübung von Verbrechen</i> .

40.

Lernertext	usb005_2006_10
Fehler im Kontext	Der erste und schlimmste Nachteil ist die <b>Sinkung</b> der Qualität.
Zielhypothese	(...) Nachteil ist <b>das Sinken</b> der Qualität.
Fehlerbeschreibung	Blockierung des <i>-ung</i> -Derivates durch ein usuelles Synonym.

41.

Lernertext	usb005_2006_10
Fehler im Kontext	Deshalb erscheint die Notwendigkeit für <b>die Kontrollierung</b> der Qualität.
Zielhypothese	(...) Notwendigkeit für <b>eine Kontrolle</b> der Qualität.
Fehlerbeschreibung	Blockierung des <i>-ung</i> -Derivates durch ein usuelles Synonym.

42.

Lernertext	usb12_2006_10
Fehler im Kontext	Im Lauf der Stunde an der Uni müssen die Studenten ihre <b>Benehmung</b> verbessern.
Zielhypothese	(...) die Studenten ihr <b>Benehmen</b> verbessern.

Fehlerbeschreibung	Blockierung des <i>-ung</i> -Derivates durch ein usuelles Synonym.
--------------------	--

43.

Lernertext	usb12_2006_10
Fehler im Kontext	Weil die Schüler oder die Studenten immer erst Mal <b>Lehrersbenehmung</b> achten.
Zielhypothese	(...) erstmal auf das <b>Benehmen des Lehrers</b> achten. (!?)
Fehlerbeschreibung	So wie die Zielhypothese formuliert ist, gilt auch hier: Blockierung des <i>-ung</i> -Derivates durch ein usuelles Synonym. Allerdings ist die Zielhypothese in diesem Fall nicht so eindeutig. Es ist irgendwie fraglich, ob der Lerner wirklich ausdrücken wollte, dass Studenten an der Universität auf das Benehmen des Dozenten achten.

44.

Lernertext	usb013_2006_10
Fehler im Kontext	Nur sie müssen während der Stunde mit guter <b>Achtung</b> sitzen. Wenn der Student alles so macht, wird alles sehr gut.
Zielhypothese	(...) mit guter <b>Aufmerksamkeit</b> sitzen.
Fehlerbeschreibung	Falscher Gebrauch: <i>Achtung</i> und <i>Aufmerksamkeit</i> sind sinnverwandt, der Bedeutungsunterschied ist hier sehr gering, aber die Gebrauchsnorm der Lexeme gibt für diesen Kontext nur die Verwendung von <i>Aufmerksamkeit</i> vor.

45.

Lernertext	usb14_2006_10
Fehler im Kontext	(...) einige Menschen können auf Essen mit <b>Vergnügung</b> verzichten, (...)
Zielhypothese	(...) Essen mit <b>Vergnügen</b> verzichten.
Fehlerbeschreibung	Falsche Semantik: <i>Vergnügen</i> ist das ein Gefühl, sinnverwandt mit <i>Freude</i> ; <i>Vergnügung</i> dagegen beschreibt einen konkreten Anlass für dieses Gefühl, etwa ein Fest oder eine Reise.

Abb. 7: Auflistung der Lerner-Fehler mit Quellenangabe, Zielhypothese und Fehlerbeschreibung

### 3.2.3. Kategorisierung der Fehler

Die Beschreibung der Fehler im vorangegangenen Abschnitt arbeitet ohne vorgegebene Fehlerkategorien, sie stützt sich lediglich auf einen Vergleich von Lerneräußerung und Zielhypothese sowie auf die Regeln und Normen, die eine deskriptive Grammatik für den Bereich der Wortbildung vorgibt, und teilweise – das ist nicht ganz unumgänglich – auf meine subjektive Einschätzung bei der Frage, ob ein Muttersprachler gewisse Ausdrücke verwenden würde oder nicht.

Um ein brauchbares Fehlertagset für die Annotation dieser Fehler in Falko erstellen zu können, muss die Vielzahl der verschiedenen Fehler auf eine überschaubare Zahl an Fehlerkategorien heruntergebrochen werden. Dazu ist ein gewisses Maß an Abstraktion nötig, es müssen zum Teil mehrere Fehlerarten in einer Kategorie zusammengefasst werden. Wie in Abschnitt 2.4.2. bereits erwähnt, stehen die Fehlerkategorien in einem Spannungsfeld zwischen zwei Desiderata: einerseits sollen sie gut ausgearbeitet und in der Lage sein, feine Unterschiede zu benennen, andererseits sollen sie auch schlicht und selbsterklärend sein.<sup>173</sup>

Diese vier Kategorien erfüllen meines Erachtens am besten die Anforderungen:

- A) Fehler bei der Bildung des Derivats
- B) Produktion von blockierten Bildungen
- C) Falscher Gebrauch des Derivats
- D) Fehler im Gebrauch des Plurals

Im Folgenden wird erklärt, welche Fehlerarten in diesen vier Kategorien zusammengefasst sind.

In Kategorie A finden sich zwei Fehlerarten vereint, zum einen die Bildung von *-ung*-Derivaten, denen ein nicht existenter Verbstamm zugrunde liegt (wie in Nr. 11, 25, 30 und 38)<sup>174</sup>, zum anderen die Bildung von Derivaten mit Verben, die aufgrund ihrer Semantik keine *-ung*-Suffigierung zulassen (wie in Nr. 39).<sup>175</sup>

In Kategorie B finden sich ebenfalls fehlerhafte Bildungen und zwar solche, die blockiert sind. Darunter einzuordnen sind vor allem die blockierten Bildungen, die von Muttersprachlern überhaupt nicht produziert werden, weil es usuelle Synonyme gibt, die in den entsprechenden Kontexten aktualisiert werden (wie z.B. in Nr. 4, 7 oder 22).

Kategorie C enthält all jene Fehler, welche mit dem fehlerhaften Gebrauch eines an sich korrekt gebildeten *-ung*-Derivats zu tun haben. Darunter fallen sehr verschiedene Formen von Normverstößen. Erstens betrifft dies semantische Fehler, in denen ein *-ung*-Derivat in einer Bedeutung gebraucht wird, die es nicht trägt (wie z.B. in Nr. 3, 4 oder 12) und durch einen angemesseneren Ausdruck ersetzt werden muss. Für diesen Fall gibt es sehr viele Belege und meist sind es eher „kleine“ Fehler, so dass das fehlerhafte Wort in der Zielhypothese durch ein sinnverwandtes und oft auch morphologisch verwandtes Wort ersetzt wird, wie zum Beispiel *Schätzung* vs. *Einschätzung* (Nr. 15) oder *Wirkung* vs. *Auswirkung* (Nr. 32). Zweitens gehört dazu der Verstoß gegen die

---

<sup>173</sup> Vgl. Ellis 2005: 60

<sup>174</sup> Wobei natürlich prinzipiell nicht zu unterscheiden ist, ob der Lerner das falsche Derivat durch Suffigierung eines nicht existenten Verbstammes gebildet hat oder als falsch gespeichertes Lexem aus dem Lexikon aktiviert hat. (Vgl. Abschnitt 2.1.5.2.)

<sup>175</sup> Vgl. außerdem Abschnitt 2.1.7.1.

Gebrauchsnormen fester Kollokationen wie *die Mühe lohnt sich* (Nr. 24) oder *in Bezug auf* (Nr. 37). Drittens lassen sich in die Kategorie C auch die uneindeutigen Fälle einordnen, in denen nicht ganz klar ist, ob wirklich ein Fehler vorliegt, oder ob es eher eine Stilfrage ist, die vom Lerner produzierte Äußerung für verbesserungswürdig zu halten (vgl. Nr. 8 und Nr. 17).

Kategorie D ist sicher diskussionswürdig. Die zentrale Frage, die sich stellt, lautet: Inwiefern sind Fehler bei der Pluralbildung als Fehler im Zusammenhang mit der *-ung*-Suffigierung zu sehen? Es gibt mindestens einen relativ klaren Fall, nämlich den, dass ein *nomen actionis*, das ein Geschehen als Tatsache ausdrückt, im Plural verwendet wird (wie in Nr. 6 und Nr. 34). Weniger klar ist das Funktionsverbgefüge *Verantwortung übernehmen*. Ist hier der Kollokationscharakter der Wortverbindung entscheidend dafür, dass der Plural inakzeptabel scheint? Und wie ist der Fall in Nr. 27? Handelt es sich hier vielleicht nur um eine Stilfrage und wäre der Plural eventuell möglich, die Lerneräußerung also gar nicht als Fehler zu markieren? Die Fragen müssen an dieser Stelle leider unbeantwortet bleiben, sie bieten eventuell Anlass für weitere Untersuchungen über den Zusammenhang von *-ung*-Derivaten und Pluralbildung.

Ein einziger Fehler aus der obigen Auflistung, der keiner Kategorie zugeordnet werden kann, bleibt nun übrig: die falsche Bildung des Kompositums *\*Stellungnahme* (Nr. 20). Dies kann als Übergeneralisierung der Regel, dass nach *-ung*-Derivaten als Kompositumerstglied eine *-s*-Fuge eingefügt werden muss, interpretiert werden; *Stellungnahme* ist eine der wenigen Ausnahmen von dieser Regel. Ich hatte anfangs damit gerechnet, dass eine fünfte Kategorie für Fehler bei der Bildung von Komposita mit *-ung*-Derivaten erforderlich sein würde und hatte fehlerhafte Bildungen, in denen die Fuge fehlt, erwartet. Diese Erwartung wurde von den Daten nicht bestätigt, das Gegenteil ist der Fall, die Kompositabildung scheint den Lernern im Prinzip keine Schwierigkeiten zu bereiten. Es bleibt zu überlegen, wie der Beleg *Stellungnahme* zu kategorisieren ist. Eine eigene Kategorie dafür zu eröffnen – und dafür anschließend auch ein Fehlertag ins Tagset aufzunehmen – halte ich zum jetzigen Zeitpunkt für verfrüht. Da das Falko-Korpus kein statisches Korpus ist, sondern noch wächst, besteht die Möglichkeit, zu einem späteren Zeitpunkt diese Fehleranalyse zu wiederholen und, falls sich noch mehr Fehler dieser Art finden, dann eine Kategorie für die Kompositumbildung anzunehmen. Denkbar wäre ansonsten auch eine Art „Rest-Kategorie“, in die alle Fehler aufgenommen werden, die nicht den Kategorien A-D zugeordnet werden können. Da es sich zum jetzigen Zeitpunkt aber nur um genau einen Fall handelt, halte ich eine Beschränkung auf die vier Kategorien zunächst für sinnvoll.

Die Belege Nr. 1-45 verteilen sich wie folgt auf die einzelnen Kategorien:

Kategorie A: Beleg Nr. 11, 30, 38 und 39

Kategorie B: Beleg Nr. 4, 7, 18, 22, 27, 29, 40-43

Kategorie C: Beleg Nr. 1-3, 5, 8, 12, 13, 15-17, 19, 21, 23-26, 28, 31-33, 35-37, 43 und 44

Kategorie D: Beleg Nr. 6, 9, 10, 14, 34

### 3.2.4. Ein mögliches Fehlertagset für Falko

Ausgehend von den vier Fehlerkategorien im vorangegangenen Abschnitt und Bezug nehmend auf das, was in Abschnitt 2.4.3. bereits über die Entwicklung eines Fehlertagsets gesagt wurde, möchte ich nun einen Vorschlag machen für ein mögliches Tagset, mit dem die Wortbildungsfehler im Bereich der *-ung*-Suffigierung in Falko annotiert werden können. Da es in Falko bisher noch kein Tagset für Wortbildungsfehler gibt, ist keine Orientierung an bestimmten Vorgaben oder Standards möglich. Lediglich zwei Dinge möchte ich aus der bestehenden Annotation von Wortarten und syntaktischen Feldern in Falko übernehmen: alle Buchstaben in den Tags werden groß geschrieben und die Tags sind so aufgebaut, dass sie – von links nach rechts gelesen – zuerst die allgemeineren Informationen kodieren und dann die spezielleren. Ein Beispiel dafür sind die „Part-of-speech-Tags“ für das Adjektiv, wo zuerst die Information „Adjektiv“ (ADJ) genannt wird und dann die Unterscheidung zwischen der prädikativen (ADJD) oder attributiven (ADJA) Verwendung getroffen wird.

Entsprechend sind meine Tags aufgebaut. Zuerst wird mit der Abkürzung „WB“ angezeigt, dass es sich um einen **W**ort**B**ildungsfehler handelt. „DE“ steht für **D**erivation (auf dieser Ebene wären später weitere Kürzel für andere Wortbildungsprozesse wie Konversion und Komposition denkbar). Mit „-UNG“ wird der spezielle Fall der *-ung*-Suffigierung benannt. Zuletzt stehen Abkürzungen für die Fehlerkategorien A-D: „BI“ für Fehler bei der **B**ildung, „BL“ für **b**lockierte Bildungen, „GE“ für Fehler im **G**ebrauch und „PL“ für Fehler bei der Bildung der **P**luralformen der Derivate. Im Endergebnis stehen folgende vier Fehlertags, deren Fehlerexponent jeweils das fehlerhafte *-ung*-Derivat in der Lerneräußerung ist:

1. **WBDE-UNGBI**
2. **WBDE-UNGBL**
3. **WBDE-UNGGE**
4. **WBDE-UNGPL**

Fehlerart	Fehlertag
Fehler bei der Bildung	WBDE-UNGBI
Fehler „blockierte Bildung“	WBDE-UNGBL
Fehler im Gebrauch des Derivats	WBDE-UNGGE

### 3.2.5 Diskussion der Ergebnisse der qualitativen Analyse

In der qualitativen Analyse wurden 45 Belege für die fehlerhafte Bildung bzw. den fehlerhaften Gebrauch von *-ung*-Derivaten untersucht (bei einer Gesamtzahl von 873 Tokens sind also etwas über 5% der belegten *-ung*-Derivate fehlerhaft). Ausgehend von einer konkreten Zielhypothese, die jedem Fehler zugewiesen wurde, wurden die Fehler beschrieben und anschließend kategorisiert. Die Ergebnisse der Fehleranalyse lassen eine Einteilung der Fehler in vier Kategorien (A-D) zu.

Die mit Abstand größte Kategorie ist die Kategorie C, der 26 Belege zugeordnet wurden. Hier finden sich alle *-ung*-Derivate, die in ihrem jeweiligen Kontext (Kollokation, Satz oder kompletter Text) in einer Art und Weise gebraucht sind, wie sie ein Muttersprachler nicht gebrauchen würde, etwa weil sie eine bestimmte Bedeutung nicht tragen oder in bestimmten festen Kollokationen nicht vorkommen. Die Belege in dieser Kategorie sind sicher nicht immer eindeutig, da sich die Fehleridentifikation hier auf Normen des Sprachgebrauchs beruft, welche meist weitaus weniger klar linguistisch beschrieben, bzw. beschreibbar sind als etwa Normen der Rechtschreibung, die eindeutig festgelegt sind. Man stößt hier auch wieder auf das Problem, dass man nicht mit Gewissheit sagen kann, wie ein Lerner ein bestimmtes komplexes Wort in der Äußerung aktualisiert hat (Lexikon vs. produktive Bildung). Wie ist ein Fehler wie etwa der falsche Gebrauch von *Verfolgungen* (nämlich im Sinne von *Folgen*) zu erklären (vgl. Nr. 31)? Hat der Lerner bereits die Bedeutung des Basisverbs *verfolgen* falsch abgespeichert und bildet damit produktiv das nomen acti *Verfolgungen*? Oder ist dem Lerner die Bedeutung von *verfolgen* bewusst, aber er aktiviert *Verfolgungen* aus seinem Lexikon als ein Lexem, das er mit falscher Bedeutung bzw. Gebrauchsnorm abgespeichert hat? Diese Fragen müssen hier unbeantwortet bleiben.

Immerhin neun Belege fallen in die zweitgrößte Kategorie B, die die blockierten Bildungen enthält. Das Auftreten dieses Fehlertypus kann als Beleg angesehen werden für die These, dass die fortgeschrittenen Lerner die Wortbildungsmuster der *-ung*-Suffigierung verinnerlicht haben und produktiv anwenden können, da davon auszugehen ist, dass sie den blockierten Bildungen nicht in zielsprachlichen Äußerungen, also in ihrem Input, begegnet sind. Sie beherrschen grundsätzlich die Regeln der *-ung*-Suffigierung, haben aber in ihrem Lexikon noch nicht alle Ausnahmen abgespeichert und deswegen kommt es zu Übergeneralisierungen.

Die Bildung der *-ung*-Derivate als solche bereitet den Lernern offenbar keine großen Schwierigkeiten. Lediglich vier Belege konnten Kategorie A zugeordnet werden und darunter sind sogar noch

zwei Belege, bei denen eine Interpretation als Rechtschreibfehler durchaus plausibel erscheint.

Ähnlich verhält es sich in Kategorie D, nur zwei der fünf Fehler bei der Pluralbildung konnten wirklich als Fehler im Zusammenhang mit dem, was in der Literatur zur *-ung*-Suffigierung und der Pluralbildung von *nomen actionis* zu finden ist, interpretiert werden. Bei den anderen drei Fehlern ist dieser Zusammenhang weniger klar. Insgesamt scheint auch dieser Bereich für die Lerner kein allzu großes Problem darzustellen.

Die untersuchten Daten zeigen eindeutig, dass bei den fortgeschrittenen Lernern völlige Klarheit herrscht über die kategoriellen Eigenschaften der *-ung*-Derivate (Nomen, Femininum, Singular), hier waren keine Fehler zu beobachten. Auch die Regel, dass bei der Kompositumbildung mit *-ung*-Derivat als Erstglied eine *-s*-Fuge eingefügt werden muss, wird beherrscht; lediglich ein einzelner Fall von Übergeneralisierung dieser Regel (*\*Stellungnahme*) war unter den Belegen.

Im Anschluss an die Fehlerkategorisierung wurde ein Fehlertagset vorgeschlagen, mit dem zukünftig die Fehler im Zusammenhang mit der *-ung*-Suffigierung in Falko annotiert werden könnten.

## 4. Fazit und Ausblick

Dieses abschließende Kapitel soll noch einmal die wichtigsten Erkenntnisse aus dieser Studie zusammenfassen und einen Ausblick geben auf daran anknüpfende Fragestellungen und Forschungsdesiderata.

In einer Korpusanalyse mit dem Lernerkorpus Falko wurde der Umgang von fortgeschrittenen Lernern des Deutschen als Fremdsprache mit einer produktiven morphologischen Wortbildungsregularität, der *-ung*-Suffigierung, untersucht. Zwei Teilanalysen, eine quantitative und eine qualitative, sollten zum einen Aufschluss darüber geben, wie produktiv die DaF-Lerner im Vergleich zu Muttersprachlern die produktiven Wortbildungsmuster verwenden, und zum anderen, welche Fehler sie dabei machen.

Als Ergebnis lässt sich zunächst festhalten, dass die Lerner die produktiven Muster zur Bildung von *nomina actionis* und *nomina acti* signifikant weniger häufig benutzen als Muttersprachler. Das belegen die Werte der *potential productivity* für diese beiden Muster. Dennoch spricht vieles dafür, dass die Lerner die Wortbildungsmuster verinnerlicht haben und produktiv für Neubildungen nutzen können. Dies zeigt vor allem das Auftreten von regelkonform gebildeten *-ung*-Derivaten, die nach der zielsprachlichen Gebrauchsnorm durch usuelle Synonyme blockiert sind. Sowohl die Bildung von *-ung*-Derivaten an sich als auch die Bildung von deren Pluralformen und von Komposita mit *-ung*-Derivaten als Erst- oder Zweitglied bereitet den Lernern keine Schwierigkeiten. Fehler machen sie dagegen noch beim Gebrauch der Derivate, die sie öfter mit falschen Bedeutungen verwenden. Vor allem im Zusammenhang mit dieser Art von Fehler, aber auch an anderen Stellen ist deutlich geworden, dass es zukünftig Klärungsbedarf gibt in der Frage, wie ein komplexes Wort in der Lernerproduktion aktualisiert wurde – durch Abruf aus dem Lexikon oder durch produktive

Bildung. Des Weiteren besteht noch grundsätzlicher Bedarf an einer präzisen und umfangreichen linguistischen Beschreibung der Regeln und Beschränkungen der *-ung*-Suffigierung, hier liefert die theoretische Literatur noch unzureichende Informationen.

Ein weiteres Ergebnis dieser Arbeit ist der Vorschlag für ein Fehlertagset, mit dem zukünftig in Falko Lernerfehler im Bereich der *-ung*-Suffigierung annotiert werden könnten. Dieses Fehlertagset erlaubt meines Erachtens eine genügend differenzierte, aber dennoch überschau- und nachvollziehbare Annotation und bietet die Möglichkeit sehr konkreter Suchabfragen, z.B. die Suche nach allen von den Lernern produzierten *-ung*-Derivaten, die von der zielsprachlichen Norm eigentlich blockiert sind. Für eine abschließende Bewertung des Fehlertagsets ist allerdings erforderlich, in einem kontrollierten Verfahren eine Urteilerübereinstimmung<sup>176</sup> zu ermitteln. Dies steht noch aus. Anknüpfend an diese Arbeit sind weitere vergleichbare Studien denkbar, in denen der Gebrauch anderer Wortbildungsaffixe bei fortgeschrittenen DaF-Lernern untersucht wird, um so zu einer umfassenden Vorstellung über den Erwerb von morphologischen Wortbildungsregularitäten zu gelangen. Ein Vergleich mit Produktionen von Lernern verschiedener Niveaustufen könnte Aufschluss darüber geben, ab welcher Erwerbsstufe Lerner zum produktiven Gebrauch von Wortbildungsmustern in der Lage sind und inwieweit dies ein Kriterium für die Feststellung von Fortgeschrittenheit sein kann. Die Ergebnisse der quantitativen Analyse, die einen Underuse von *-ung*-Derivationen in den Lernerproduktionen nachweisen, geben Anlass darüber nachzudenken, was der Grund dafür ist und ob es didaktische Möglichkeiten gibt, wie man die weit fortgeschrittenen Lerner einem zielsprachlichen Umgang mit der *-ung*-Suffigierung näher bringt.

Des Weiteren bieten sich Studien an, die als Ziel haben, eine Erklärung für die Fehler, die in der Fehleranalyse beschrieben wurden, zu liefern. Beispielsweise ließe sich untersuchen, ob Transfer aus der Muttersprache der Lerner hier eine Rolle spielt. Dafür ist allerdings eine Erweiterung der Datengrundlage erforderlich, denn eine Aufteilung der wenigen Fehler, die gefunden wurden, auf die verschiedenen Muttersprachen der Lerner, würde zu nicht aussagekräftigen Ergebnissen führen. Allerdings wächst das Falko-Korpus ja noch und es wäre sicher durchaus angebracht, die hier gefundenen Ergebnisse irgendwann auf der Grundlage einer größeren Datenmenge zu überprüfen.

---

<sup>176</sup> Vgl. Lüdeling/Walter 2009

## 5. Quellenverzeichnis

### 5.1 Literatur

Bauer, Laurie: Morphological Productivity. Cambridge: Cambridge University Press 2001.

Baayen, Rolf Harald: Corpus linguistics in morphology: Morphological productivity. In: Corpus Linguistics. An International Handbook. Vol. 2. Hrsg. von Anke Lüdeling & Merja Kytö. Berlin/New York: Walter de Gruyter 2009 (=Handbücher zur Sprach- und Kommunikationswissenschaft 29.2). S. 899-919.

Becker, Angelika: The semantic knowledge base for acquisition of negation and the acquisition of finiteness. In: Hendricks, Heriette: The Structure of Learner varieties. Berlin/New York: Mouton de Gruyter 2005. S. 263-314.

Chomsky, Noam: Syntactic Structures. Den Haag: Mouton de Gruyter 1957.

Chomsky, Noam: Syntactic Structures. Den Haag (u.a.): Mouton de Gruyter 1965.

Corder, Pit (1967): The significance of learners' errors. In: Corder, Pit: Error analysis and interlanguage. Oxford: Oxford University Press 1981. S. 5-13.

Corder, Pit: Error analysis and interlanguage. Oxford: Oxford University Press 1981.

Corpus Linguistics. An International Handbook. Vol. 1. Hrsg. von Anke Lüdeling & Merja Kytö. Berlin/New York: Walter de Gruyter 2008 (=Handbücher zur Sprach- und Kommunikations-

wissenschaft 29.1).

Corpus Linguistics. An International Handbook. Vol. 2. Hrsg. von Anke Lüdeling & Merja Kytö. Berlin/New York: Walter de Gruyter 2009 (=Handbücher zur Sprach- und Kommunikationswissenschaft 29.2).

Corpus Linguistics. Critical Concepts in Linguistics. Bd. 2. Hrsg. von Wolfgang Teubert & Ramesh Krishnamurthy. London/New York: Routledge 2007.

Deutsch als Fremdsprache. An den Quellen eines Faches. Hrsg. von Heidrun Popp. München: Iudicium Verlag 1995.

Diehl, Erika: Grammatikunterricht: Alles für der Katz? Tübingen: Max Niemeyer 2000.

Donalies, Elke: Die Wortbildung des Deutschen. Tübingen: Narr 2005.

Draeger, Kerstin: Die semantische Leistung der suffixalen Wortbildungsmorpheme der Substantive in der deutschen Gegenwartssprache. Aachen: Shaker 1996.

Duden Bd. 1: Die deutsche Rechtschreibung. 22. Auflage. Hrsg. von der Dudenredaktion. Mannheim: Dudenverlag 2000.

Duden Bd. 5: Das Fremdwörterbuch. 7. Auflage. Hrsg. von der Dudenredaktion. Mannheim: Dudenverlag 2001.

Duden Bd. 10: Das Bedeutungswörterbuch. 2. Auflage. Hrsg. von Wolfgang Müller. Mannheim: Dudenverlag 1985.

Ellis, Rod: Understanding Second Language Acquisition. Oxford: Oxford University Press 1985.

Ellis, Rod: The Study of Second Language Acquisition. Oxford: Oxford University Press 1995.

Ellis, Rod: Analysing Learner Language. Oxford: Oxford University Press 2005.

- Erben, Johannes: Einführung in die deutsche Wortbildungslehre. Berlin: Erich Schmidt 1975.
- Erben, Johannes: Wortbildung und Textbildung. In: Deutsch als Fremdsprache. An den Quellen eines Faches. Hrsg. von Heidrun Popp. München: Iudicium Verlag 1995. S. 545-552.
- Erben, Johannes: Einführung in die deutsche Wortbildungslehre. 4. Aufl. Berlin: Erich Schmidt 2000.
- Error Analysis. Perspectives in second language acquisition. Hrsg. von Jack C. Richards. London: Longman 1974.
- Fleischer, Wolfgang und Irmhild Barz: Wortbildung der deutschen Gegenwartssprache. Tübingen: Max Niemeyer 1992.
- Fortgeschrittene Lernervarietäten. Korpuslinguistik und Zweitsprachenerwerbsforschung. Hrsg. von Maik Walter & Patrick Grommes. Tübingen: Max Niemeyer 2008.
- Fremde Sprache Deutsch: grammatische Beschreibung - Erwerbsverläufe – Lehrmethodik Hrsg. von Brigitte Handwerker. Tübingen: Narr 1995.
- Fries, Charles C.: Teaching and learning English as a foreign language. Ann Arbor: University of Michigan Press 1945.
- Hendricks, Heriette: The Structure of Learner varieties. Berlin/New York: Mouton de Gruyter 2005.
- Henzen, Walter: Deutsche Wortbildung. Halle an der Saale: Niemeyer 1947.
- Klein, Wolfgang & Clive Perdue (1997): The Basic Variety (or: Couldn't natural languages be much simpler?). In: Second Language Research 13, 301-347.
- Köhlmyr, Pia: To Err Is Human. Göteborg: Acta Universitatis Gothoburgensis 2003.
- Kohn, Kurt: Dimensionen Lernaltsprachlicher Performanz: theoretische und empirische Untersuchungen zum Zweitspracherwerb. Tübingen: Narr 1990.

Lado, Robert: Linguistics across cultures. Ann Arbor: University of Michigan Press 1945.

Lehmann, Christian: Daten – Korpora – Dokumentation. In: Sprachkorpora – Datenmengen und Erkenntnisfortschritt. Hrsg. von Werner Kallmeyer & Gisela Zifonun. Berlin/New York: Walter de Gruyter 2007. S. 9-27.

Lemnitzer, Lothar und Heike Zinsmeister: Korpuslinguistik. Eine Einführung. Tübingen: Narr 2006.

Lüdeling, Anke; Maik Walter; Emil Kroymann; Peter Adolphs et al.: Multi-level error annotation in learner corpora. In: Proceedings of Corpus Linguistics. Birmingham 2005.

Lüdeling, Anke (2007): Das Zusammenspiel von qualitativen und quantitativen Methoden in der Korpuslinguistik. In: Sprachkorpora – Datenmengen und Erkenntnisfortschritt. Hrsg. von Werner Kallmeyer & Gisela Zifonun. Berlin/New York: Walter de Gruyter 2007. S. 28-48.

Lüdeling, Anke: Mehrdeutigkeiten und Kategorisierung: Probleme bei der Annotation von Lernerkorpora. In: Fortgeschrittene Lernervarietäten. Korpuslinguistik und Zweitsprachenerwerbsforschung. Hrsg. von Maik Walter & Patrick Grommes. Tübingen: Max Niemeyer 2008. S. 119-140.

Lüdeling, Anke et. al.: Das Lernerkorpus Falko. In: Deutsch als Fremdsprache 2. München/Berlin 2008. S. 67-73.

Lüdeling, Anke & Walter Maik (2009): Korpuslinguistik. Veröffentlicht als pdf-Datei unter [http://www.linguistik.hu-berlin.de/institut/professuren/korpuslinguistik/mitarbeiterinnen/anke/anke\\_veroeffentlichungen](http://www.linguistik.hu-berlin.de/institut/professuren/korpuslinguistik/mitarbeiterinnen/anke/anke_veroeffentlichungen)

Mikolajczyk, Monika: Von der linguistischen Beschreibung zur Erfassung eines Lerngegenstands. '-ung'-Nominalisierung im Deutschen. Unveröffentlichte Magisterarbeit. Berlin 2002.

Motsch, Wolfgang: Semantische und pragmatische Aspekte der Wortbildung. In: Deutsch als

Fremdsprache. An den Quellen eines Faches. Hrsg. von Heidrun Popp. München: Iudicium Verlag 1995. S. 513-531.

Motsch, Wolfgang: Deutsche Wortbildung in Grundzügen. Berlin/New York: Walter de Gruyter 1999.

Naumann, Bernd: Einführung in die Wortbildungslehre des Deutschen. Tübingen: Max Niemeyer 1986.

Neologie und Korpus. Hrsg. von Wolfgang Teubert. Tübingen: Narr 1998.

Olsen, Susan: Wortbildung im Deutschen. Eine Einführung in die Theorie der Wortstruktur. Stuttgart: Kröner 1986.

Plag, Ingo: Morphological productivity: structural constraints in English derivation. Berlin/New York: Mouton de Gruyter 1999.

Putzer, Oskar: Fehleranalyse und Sprachvergleich: linguistische Methoden im Fremdsprachenunterricht am Beispiel Italienisch-Deutsch. Ismaning: Max Hueber Verlag 1994.

Rainer, Franz: Spanische Wortbildungslehre. Tübingen: Max Niemeyer 1993.

Scherer, Carmen: Wortbildungswandel und Produktivität. Eine empirische Studie zur nominalen *-er*-Derivation im Deutschen. Tübingen: Max Niemeyer 2005.

Scherer, Carmen: Korpuslinguistik. Heidelberg: Universitätsverlag Winter 2006.

Sinclair John: Korpustypologie. Ein Klassifikationsrahmen. In: Neologie und Korpus. Hrsg. von Wolfgang Teubert. Tübingen: Narr 1998. S. 111-128.

Sinclair, John: Intuition and annotation. The discussion continues. In: Corpus Linguistics. Critical Concepts in Linguistics. Bd. 2. Hrsg. von Wolfgang Teubert & Ramesh Krishnamurthy. London/New York: Routledge 2007. S. 415-435.

Selinker, Larry (1972): Interlanguage. In: Error Analysis. Perspectives in second language

acquisition. Hrsg. von Jack C. Richards. London: Longman 1974. S. 31-54.

Selinker, Larry: Rediscovering Interlanguage. London / New York: Longman 1992.

Sprache der Gegenwart. Hrsg. vom Institut für deutsche Sprache. Band 79. Berlin/New York: Walter de Gruyter 1991 (=Deutsche Wortbildung. Typen und Tendenzen in der Gegenwartssprache. Vierter Hauptteil: Substantivkomposita).

Sprachkorpora – Datenmengen und Erkenntnisfortschritt. Hrsg. von Werner Kallmeyer & Gisela Zifonun. Berlin/New York: Walter de Gruyter 2007.

Storch, Günther: Deutsch als Fremdsprache – eine Didaktik. München: Wilhelm Fink Verlag 2001.

Teubert, Wolfgang: Korpus und Neologie. In: Neologie und Korpus. Hrsg. von Wolfgang Teubert. Tübingen: Narr 1998. S. 129-171.

Vogel, Klaus: Lernersprache: linguistische und psycholinguistische Grundlagen zu ihrer Erforschung. Tübingen: Narr 1990.

Wellmann, Hans: Deutsche Wortbildung. Typen und Tendenzen in der Gegenwartssprache. Bd. 2: Substantiv. Düsseldorf: Pädagogischer Verlag Schwann 1975.

Walter, Maik & Patrick Grommes: Die Entdeckung des fortgeschrittenen Lernalters in der Varietätenlinguistik. In: Fortgeschrittene Lernervarietäten. Korpuslinguistik und Zweitsprachenerwerbsforschung. Hrsg. von Maik Walter & Patrick Grommes. Tübingen: Max Niemeyer 2008. S. 3-27.

## **5.2. Untersuchte DaF-Lehrwerke:**

Aufderstraße, Hartmut; Jutta Müller und Thomas Storz: Delfin. Lehrwerk für Deutsch als Fremdsprache. Ausgabe in 2 Bd. Ismaning: Hueber 2002.

Aufderstraße, Hartmut; Werner Bönzli und Walter Lohfert: Themen neu. Lehrwerk für Deutsch als Fremdsprache. Kursbuch 3. Ismaning: Max Hueber Verlag 1994.

Baumann, Evgenija: Gute Deutsche Grammatik. Konstanz: Baumann Verlag 2004.

Dodd, Bill; Christine Eckhard-Black et al.: Modern German Grammar. A Practical Guide. 2. Auflage. London: Routledge 2003.

Lemcke, Christiane; Lutz Rohrmann, Theo Scherling: Berliner Platz 1: Deutsch im Alltag für Erwachsene. Lehr- und Arbeitsbuch. Berlin (u.a.): Langenscheidt 2000.

Neuner, Gerd: Deutsch aktiv Neu. Ein Lehrwerk für Erwachsene. Bd. 1A, Lehrbuch. Berlin (u.a.): Langenscheidt 1994.

Perlmann-Balme, Michaela: EM-Hauptkurs: Deutsch als Fremdsprache für die Mittelstufe. Ismaning: Hueber 2000.

Wege. Deutsch als Fremdsprache: Mittelstufe und Studienvorbereitung. Hrsg. von Dietrich Eggers. Ismaning: Hueber 1988.

### **5.3. Internet-Links**

<http://www.c-test.de>

<http://www.goethe.de/z/50/commeuro>

<http://www.linguistik.hu-berlin.de/institut/professoren/korpuslinguistik/forschung/falko>

<http://www.ims.uni-stuttgart.de/projekte/corplex/TreeTagger/>

(Alle URLs wurden überprüft am 15.09.2009)

## **6. Anhang**

### **6.1. Abstract**

Die Fähigkeit, Wortbildungsregularitäten zu erkennen und für die Rezeption und Produktion von komplexen Wörtern zu nutzen, gehört zur Sprachkompetenz eines Muttersprachlers. Wortbildungsmuster sind ein Bestandteil der Lexikonkomponente der mentalen Grammatik. Wie ist das bei Lernern des Deutschen als Fremdsprache? Haben zumindest fortgeschrittene Lerner diese Kompetenz ebenfalls ausgebildet?

Die vorliegende Studie untersucht dies anhand einer Korpusanalyse mit Falko am Beispiel der produktiven Wortbildungsmuster der *-ung*-Suffigierung. Die Ergebnisse zeigen, dass die Lerner durchaus die Wortbildungsmuster erkennen und produktiv anzuwenden. Sie tun dies allerdings weitaus weniger häufig als Muttersprachler. Außerdem machen sie eine Reihe von Fehlern, die in der Studie genauer beschrieben und kategorisiert werden. Anhand dieser Fehlerkategorien wird zuletzt ein mögliches Fehlertagset für Falko entwickelt.

## 6.2. Eigenständigkeitserklärung:

Ich erkläre, dass ich die vorstehende Seminararbeit mit dem Titel

**„Treffungen, Sinkung und Benützung -  
Korpuslinguistische Untersuchung des Erwerbs von derivationsmorphologischen  
Wortbildungsregularitäten bei fortgeschrittenen Lernern des Deutschen als Fremdsprache“**

selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

---

Ort, Datum

---

Unterschrift

### **6.3. Lebenslauf**

#### **Persönliche Daten:**

Name: Michael Berth  
Adresse: Gustav-Freytag-Str. 1/2, 10827 Berlin  
Telefon: 030-21463908 und 0179-5115693  
Email-Adresse: mizzla@gmx.de  
Familienstand ledig  
Staatsangehörigkeit deutsch  
Geburtsdaten: 31. Oktober 1978 in Reutlingen

#### **Wichtigste Stationen:**

Seit November 2008: Nebentätigkeit als DaZ-Lehrer an der Staatlichen Ballettschule Berlin und Schule für Artistik

März 2003 – heute: Magister-Studium an der Humboldt Universität zu Berlin, Fächerkombination Linguistik und Kulturwissenschaft; voraussichtlicher Abschluss: Frühjahr 2010

September 2001 – August 2007: Autor der Reggae-Sendung „JAMaican VibeZ“ (Moderation, musikalische und inhaltliche Planung) beim Radiosender JAM FM (Berlin)

Juli 2000 – März 2003: Festanstellung als Redakteur beim Radiosender  
JAM FM der Skyline Medien GmbH, Berlin

August 1998 – August 1999: Zivildienst bei den „Johannitern“ in Reutlingen

Juni 1998: Abitur am Gymnasium des Bildungszentrums  
Reutlingen-Nord (Abschlussnote 1,8)