

Über Wortarten im Hawaiischen

-

Eine korpusgestützte Analyse

Inaugural-Dissertation

zur Erlangung des Doktorgrades der Philosophie
der Ludwig-Maximilians-Universität
München

vorgelegt von

Benjamin Weber

aus

Hemer

2024

Referent: PD Dr. Peter-Arnold Mumm

Korreferentin: Prof. Dr. Katja Hannß

Tag der mündlichen Prüfung: 15.07.2022

• This work is licensed under CC BY 4.0. •

Inhaltsverzeichnis

Tabellenverzeichnis	iii
Abbildungsverzeichnis	iv
Abkürzungsverzeichnis	vi
1 Einleitung	1
1.1 Untersuchungsgegenstand	1
1.2 Aufbau der Arbeit	3
2 Forschungshintergrund	4
2.1 Die griechisch-lateinische Tradition und die Grenzen der Uni- versalität	4
2.2 Wortarten in der Beschreibung polynesischer Sprachen	8
2.2.1 Übertragbarkeit der tradierten Wortartkategorien	11
2.2.2 Wortarten als syntaktische Kategorien	14
2.2.3 Ablehnung von Wortartkategorien	16
2.2.4 Prototypische Abdeckung semantischer Klassen	25
2.3 Zusammenfassung und These	55
3 Methodik	61
3.1 Motivation für eine Korpusanalyse	62
3.2 Überblick Korpusanalyse	65
3.3 Korpusdesign	66
3.3.1 Datenerschließung	66
3.3.2 Korpuserstellung	67
3.4 Automatisiertes Verfahren	78
3.4.1 Grundlagen und Überblick	78
3.4.2 Normalisierung	80
3.4.3 Tokenisierung	82
3.4.4 Tagging	83
3.4.5 Chunking	90
3.4.6 Weitere Schritte und Typisierung	99
3.5 Zusammenfassung	104
4 Auswertung und Ergebnisse	107
4.1 Wortverteilung und Distributionsklassen	107
4.1.1 Überblick der Wortverteilung im Korpus	108

4.1.2	Abgleich Korpusbeobachtungen und Wörterbucheinträge	118
4.1.3	Nominalisierung im Hawaiischen	131
4.1.4	<i>Verbale</i> Derivationsmorphologie	167
4.1.5	Zusammenfassung Distribution	175
4.2	Häufigkeitsverteilung und Frequenzklassen	176
4.2.1	Frequenz und Prototypikalität	177
4.2.2	Frequenz und Klassenbildung	179
4.2.3	Zusammenfassung Frequenzklassen	196
4.3	semantische Klassen	198
4.3.1	Grundannahmen	198
4.3.2	Bestimmung einer semantischen Klasse eines Wortes . . .	200
4.3.3	Bestimmung mehrerer semantischer Klassen eines Wortes	211
4.3.4	Zusammenfassung semantische Klassen	216
5	Fazit und Ausblick	218
	Literaturverzeichnis	221

Tabellenverzeichnis

1	Vergleich der Verb-Nomen-Unterscheidung in Einzelsprachen . . .	7
2	Metaanalyse Wortarten in polynesischen Sprachen	10
3	prototypische Abdeckung semantischer Klassen auf pragmatische Funktionen (vgl. Croft 2001: 88)	26
4	Definition semantischer Klassen durch semantische Merkmale (vgl. Croft 2001: 88)	27
5	Gegenüberstellung von prototypischen und nicht-prototypischen Fällen bei Croft (2001)	31
6	Zusammenhang von semantischer Klasse und Flexibilität eines Wortes (vgl. Völkel 2017: 465)	46
7	Vergleich lexikalisch-semantischer Theorien	59
8	Historische Entwicklung des Hawaiischen	70
9	Ausschnitt Tabelle mit Frequenzen von Types in syntaktischen Konstruktionen	99
10	Vergleich der Angaben zu morphologisch markierten Pluralformen von Elbert und Pukui (1979: 106) und der im Korpus gefundenen Formen	102
11	Überblick Wortverteilung im Korpus	108
12	Vergleich erfasste und nicht erfasste Wörter im Korpus	117
13	Angleichung der im Korpus und in Wörterbüchern verwendeten Kategorien	119
14	Ausschnitt Vergleichstabelle zum Wortauftreten	121
15	Zusammenfassung Vergleich Korpus und Wörterbücher	122
16	Korpus-Wörterbuch-Vergleich flexibler Wörter	125
17	erfasste und nicht-erfasste klassifizierte Wörter	131
18	Vorkommenshäufigkeit von Syntagmen im Korpus	132
19	Reflexe von <i>*(C)anga</i> in polynesischen Sprachen	135
20	Vorkommen Basen und <i>-na</i> -Derivate im Korpus	140
21	Zusammenfassung Basis-Derivat-Kategorien	142
22	Überblick <i>verbale</i> Derivationsmorphologie	168
23	Verteilung <i>verbaler</i> Derivate im Korpus	171
24	Korrigierte Kategorien <i>verbaler</i> Basen und <i>verbaler</i> Derivate . .	172
25	Nicht-kategoriale derivationsmorphologische Prozesse	174
26	Kombinationen gleiche Anteile in Syntagmen	183

Abbildungsverzeichnis

1	Schematisierung von Ereignisvorstellungen (vgl. Schulze 2012)	29
2	Zusammenfassung des Ansatzes von Croft (2001)	38
3	Prädisposition semantischer Klassen	51
4	Vergleich von Völkel (2017), Don und Van Lier (2013) und Croft (2001)	53
5	Entwicklung Wortarttheorien	56
6	Alternative Ansätze zur Wortartkategorisierung in den polyne- sischen Sprachen	57
7	Überblick Schritte in der Korpusanalyse	65
8	Einzelschritte im automatisierten Verfahren	79
9	Schritte beim Trainieren eines Taggers	84
10	Tagging-Modell	88
11	Beispiel für kombiniertes Tagging-Modell	90
12	schrittweise Entwicklung zum Chunking	91
13	Aufbau des TAM-Syntagmas im Hawaiischen	93
14	Aufbau des Artikel-Syntagmas im Hawaiischen	94
15	Berechnung der Kombinationsmöglichkeiten von präponierten Elementen	96
16	Beispiel Chunking-Ablauf	97
17	Zusammenfassung Korpusanalyse	105
18	Wortartsystem von Elbert und Pukui (1979)	113
19	Venn-Diagramm Distributionsklassen	115
20	Venn-Diagramm korrigierter Distributionsklassen	129
21	Vergleich Nominalisierungsstrategien in polynesischen Sprachen	137
22	Kategorien identifizierbarer Basis-Derivat-Paare im Hawaiischen	144
23	Kategorien in „Nominalisierungskonstruktion“	151
24	Zusammenfassung Beschreibungsversuch <i>Nominalisierung</i>	163
25	beobachtete Auftretensweisen <i>nominalisierter</i> Wörter	165
26	Endgültig aus Distributionsklassen gebildete Kategorien	166
27	Vergleich Distributionsklassen und Frequenzklassen	178
28	Vergleich Anteile in Syntagmen	180
29	Vergleich Anteile in Artikel-Syntagma mit und ohne NMLZ	181
30	Vergleich Anteile in Syntagmen und Flexibilität	184
31	Vergleich <i>Attraction</i> in Syntagmen	187
32	Rang-Frequenz-Plot auf einer logarithmischen Skala	188
33	<i>Attraction</i> und Flexibilität	190

34	<i>Attraction</i> und Flexibilität	191
35	Relative Chancen Wörter und Syntagmen	194
36	<i>Odds Ratio</i> und Flexibilität	195
37	Odds und Flexibilität	196
38	Lexikalisierte Grundbedeutung und Auftreten	202
39	Wortarten als <i>Typological Prototypes</i>	212

Abkürzungsverzeichnis

1 Erste Person	LOC Lokativ
2 Zweite Person	NEG Negation
3 Dritte Person	NMLZ Nominalisierer
ABS Absolutiv	PASS Passiv
ANAPH Anaphorisch	PFV Perfektiv
ART Artikel	PL Plural
ATTR Attributiv	POSS Possessiv
DEF Definit	PREP Präposition
DEM Demonstrativ	PRES Präsens
DIR Direktional	PROX Proximal
DIST Distal	PST Past
IDF Indefinit	SG Singular
INF Infinitiv	TAM Tempus-Aspekt-Modus
INTV Intentiv	TOP Topik
IPFV Imperfektiv	
ITSF Intensivierer	

1 Einleitung

1.1 Untersuchungsgegenstand

Es fragt sich nämlich [...], ob da, wo verbum und nomen nicht in lautlich gesonderter weise existieren, diese unterscheidung auch in der function fele, also überhaupt nicht vorhanden sei, oder wir ein recht haben, auch in solchen sprachen, die nomen und verbum lautlich nicht unterscheiden, dennoch das vorhandensein dises gegensatzes an zu nemen.

(Schleicher 1865: 501)

Der Untersuchungsgegenstand der vorliegenden Arbeit soll an folgenden Beispielen aus dem Hawaiischen erläutert werden:

- (1a) *E hana ana 'o Herman me ka hale*
IPFV √arbeit DIST TOP Herman bei DEF Haus
pau-ahi in Hilo.
zerstört-Feuer in Hilo

‘Herman arbeitete bei der Feuerwache in Hilo.’

(Kauhi 1996:26)

- (1b) ‘*O ia nō ka hana a nā po‘e kāne*
TOP 3SG ITSF DEF √arbeit POSS PL Mensch männlich\PL
‘Dies (war) die Arbeit der Männer.’
(Kauhi 1996:6)

Vergleicht man die beiden Sätze miteinander, kann aufgrund der fehlenden Flexionsmorphologie formal nicht zwischen *hana* in Satz (1a) und *hana* in Satz (1b) unterschieden werden. So konnte *hana* in der gewählten Metasprache

Deutsch nicht besser als mit der Wurzel $\sqrt{\text{arbeit}}$ glossiert werden.¹ Bereits 1837 hat Adelbert von Chamisso in seiner Grammatikskizze *Über die Hawaiische Sprache* festgestellt: „Die Wurzelwörter [...] sind durchaus unveränderlich. Sie sind ohne Unterschied Haupt- und Beiwort, Substantiv und Verbum, Adjectiv und Adverbium, und das Fachwerk unserer Grammatik findet auf sie keine Anwendung.“ (von Chamisso 1837: 8). Zwar seien nach von Chamisso Wurzelwörter unter anderem sowohl Verb als auch Substantiv, allerdings: „[...] ihre jedesmalige Geltung in dieser Hinsicht erhalten sie von ihrer Stellung im Satze“ (von Chamisso 1837: 8). Als bald lässt sich diese Annahme auf alle polynesischen Sprachen verallgemeinern. So schreibt Eduard Buschmann (1839: 842) in seiner vergleichenden Grammatik: „Die Südsee-Sprachen haben die Ununterscheidbarkeit der Redetheile mit den westlichen [Sprachen des Malayischen Sprachstammes] gemein; dasselbe Wort kann die Eigenschaft eines Subjt., Adj. [sic], Verbums u. s. w. in sich vereinen.“ Die Feststellung von Chamissos mag die Beobachtung widerspiegeln, dass Wörter im Hawaiischen in ihrem aktuellen Auftreten zwar als Verb wie *hana* in Satz (1a) oder als Nomen wie *hana* in Satz (1b) in eine Sprache mit Verb-Nomen-Unterscheidung übertragen werden können. Damit lassen sich jedoch anders als im Deutschen mit *arbeiten* und *Arbeit* nicht zwei Wörter *hana* und *hana* gegenüberstellen. Für das Deutsche gilt es, zwei Redeweisen von *Wort* zu unterscheiden, während dies für das Hawaiische nicht sinnvoll ist:² erstens die Redeweise im Sinne von *Wortform* und zweitens die Redeweise im Sinne von *Lexem*, zu dem eine Wortform gehört. Durch diese Unterscheidung wird zugleich die Grundlage für die Unterscheidung von Wortarten als Lexemklassen gelegt.³ Ein Lexem ist die Abstraktion über Wortformen eines Flexionsparadigmas. Die Wortform *arbeitete* als Teil desjenigen Flexionsparadigmas, in dem morphologisch die grammatischen Kategorien Person, Numerus, Tempus und Modus ausgedrückt werden, gehört zu einem verbalen Lexem, das morphologisch repräsentiert wird durch den Stamm *arbeit* und qua Konvention im Deutschen mit der Infinitivform *arbeiten* zitiert wird. Die Wortform *Arbeit* hingegen gehört als Teil desjenigen Flexionspara-

¹Der besseren Lesbarkeit halber sind die anderen Wörter in den Sätzen nicht als Wurzeln glossiert, obwohl dies analog zu *hana* richtig wäre.

²Mir geht es hier um das *Wort* als grammatisches Wort und ich sehe daher von weiteren Definitionen wie *Wort* als phonologisches Wort ab.

³Die folgenden Ausführungen können gewiss keine erschöpfende Behandlung des Themenkomplexes Morphologie im Deutschen oder der Flexion allgemein darstellen. Sie dienen lediglich der exemplarischen Gegenüberstellung des Hawaiischen und Deutschen, um die Problematik in der Beschreibung des Hawaiischen herauszustellen.

digmas, in dem morphologisch die grammatischen Kategorien Genus, Numerus und Kasus ausgedrückt werden, zu einem nominalen Lexem, das morphologisch repräsentiert wird durch den Stamm *arbeit* und qua Konvention im Deutschen mit der Form des Nominativ Singular *Arbeit* zitiert wird.

Eine analoge Analyse ist im Hawaiischen nicht möglich. Auch die morphologische Repräsentation beider Lexeme *arbeiten* und *Arbeit* durch den identischen Stamm *arbeit* stellt einen anderen Fall dar als das hawaiische *hana*. Denn unabhängig von der theoretischen Behandlung im Deutschen – ich verzichte hier bewusst auf Begriffe wie *Konversion* oder *Nullderivation* – fehlt im Hawaiischen die Unterscheidbarkeit eines Wortes im obigen Sinne von Lexem und seinen in einem Paradigma erscheinenden Formen. Zudem gilt: „Problematisch wird es dann, wenn die Null-Ableitung regelhaft ist, und wenn es keine anderen unabhängigen Kriterien gibt für die Etablierung von Nomen und Verben als Wortarten.“ (Helmbrecht 2005: 10).

Es lässt sich sodann unter Rückgriff auf das vorangestellte Zitat aus der 1865 erschienenen Abhandlung *Die Unterscheidung von Nomen und Verbum in ihrer lautlichen Form* von August Schleicher der Untersuchungsgegenstand dieser Arbeit in der folgenden Frage formulieren: Müssen Wortarten als Beschreibungsgrößen in einer flexionslosen Sprachen wie dem Hawaiischen, in der Wörter nicht in Formklassen geschieden sind, abgelehnt werden oder nicht?

1.2 Aufbau der Arbeit

Die Arbeit gliedert sich in fünf Kapitel. Nach der vorangegangenen Klärung des Untersuchungsgegenstandes wird zunächst ein Überblick über den Forschungshintergrund gegeben, um so die zu überprüfenden Annahmen präziser zu fassen und die eigene These zu entwickeln (Kapitel 2). Es folgt die Begründung und Darstellung der eingesetzten Methoden (Kapitel 3). Im vierten Kapitel werden die Ergebnisse aus der durchgeführten automatisierten Korpusanalyse präsentiert und in weiteren Analysen problematisiert.

Die Arbeit schließt mit einer Zusammenfassung der wichtigsten Ergebnisse und den sich daraus ergebenden Konsequenzen für die Wortarttheorie.

2 Forschungshintergrund

Da es nicht möglich ist, die Beschreibungsversuche von Wortarten in den polynesischen Sprachen zu verstehen, ohne den größeren theoretischen Rahmen darzustellen, aus dem heraus diese entstanden sind, soll hier zuerst die Entwicklung der Wortarttheorie aus der griechisch-lateinischen Tradition bis zur Frage nach den Grenzen ihrer universellen Geltung skizziert werden.

2.1 Die griechisch-lateinische Tradition und die Grenzen der Universalität

Die Grundlegung dessen, was von Chamisso (1837: 8) das „Fachwerk unserer Grammatik“ nennt und man als traditionelle Wortartlehre der abendländischen Sprachwissenschaft bezeichnen kann, beginnt in der Grammatikschreibung des antiken Griechenlands (zum Begriff außerhalb Europas vgl. Kaltz 2000: 702–704). Als *locus classicus* gilt hier die *Téchnē grammatiké* von Dionysius Thrax (vgl. Uhlig 1883, Kürschner 1996), in der die zuvor von Platon vorgenommene Differenzierung des λόγος (lógos) „Rede, Satz“ in ὄνομα (ónoma) „Name“ und ῥῆμα (rhêma) „Wort, Spruch“ und dessen Weiterentwicklungen durch Aristoteles und die Stoiker als erstes systematisiert wird (vgl. Robins 1966: 6 und passim). Dionysius (Téchnē, §11–§19) unterscheidet für das Griechische seiner Zeit die acht Redeteile ὄνομα (ónoma), i.e. Nomen, ἐπίρρημα (epírrhēma), i.e. Interjektion oder Adverb, ῥῆμα (rhêma), i.e. Verb, μετοχή (metoché), i.e. Partizip, πρόθεσις (próthesis), i.e. Präposition, σύνδεσμος (sýndesmos), i.e. Konjunktion, ἀντωνυμία (antōnymía), i.e. Pronomen und ἄρθρον (árthron), i.e. Artikel, auf Basis semantischer und morphosyntaktischer Kriterien⁴ (vgl. Kürschner 1996: 187–211, Robins 1966: 13–18, Robins 1986: 26). Das Verb wird z.B. als Redeteil definiert, das eine Tätigkeit bezeichnet und unter anderem Tempus-, Modus- und Personmarkierung trägt (vgl. Kürschner 1996: 197–199, Robins 1966: 15) und das Nomen als Redeteil, das ein Ding bezeichnet und unter anderem Kasus, Genus und Numerus markiert (vgl. Kürschner 1996: 189–197, Robins 1966: 14). Diese Wortarteinteilung wurde insbesondere unter dem Einfluss von Donatus’ Schulgrammatiken *Ars minor* (vgl. Schönberger 2008) und

⁴Dionysius unterscheidet nicht genau zwischen morphologischer und syntaktischer Ebene (vgl. Sasse 1993: 189 Anmerkung 2) – daher auch keine terminologische Unterscheidung zwischen Wortarten (vgl. engl. *word classes*) und Redeteilen (*méré lógou*, vgl. lat. *partes orationis* oder engl. *parts of speech*).

Ars maior (vgl. Schönberger 2009) und von Priscians *Institutiones Grammaticae* (vgl. Krehl 1819–1829) auf das Lateinische übertragen (vgl. Robins 1966: 1–5) und schließlich ohne wesentliche Modifikationen als Modell für andere europäische Sprachen etabliert (vgl. Robins 1966, Kaltz 2000), wenn auch im Laufe der Geschichte immer wieder eine kritische Auseinandersetzung erfolgte (vgl. Kaltz 2000: 698–702). Hermann Paul (1995: 352 [¹1880]) gar bescheinigt aufgrund der fehlenden logischen Geschlossenheit der tradierten Wortarteinteilung in den indogermanischen Sprachen den „Charakter der Willkürlichkeit an sich“, um jedoch sogleich hinzuzufügen: „Es würde aber nicht möglich sein etwas wesentlich Besseres an die Stelle zu setzen, so lange man darauf ausgeht, jedes Wort in eine bestimmte Klasse unterzubringen.“ (Hermann Paul 1995: 352 [¹1880]).

Für die Unterbringung eines Wortes in eine bestimmte Klasse haben sich aus der dionysischen Beschreibungstradition jedenfalls folgende Kriterien entwickelt:

1. formales Kriterium:
Zuordnung zu morphologischen Formklassen
2. syntaktisches Kriterium:
Zuordnung zu syntaktischen Kategorien wie Satzglied oder Phrase
3. ontologisch-semantisches Kriterium:
Zuordnung zu Konzeptklassen wie OBJEKT, EREIGNIS und EIGENSCHAFT
4. diskurspragmatisches Kriterium:⁵
Zuordnung zu den Funktionen Referenz, Prädikation und Modifikation

An dieser Stelle sind die Kriterien nicht ausführlich erläutert, da es ausschließlich darum geht, das grundlegende Modell aufzuzeigen.

Mir scheint es von besonderer Bedeutung, drei wesentliche Aspekte dieses Modells hervorzuheben: Erstens handelt es sich bei der so vorgenommenen Wortarteinteilung um eine hybride Kategorisierung in dem Sinne, dass keines der angeführten Kriterien für sich genommen hinreicht, um eine Wortart zu etablieren, sondern eine systematische Korrelation zwischen diesen bestehen soll (vgl. Lyons 1977: 425–430, Sasse 1993: 195–203, Schachter 2007: 1–22). Im

⁵Das diskurspragmatische Kriterium hat sich erst später in die antike Definition eingefügt.

Kern geht es bei der Wortarttheorie also – und ich kann dies nicht besser als in solch allgemeiner Form ausdrücken – um die Prädisposition von Wörtern für den Gebrauch im Satz.

Zweitens kommt der Zuweisung in formale Klassen dabei aber eine zentrale Rolle zu. Denn wenn die Motivation für die Wortarttheorie in dem Versuch begründet liegt, den Wortschatz einer Sprache in Klassen einzuteilen, setzt dies allein in der Diskriminierung von Einheiten des Wortschatzes das Potential für eine Klassenbildung voraus. Das heißt, „[...] dass bei Sprachen mit Flexionsmorphologie zwei Klassifikationsphasen zu unterscheiden sind, wobei durch die erste, die Zusammenfassung von Wortformen zu paradigmatisch strukturierten Lexem-Einheiten, für die zweite, die Klassifizierung von Lexemen nach ihren grammatischen Verwendungspotentialen, ein Rahmen vorgegeben wird.“ (Plank 1984: 495).

Drittens sind solche Klassen dann nicht *a priori* festgelegt, sondern basieren auf einzelsprachlicher Evidenz (vgl. Haspelmath 2007). In der Annahme einer schwachen Äquivalenzhypothese (vgl. Sasse 1993: 190–191) kann aber die prinzipielle Übertragbarkeit auf die indogermanischen Sprachen angenommen werden. Auch das tradierte Wortartmodell bleibt im Detail nicht frei von Problemen – dies betrifft z.B. Fragen nach Subkategorisierung (vgl. Sasse 2001, Simone 2003) oder Mehrfachkategorisierung (vgl. Ross 1972, Dixon und Aikhenvald 2004, Levin 1993), ohne es jedoch *grosso modo* abzulehnen.

Erst in der Auseinandersetzung mit nicht-indogermanischen Sprachen beginnt die eigentliche Diskussion um die Grenzen der Generalisierbarkeit dieses Modells.

Bossong (1992: 6 und *passim*) führt das früheste erhaltene Zeugnis des Zweifels an der Universalität von Wortarten auf das Jahr 1638 zurück, namentlich in der *Grammaticalia* des kalabrischen Philosophen Tommaso Campanella, in der die Übertragung der traditionellen Wortartenlehre auf ost-asiatischen Sprachen infrage gestellt wird. Unter dem Eindruck sich häufender Beschreibungen nicht-indogermanischer Sprachen scheint sich bis ins 19. und frühe 20. Jahrhundert die Annahme durchzusetzen, dass der tradierte Wortartenbegriff nur bei flektierenden Sprachen angesetzt werden kann (z.B. Humboldt 1836, Müller 1861–1864, Boas 1911, Sapir 1921, vgl. auch Haspelmath 2012: 125–126). Seit der zweiten Hälfte des 20. Jahrhunderts nimmt die Diskussion um die Universalität von Wortarten neue Fahrt auf. Dies geschieht zum einen im sich fortführenden Streit um die Analyse verschiedener Einzelsprachen – Tabelle 1 (S. 7) bietet am Beispiel der Verb-Nomen-Distinktion einen Überblick

über prominent diskutierte Sprachen⁶ – zum anderen durch neue sprachwissenschaftliche Theoriebildung wie der generativen Grammatik (vgl. Chomsky 1957, Jackendoff 1983, Baker 2003), der Functional Grammar (vgl. Dik 1978, Hengeveld 1992), der Konstruktionsgrammatik (Goldberg 1995, Croft 2001) oder der Kognitiven Linguistik (vgl. Langacker 1987a, 1991), unter dessen jeweiliger Federführung die Wortarthematik behandelt wird.

Sprachfamilie	Sprache	Verb-Nomen-Unterscheidung	
		ja	nein
Nootka	Wakash	Jacobson (1979)	Sapir (1921) Swadesh (1939)
Irokesisch	Cayuga	Mithun (2000)	Sasse (1993)
Salish	Squamish	van Eijk und Hess (1986) Davies und Matthewson (1999) Beck (2013)	Kuipers (1968) Kinkade (1983)
Austroasiatisch	Mundari	Evans und Osada (2005) Croft (2005)	Hoffmann (1903) Hengeveld und Rijkhoff (2005)
Austronesisch	Tagalog	Aldridge (2009) Richards (2009)	Scheerer (1924) Bloomfield (1917) Capell (1964) Gil (2000) Himmelmann (2008) Kaufmann (2009)
	Samoanisch	Vonen (2000) Dixon (2010)	Mosel und Hovdhaugen (1992)
	Tongaisch	Croft (2001)	Tchekhoff (1984) Hengeveld (1992) Broschart (1997)

Tabelle 1: Vergleich der Verb-Nomen-Unterscheidung in Einzelsprachen

Die in Tabelle 1 vorgenommene Ja-Nein-Gegenüberstellung sollte dann auch als idealisiert betrachtet werden, da den einzelsprachlichen Beschreibungen teils sehr unterschiedliche theoretische Annahmen und Wortartdefinitionen zugrunde liegen. Beispielsweise behandelt Hengeveld (1992) das Tongaische aus funktionsgrammatischer Perspektive, Croft (2001) dagegen im Rahmen seiner *Radical Construction Grammar*, während sich Broschart (1997) neben Henge-

⁶Mir scheint zu nahezu jeder nicht-indogermanischen Sprache eine Arbeit auffindbar, die die tradierte Wortarteinteilung anzweifelt.

veld (1992) an Frege (1892) und Jackendoff (1983) orientiert.

In Bezugnahme auf die tradierte Wortarteinteilung entstehen immer dann Beschreibungsprobleme in den Einzelsprachen, wenn entweder keine Korrelationsbeziehung zwischen den oben genannten Definitionskriterien (vgl. S. 5) existiert, wie dies z.B. Sasse (1993) für das Cayuga annimmt, eine andersartige Korrelationsbeziehung besteht, wie dies z.B. Nordhoff (2004) für das Guarani annimmt, oder mindestens ein Kriterium erst gar nicht analysiert werden kann, wie dies im Hawaiischen der Fall ist.

Nachdem die zu Beginn erläuterte Fragestellung zunächst im größeren Kontext der Wortarttheorie und der daraus folgenden Diskussion um die Universalität von Wortarten betrachtet werden musste, soll im Folgenden der Blick genauer auf die Beschreibungsversuche in den polynesischen Sprachen gerichtet werden.

2.2 Wortarten in der Beschreibung polynesischer Sprachen

Ähnlich dem bis hierhin skizzierten Verlauf in der Wortarttheorie lassen sich die Beschreibungsversuche in den polynesischen Sprachen grob in drei Phasen unterteilen:

1. Frühphase:

Im frühen 19. Jahrhundert nehmen erste (westliche) Beschreibungsversuche ihren Anfang, oftmals in missionarischem Kontext. Die Orientierung an der griechisch-lateinischen Tradition ist unverkennbar. Wortartkategorien sollen entweder uneingeschränkt übertragbar sein oder werden abgelehnt.

2. Adäquate Beschreibung:

Ab Mitte des 20. Jahrhunderts beginnen Bemühungen um eine bessere Beschreibung der polynesischen Sprachen, die sich stärker an den spezifischen Strukturen der jeweiligen Sprache orientieren.

3. Typologische Modelle:

Seit Ende des 20. Jahrhunderts rückt der gegenseitige Einfluss typologisch orientierter Wortarttheorien und einzelsprachlicher Beschreibungen deutlich in den Vordergrund.

In allen drei Phasen werden Wortarten als adäquate Beschreibungsgrößen für die polynesischen Sprachen sowohl angenommen als auch abgelehnt. Zwar un-

terscheidet sich der theoretische Rahmen, in dem die Arbeiten entstanden sind, es lassen sich jedoch über die Phasen hinweg maßgebliche Gemeinsamkeiten finden.

In Tabelle 2 (S. 10) sind die Ergebnisse aus einer Metaanalyse zur Wortartkategorisierung in verschiedenen polynesischen Sprachen in Einzelgrammatiken und theoretischen Arbeiten zusammengefasst.

Zu Beginn sind in Tabelle 2 (S. 10) vier Arbeiten gelistet, die explizit Aussagen über alle polynesischen Sprachen machen. Bei der Auswahl der Einzelsprachen wurde darauf geachtet, dass mehrere und in allen drei oben beschriebenen Phasen (vgl. S. 8) entstandene Arbeiten vorhanden sind, um etwaige Ähnlichkeiten wie Divergenzen in den Beschreibungen einer einzelnen Sprache zu untersuchen. Insgesamt finden sich über die Arbeiten hinweg fünf Ebenen der Analyse. Das sind die morphologische, semantische, distributionale, syntaktische und pragmatische Ebene. Unterschiede zeigen sich darin, ob auf einer der Ebenen Klassen differenziert werden können (in Tabelle 2 mit Eintrag +) oder nicht (in Tabelle 2 mit Eintrag –), wobei in manchen Arbeiten einzelne Ebenen keine Berücksichtigung finden (in Tabelle 2 mit Eintrag 0) oder ihnen ein besonderer Status zugeschrieben wird (in Tabelle 2 mit Eintrag *). Zunächst einmal geht es darum, ob überhaupt morphologische Klassen, semantische Klassen, syntaktische Slots oder diskurspragmatische Funktionen analysiert werden können (vgl. die Wortartkriterien S. 5). Allein die Unterscheidung in morphologische Klassen wird mehrheitlich abgelehnt, obwohl vereinzelt mittels Konversions- oder Homonymieanalyse Wörter auf lexikalischer Ebene unterscheidbar sein sollen (z.B. Vonen 2000 oder Dixon 2010). Die Differenzierung von semantischen Klassen lässt sich in allen Arbeiten mehr oder weniger deutlich erkennen. Mit der Unterscheidung von syntaktischen Slots und – falls berücksichtigt – diskurspragmatischen Funktionen wird in manchen Arbeiten implizit oder explizit Wörtern im aktuellen Auftreten im Satz ein eigenständiger kategorialer Status zugeschrieben (z.B. von Chamisso 1837 oder Mosel und Hovdhaugen 1992).

Für die Ansetzung von Wortartkategorien aber spielt mehrheitlich die Distribution die ausschlaggebende Rolle (z.B. Hengeveld 1992 oder Lazard und Peltzer 2000). Entscheidend ist dabei die Frage, ob alle Wörter auf alle syntaktischen Slots oder pragmatische Funktionen verteilt sind und ob beispielsweise ein Zusammenhang zu semantischen Klassen angenommen werden muss (z.B. Elbert und Pukui 1979 oder Völkel 2017).

	morphol.	semant.	distrib.	syntakt.	pragm.
Polynesisch					
Buschmann (1839)	–	+	–	*	0
Steinthäl (1860)	–	+	–	+	0
Schleicher (1865)	–	+	–	+	0
Vonen (2000)	*	+	+	+	0
Hawaiisch					
Chamisso (1837)	–	+	–	*	0
Andrews (1854)	–	+	+	+	0
Elbert und Pukui (1979)	–	+	+	+	0
Samoaanisch					
Pratt (1893)	–	+	+	+	0
Pawley (1966)	–	+	+	+	0
Mosel und Hovdhaugen (1992)	–	+	–	*	*
Hengeveld (1992)	–	0	–	+	+
Vonen (2000)	*	+	+	+	0
Dixon (2010)	*	+	+	+	+
Tongaisch					
Churchward (1953)	–	+	+	+	0
Tchekhoff (1981)	–	+	–	+	0
Broschart (1997)	–	*	–	*	*
Croft (2001)	–	+	+	+	+
Völkel (2017)	–	+	+	+	+
Maorisch					
Williams (1862)	–	+	+	+	0
Schleicher (1865)	–	+	–	+	0
Biggs (1960, 1961, 1971)	–	+	+	+	0
Bauer (2001)	–	+	+	+	0
Tahitisch					
Davies (1821, 1851)	–	+	+	+	0
Tryon (1970)	–	+	+	+	0
Lazard und Peltzer (2000)	–	+	+	+	0
Marquesisch					
Drodillion (1904)	–	+	+	+	0
Cablitz (2006)	–	+	+	+	0
Rapanui					
DuFeu (1996)	–	+	–	+	0
Kiviet (2017)	–	+	+	+	+

Tabelle 2: Metaanalyse Wortarten in polynesischen Sprachen

Im Wesentlichen lassen sich vier grundlegende Ansätze herausarbeiten, mit Wortarten in den polynesischen Sprachen umzugehen:

1. Übertragung des tradierten Wortartenmodells
2. Annahme von Wortarten als syntaktische Kategorien unabhängig von lexikalischen Kategorien
3. Ablehnung von Wortartkategorien
4. Prototypische Abdeckung semantischer Klassen auf syntaktisch-funktionale Slots

Diese Ansätze sollen nachfolgend genauer betrachtet werden.

2.2.1 Übertragbarkeit der tradierten Wortartkategorien

Insbesondere in Grammatiken aus der Frühphase der Beschreibung der polynesischen Sprachen im 19. Jahrhundert (vgl. S. 8) findet eine unreflektierte Anwendung der Beschreibungsbegriffe aus der griechisch-lateinischen Tradition statt, mit der der Versuch einer isomorphen Übertragung von Wortartkategorien einhergeht. Für das Hawaiische beispielsweise sei Andrews (1854) genannt, der unter anderem schreibt: „Hawaiian Nouns have *Person, Number, Gender, and Case*.“ (Andrews 1854: 33 §81). Andrews (1854: 43–55) spricht dann z.B. trotz nicht unterscheidbarer Formklassen im Hawaiischen in Analogie zur Beschreibung indogermanischer Sprachen von der Deklination von Nomen oder setzt die Unterscheidung der grammatischen Kategorien von femininem und maskulinem Genus mit der Unterscheidung des Sexus im Hawaiischen durch Adposition der Wörter *kane* „Mann, männlich“ oder *wahine* „Frau, weiblich“ wie in *makua kane* „männliches Elternteil, Vater“ gleich (vgl. Andrews 1854: 41–42 § 95).⁷ An solchen Beispielen wird deutlich, dass viele Grammatiken aus der Frühphase der Beschreibung mehr dem Diktat der lateinischen Kategorien folgen als spezifische einzelsprachliche Strukturen zu beachten. So

⁷Schütz (1994: 241–288) bietet einen guten Überblick über hawaiische Grammatiken und die einzelnen Versuche der Wortartübertragung.

schreibt Mutu (1989: 399): „[...] prior to 1960, all published grammars of Polynesian languages were ‘traditional’ grammars, sometimes referred to as Latinate.“ Mutu (1989: 399) bezieht sich hier auf die Arbeiten von Biggs (1960, 1961, 1969, 1971) zum Maorischen. Zwar versucht Biggs (1960, 1961, 1969, 1971) weitgehend auf die Terminologie der lateinisch-griechischen Tradition zu verzichten, legt aber zugleich den Grundstein für das, was man als eine Art Standard-Wortartdefinition aus den verschiedenen Grammatiken zu einzelnen polynesischen Sprachen, die seit Mitte des 20. Jahrhunderts entstanden sind (vgl. S. 8), extrahieren kann. Als Ausgangspunkt dient hier die Unterscheidung von Tempus-Aspekt-Modus-Syntagma wie [*i hana ana*] in Satz (1a) (S. 1) und Artikel-Syntagma wie [*ka hana*] in Satz (1b) (S. 1), wobei diese in vielen Grammatiken einfach als *Verbal-* und *Nominalphrasen* und entsprechend Tempus-Aspekt-Modus-Marker als *Verbalmarker* und Artikelwörter als *Nominalmarker* bezeichnet werden (vgl. z.B. Elbert und Pukui 1979: 41 und passim). *Verben* seien demzufolge Wörter, die in Tempus-Aspekt-Modus-Syntagma auftreten und *Nomen* Wörter, die in Artikel-Syntagma auftreten (vgl. z.B. Elbert und Pukui 1979: 43 und passim). Wörter, die sowohl in Tempus-Aspekt-Modus-Syntagma als auch Artikel-Syntagma auftreten, werden dann entweder als eigenständige Klasse behandelt, etwa als Klasse der „universals“ (Biggs 1971: 11, vgl. auch Pawley 1966 zum Samoanischen) oder als eine Wortart, die als andere verwendet wird. Elbert und Pukui (1979: 43 und passim) sprechen im Hawaiischen etwa von „noun-verbs“. Das sollen Verben sein, die als Nomen verwendet werden (vgl. auch Churchward 1953 zum Tongaischen oder Tryon 1973 zum Tahitischen).

Entscheidend ist hier, dass eine Distributionshypothese aufgestellt wird, die besagt, dass nicht alle Wörter in allen Syntagmen auftreten und – wenn auch nicht immer explizit ausgearbeitet – eine Prädispositionshypothese, die besagt, dass die Distribution durch die lexikalisierte Wortbedeutung motiviert ist, vgl. z.B. Elbert und Pukui (1979: 43) zum Hawaiischen: „Nouns [denote] names of persons, places, or things; potential occurrence after the articles [...]“.

Die Parallelität zur dionysischen Wortartdefinition ist hier offenkundig: Nomen sind Wörter, die ein Ding bezeichnen und prädisponiert sind für das Auftreten in Artikel-Syntagma, Verben sind Wörter, die eine Tätigkeit bezeichnen und prädisponiert sind für das Auftreten in Tempus-Aspekt-Modus-Syntagma. Damit wird letztlich wie schon in Arbeiten aus der Frühphase der Beschreibung der polynesischen Sprachen die Unterscheidung von morphologischen Paradigmen mit dem Auftreten in Tempus-Aspekt-Modus- und Artikel-Syntagma

gleichgesetzt. Die theoretische Behandlung der Frage nach der Klassifizierung von Wörtern, die sowohl in Tempus-Aspekt-Modus-Syntagma als auch Artikel-Syntagma auftreten, endet dann einfach darin, sie distributional als *universals* oder semantisch in Analogie zu anderen Sprachen als *Verben* oder *Nomen* zu analysieren, vgl. Tryon (1970: 73) zum Tahitischen: „[...] verbs expressing movement or state may become abstract nouns when preceded by the article te.“ Die Distributionshypothese wird in späteren Grammatiken immer wieder aufgenommen (vgl. z.B. Lazard und Peltzer 2000 zum Tahitischen). Dixon (2010) sieht in seiner *Basic Linguistic Theory* darin sogar den Beweis für die Universalität von Nomen und Verben, da davon ausgegangen werden könne, dass auch in Sprachen wie dem Samoanischen nur wenige Lexeme tatsächlich sowohl in Nominal- als auch Verbalphrase auftreten (vgl. Dixon 2010: 45). Solche Fälle könnten dann entweder als graduelle Überschneidung von Nomen- und Verbkategorie oder als Homonymie behandelt werden (vgl. Dixon 2010: 45). Dass Wörter, die in mehr als einem syntaktischen Slot auftreten, die Ausnahme darstellen, bleibt bei Dixon (2010) allerdings bloße Vermutung, und zudem ist fraglich, ob in Fällen wie *hana* in Satz (1a) und Satz (1b) (S. 1) zwei homonyme Wörter angesetzt werden können oder wie die Nomen- und Verbkategorie definiert sind, die sich darin überschneiden.

Unabhängig davon, wie häufig solche Fälle sind, wird in wenigen, späteren Arbeiten mittels Konversionsanalyse, das heißt der kategorialen Unterscheidung zweier Wortformen trotz formaler Identität, versucht, die Übertragbarkeit der tradierten Wortartkategorien beizubehalten. So ließe sich Vonen (2000) und in Teilen Don und Van Lier (2013) zufolge die formale Identität von Wörtern wie *hana* in den Sätzen (1a) und (1b) (S. 1) als kategorieverändernde Ableitung analysieren, mit einem verbalen Lexem *hana*₁ wie im Deutschen *arbeiten* in Satz (1a) und einem nominalen Lexem *hana*₂ wie im Deutschen *Arbeit* in Satz (1b) (S. 1).

Die größte Unklarheit besteht hier in der Analyse der Derivationsrichtung ohne formale Markierung (vgl. auch Sasse 2011: 175). Für Vonen (2000: 484) stellt sich dies allein als Frage dar, wie ambitioniert an die Beschreibung der einzelnen Sprache herangegangen wird. Prinzipiell wäre es möglich, in einer genauen semantisch-lexikalischen Analyse auch Regeln der Konversion aufzuzeigen. Don und Van Lier (2013) versuchen dazu, lexikalische von syntaktischer Ableitung zu unterscheiden. Welche Probleme bei der semantischen Analyse und in der Folge der Unterscheidung von Konversion, Vagheit und Polysemie auftreten, ist im Detail Gegenstand von Kapitel 2.2.4 (S. 25). Ob mit einer Konversionsana-

lyse selbst mehr getan ist, als der Versuch, *ad hoc* tradierte Wortartkategorien zu retten, weist Vonen (2000) jedenfalls nicht nach. So wurde bereits darauf hingewiesen (vgl. Kapitel 1.1, S. 3), dass die Annahme von Konversion ohne unabhängige Möglichkeit der Klassenbildung nicht plausibel erscheint. Für die Analyse von kategorienverändernder Derivation müsste zunächst geklärt sein, wie die Wortartkategorien, die sich verändern, in den polynesischen Sprachen definiert sind. Vonen (2000) setzt ähnlich wie Andrews (1854) oder Dixon (2010) letztlich syntaktisch oder semantisch äquivalente Kategorien voraus, die dann auch lexikalisch gegeben sein sollen.

Zusammenfassend besteht bei all den bis hierhin besprochenen Ansätzen die Gefahr, einfach die Namen der Kategorien aus den bekannten Sprachen für die Übersetzungsäquivalente aus der anderen zu verwenden, ohne die Möglichkeit der Konstituierung der Kategorien aus der Beschreibung der einzelnen Sprache heraus zu berücksichtigen, vgl. Broschart (1997: 160 Anmerkung 2): „Most linguists simply take a practical position implying that whatever translates as a noun in Indo-European will be called a noun.“

2.2.2 Wortarten als syntaktische Kategorien

Geht man anders als Dixon (2010) davon aus, dass prinzipiell alle Wörter in allen Syntagmen auftreten können, wie dies Mosel und Hovdhaugen (1992) im Samoanischen oder Broschart (1997) im Tongaischen tun, verzichtet aber darauf, wie Vonen (2000) oder Don und Van Lier (2013), bei formaler Identität kategorial verschiedene Einheiten des Lexikons anzusetzen, bietet sich als weitere Möglichkeit, Wortarten als rein syntaktische Kategorien aufzufassen. Das deutet sich schon bei von Chamisso (1837: 8) an, der im Hawaiischen von der Geltung von Wörtern als Nomen oder Verb im Satz spricht, ohne dies weiter zu präzisieren, oder bei Buschmann (1839: 842 und *passim*), der Wörtern in den polynesischen Sprachen auch ohne Einteilung in Formklassen die Eigenschaft von Substantiven, Adjektiven und Verben zuspricht. In ähnlicher Weise schreiben Mosel und Hovdhaugen (1992: 77) zur formalen Identität von Wörtern im Samoanischen:

This does not mean that a noun can be used as a verb or a verb as a noun or that we have two homophonous words [...]. Rather it means that in Samoan the categorisation of words into nouns and verbs is not given a priori in the lexicon. It is only their actual occurrence in a particular syntactic environment which gives them the status of a verb or noun.
(Mosel und Hovdhaugen 1992: 77)

Unklar ist hier, was genau unter der Geltung, der Eigenschaft oder dem Status eines Nomens oder Verbs zu verstehen ist. Zugrunde mag hier ähnlich wie in den im vorangegangenen Kapitel besprochenen Ansätzen (vgl. S. 11–14) der Vergleich liegen, dass manche der grammatischen Kategorien, die in einer Sprache mit Nomen-Verb-Unterscheidung an der entsprechenden Wortart, etwa Tempus an Verben, in den polynesischen Sprachen im Satz mit entsprechenden Partikeln ausgedrückt werden, vgl. Völkel (2017: 447): „Crosslinguistically, TAM is a typical verbal category and definiteness (as expressed by articles) is a typical nominal one.“ Wenn aber Nomen und Verb im Lexikon nicht geschieden sein sollen und es sich hier nicht bloß um terminologische Ungenauigkeit in der Übertragung von Kategoriennamen handelt (vgl. Kapitel 2.2.1, S. 11–14), sondern um die Postulierung von Nomen und Verb als syntaktische Kategorien, sehe ich erstens mit Sasse (1993: 199) ein methodologisches Problem, weil durch die Definition einer Wortart *Nomen* erst als Bestandteil von *Nominal*-Phrase und einer Wortart *Verb* erst als Bestandteil von *Verbal*-Phrase die lexikalische und syntaktische Ebene vermengt werden. Eine *NP* setzt die Kategorie *N* voraus, eine *VP* ebenso die Kategorie *V* und es kann sozusagen nicht rückwirkend *N* nur in *NP* und *V* nur in *VP* angesetzt werden – siehe dazu auch die Definition von *Phrasen* bei Bloomfield (1973: 196 [1933]): „The syntactic form-classes of phrases, therefore, can be derived from the syntactic form-classes of words: the form-classes of syntax are most easily described in terms of *word-classes*.“ Zweitens kann im Samoanischen, Tongaischen oder Hawaiischen nun nicht einmal von *NP* oder *VP* gesprochen werden, wenn *N* und *V* lexikalisch nicht definiert sind. Zwar lassen sich in diesen Sprachen Tempus-Aspekt-Modus-Syntagma wie [*e hana ana*] in (1a) und Artikel-Syntagma wie [*ka hana*] in (1b) (S. 1) in klarer Weise als strukturelle Einheiten voneinander unterscheiden, da sie komplementär zueinander verteilt sind, das heißt, Artikelwörter und Tempus-Aspekt-Modus-Marker niemals miteinander kombiniert werden (vgl. Mosel und Hovdhaugen 1992: 50 und passim zum Samoanischen, Broschart 1997: 133 und Völkel 2017: 452 zum Tongaischen, Elbert und Pukui 1979: 41 und passim zum Hawaiischen). Versteht man allerdings mit Bloomfield (1973: 194–196 [1933]) *Nominal*- und *Verbalphrase* als endozentrische Phrase, das heißt als Phrase, die derselben Kategorie angehört wie ihr Zentrum oder Kopf, muss man in Sprachen wie dem Hawaiischen eben von einer *Artikel*- oder *Definitheitsphrase* mit Artikel bzw. Definitheitsmarker statt kategorisiertem Inhaltswort als Kopf sprechen und entsprechend einer *Tempus-Aspekt-Modus-Phrase* mit Tempus-Aspekt-Modus-

Marker als Kopf. Im Hawaiischen wird ein Wort wie *hana* letztlich nicht dadurch zu einem Nomen, indem es in Artikel-Phrase auftritt oder dadurch zu einem Verb, indem es in Tempus-Aspekt-Modus-Phrase auftritt, auch nicht, wenn Ähnlichkeiten zwischen Artikel-Phrase und Nominalphrase und Tempus-Aspekt-Modus-Phrase und Verbalphrase feststellbar sind.

Bei der Annahme von Wortarten als reine syntaktische Kategorien sehe ich dann keinen großen Unterschied mehr zu den Versuchen, das tradierte Wortartmodell auf die polynesischen Sprachen zu übertragen (vgl. Kap. 2.2.1, S. 11–14).

2.2.3 Ablehnung von Wortartkategorien

In der Frühphase der Beschreibung der polynesischen Sprachen (vgl. Kapitel 2.2, S. 8) finden sich neben Schleicher (1865), der die Unterscheidung von Wortarten konsequent ablehnt, wenn überhaupt, Arbeiten wie die von Buschmann (1839) und von Chamisso (1837), die zumindest eine uneingeschränkte Übertragung der tradierten Wortartkategorien anzweifeln – uneingeschränkt insofern, als Buschmann (1839) und von Chamisso (1837) zwar die fehlende Unterscheidbarkeit von Formklassen berücksichtigen, aber die Analyse von Wortarten als syntaktische Kategorien offenlassen und in der Folge weiterhin die tradierten Wortartbegriffe verwenden (vgl. Kapitel 2.2.2, S. 14–16).

Erst später in der Phase typologischer Modelle (vgl. Kapitel 2.2, S. 8) folgen Ansätze, die eine Wortartkategorisierung in den polynesischen Sprachen grundsätzlich ablehnen. Dies geschieht allerdings aus einer anderen theoretischen Perspektive. Zu Beginn steht hier Hengeveld (1992) mit einer ganzen Reihe darauf aufbauender Arbeiten (z.B. Broschart 1997, Rijkhoff 2002, Hengeveld, Rijkhoff und Sierwierska 2004, Rijkhoff und Van Lier 2013 und Van Lier 2016), auf die in der Diskussion um Wortarten in den polynesischen Sprachen immer wieder Bezug genommen wird (vgl. z.B. Croft 2001, Völkel 2017). Im Folgenden soll der von Hengeveld (1992) entwickelte Ansatz genauer betrachtet werden.

Der Ansatz von Hengeveld (1992)

Die ursprüngliche auf der *Functional Grammar* (vgl. Dik 1989) basierende Wortartdefinition von Hengeveld (1992: 58) lautet:

A *verbal* predicate is a predicate which, without further measures being taken, has a predicative use *only*.

A *nominal* predicate is a predicate which, without further measures being taken, can be used as the head of a term.

An *adjectival* predicate is a predicate which, without further measures being taken, can be used as modifier of a nominal head.

An *adverbial* predicate is a predicate which, without further measures being taken, can be used as modifier of a non-nominal head.

Hengeveld (1992: 58) bedient sich in dieser Definition terminologisch der Prädikatenlogik mit der Unterscheidung von *Prädikat* (auch *genereller Term*) und *Individuenkonstante* (auch *singulärer Term*) bzw. *Individuenvariable* (auch *genereller Term mit Quantor*). In der Prädikatenlogik spielen Wortarten als Beschreibungsgrößen von natürlichen Sprachen keine Rolle. Um Beispiele aus den Sätzen (1a) und (1b) (S. 1) zu nehmen: Sowohl *Mann*, *männlich* als auch *arbeiten* sind generelle Terme. Eigennamen wie *Herman* sind singuläre Terme. Formalisieren kann man *Herman ist ein Mann* ebenso wie *Herman ist männlich* mit $M(b)$, *Herman arbeitet* mit $G(b)$ und *ein Mann arbeitet* mit Existenzquantor $\exists x [M(x) \ \& \ G(x)]$ gelesen als „Es gibt ein x für das gilt: x ist ein Mann und x arbeitet“. Hengeveld (1992) versteht zwar wie in der Logik *Mann*, *männlich* und *arbeiten* als Prädikate. Nur kann man – das ist der Übergang von der Logik zur Beschreibung natürlicher Sprachen – unterscheiden zwischen nominalen, verbalen, adjektivischen und adverbialen Prädikaten. Demzufolge wäre im Englischen *intelligent* im Vergleich zu *sing* ein „adjectival predicate“, da *sing* als Modifizierer eines nominalen Kopfes wie in *the singing detective* oder *the detective who is singing* nur mit „further measures“ auftritt, *intelligent* dagegen ohne wie in *the intelligent detective* (vgl. Hengeveld (1992: 58)). In polynesischen Sprachen wie dem Tongaischen (vgl. Hengeveld 1992: 66 und passim), dem Samoanischen (vgl. Rijkhoff 2002: 125, Hengeveld, Rijkhoff und Sierwierska 2004: 534–537) oder Hawaiischen wäre im Gegensatz zum Englischen keine Unterscheidung zwischen den Prädikatsklassen bzw. Wortarten möglich, da keine „further measures“ dafür zur Verfügung stehen.

Unklar ist hier, was genau unter „measures“ zu verstehen ist. Rijkhoff und Van Lier (2013: 8) haben später versucht, „further measures“ mit „any additional function-indicating morphosyntactic devices“ zu reformulieren, auch wenn bei Hengeveld (1992) „further measures“ nicht ausschließlich Flexionsmorphologie umfassen. So sieht Hengeveld (1992: 58) die Relativisierung mit *who* in *the detective who is singing* als „further measure“ von *sing* an. Im Kern möchte Hengeveld damit für einen übereinzelsprachlichen Vergleich die

Zuordnung von Wortformen zu Flexionsparadigmen überwinden, vgl.: „In a morpho-syntactic approach the differences between predicates are defined in terms of the morphological categories for which they may be specified. This approach has the disadvantage of not allowing for crosslinguistic generalization.“ (Hengeveld 1992: 50). Warum aber z.B. in einem Satz wie *the detective who is intelligent* die Relativisierung mit *who* nicht als „further measure“ von *intelligent* gelten soll und wie mit *singing* als Partizip in *the singing detective* umgegangen wird, bleibt dann ungeklärt.

Genau genommen kann an der Gegenüberstellung von *sing* und *intelligent*, so wie sie Hengeveld (1992: 58) vornimmt, gar nicht gezeigt werden, dass *sing* nach der vorangegangenen Definition (S. 17) kein „adjectival predicate“ ist. Der eigentliche Vergleich müsste innerhalb der Auftretensweisen des Lexems *sing* stattfinden – und zwar im Sinne von *Lexem* als Abstraktion von Wortformen. Anders ausgedrückt setzen „further measures“ oder „additional function-indicating morphosyntactic devices“ überhaupt „measures“ oder „function-indicating morphosyntactic devices“ voraus, welche in einer flektierenden Sprache durch die Flexionsmorphologie gegeben sind und *sing* schon in der morphosyntaktischen Analyse als verbales Lexem definieren oder Formulierungen wie „modifier of a nominal head“ (Hengeveld 1992: 58) erlauben.

Lässt man den prädikatenlogischen Vorbau und die pseudo-logische Formalisierung im Rahmen der *Functional Grammar* beiseite (vgl. dazu auch Dik 1989 oder Siewierska 1991), verkürzt sich der rein logisch-semantische Prädikatsbegriff hier letztlich auf den Wort- oder Lexembegriff und die prädikatenlogischen Begriffe *Referenz* und *Prädikation* auf die diskurspragmatischen Funktionen Referenz und Prädikation (vgl. Hengeveld 1992: 48–49), wobei sich all diese nicht auf derselben Analyseebene befinden.

So haben Hengeveld, Rijkhoff und Siewierska (2004: 530) auch später die ursprüngliche Definition von Hengeveld (1992) folgendermaßen neu formuliert:

a VERB (V) is a lexeme that can be used as the head of a predicate phrase only;

a NOUN (N) is a lexeme that can be used as the head of a referential phrase;

an ADJEKTIVE (A) is a lexeme that can be used as a modifier within a referential phrase;

and a MANNER ADVERB (MAdv) is a lexeme that can be used as a modifier within a predicate phrase.

Unter *Lexem* muss hier zunächst ganz allgemein „lexical item“ (vgl. Hengeveld, Rijkhoff und Siewierska 2004: 527) oder „content word“ (Rijkhoff

2002: 16) verstanden werden. Die Ausdrücke „predicate phrase“ und „referential phrase“ beziehen sich einerseits auf die diskurspragmatischen Funktionen Prädikation und Referenz (vgl. Searle 1969: 23–24, Croft 2001: 66), andererseits auf die syntaktischen Einheiten oder Phrasen, mit der Referenz oder Prädikation hergestellt wird.

Völkel (2017: 448 Anmerkung 4) hat Hengeveld (1992) und Hengeveld, Rijkhoff und Siewierska (2004) für die Gleichsetzung von diskurspragmatisch definierten Funktionen und syntaktisch definierten Phrasen kritisiert, vgl.: „While predication and reference are pragmatic units, phrases describe syntactic ones. From a terminological perspective, it is disputable whether those two criteria may be mixed.“ Man muss also z.B. unterscheiden zwischen der diskurspragmatischen Funktion Referenz und den sprachlichen Ausdrücken, mit denen ein Sprecher im Diskurs auf etwas referieren kann (vgl. Searle 1969: 114 und passim, Lyons 1977: 177).

Problematisch ist dies für Hengeveld (1992) und Hengeveld, Rijkhoff und Siewierska (2004) insofern, als ihrer Theorie eine asymmetrische Definition von referentieller und prädikativer Phrase zugrunde liegt. So werden referentielle Phrase und Nominalphrase gleichgesetzt (vgl. z.B. Hengeveld 1992: 58 und passim, Hengeveld, Rijkhoff und Siewierska 2004: 530), nicht aber prädikative Phrase und Verbalphrase. Da Hengevelds anfängliche Motivation in der Beschreibung von „non-verbal predication“ (Hengeveld 1992: 1) liegt, möchte er letztlich zeigen, dass Nomen nicht nur in referentieller Phrase, sondern auch in prädikativer Phrase ohne „further measures“ auftreten können, z.B. in Kopulasätzen wie *John is my best friend* (vgl. Hengeveld 1992: 80) – die Kopula analysiert Hengeveld (1992) nicht als Teil der Prädikation – und dass Verben im Gegensatz dazu nur in prädikativer Phrase ohne „further measures“ auftreten können – daher auch der Zusatz „only“ in der Definition von Hengeveld (1992: 58).

Wenn aber mit der Nominalphrase [*my best friend*]_{NP} John in *John is my best friend* prädiziert wird, können referentielle Phrase und Nominalphrase nicht identisch sein. Sonst würde man ebenso referentielle Phrase und prädikative Phrase gleichstellen. Unterscheidet man hingegen die Ebene der syntaktischen Einheit Phrase und der diskurspragmatischen Funktion, bedeutet das, Nomen sind zwar für das Auftreten in Nominalphrase prädisponiert, aber ein Sprecher kann im Diskurs mit einer Nominalphrase nicht nur auf etwas referieren sondern ähnlich wie mit einer Verbalphrase etwas prädizieren, mit einer Ver-

balphrase aber nicht auf etwas referieren.⁸

In ähnlicher Weise kann man nun auch in Sprachen wie dem Hawaiischen unterscheiden: Artikel-Phrasen können wie Nominalphrasen sowohl referentiell als auch prädikativ wie [*ka hana*] in Satz (1b) (S. 1) gebraucht werden, Tempus-Aspekt-Modus-Phrasen wie Verbalphrasen jedoch nicht, unabhängig von der Unterscheidung zwischen Nomen und Verb.

Worauf ich hier hinaus will, ist folgendes: Hengeveld (1992) kann nicht einfach diskurspragmatische Funktionen als definitorenischen Ausgangspunkt für eine universelle Wortarttheorie setzen. Vielmehr muss er, wenn er über „non-verbal predication“ spricht, das „verbal“ also Verben und Nomen als lexikalische Kategorien voraussetzen. Er mag zeigen, dass Nominalphrasen nicht nur, aber eher referentiell, Verbalphrasen nur prädikativ gebraucht werden und damit Nomen eher referenzfähig, Verben prädikationsgeeignet sind. Wenn aber nichts anderes gesagt wird, als dass manche Sprachen im Gegensatz zum Hawaiischen separate Lexemklassen haben, die formal unterscheidbar und prädisponiert für den Gebrauch in syntaktischen Slots mit referentieller oder prädikativer Funktion sind (vgl. Hengeveld 2013: 32), sehe ich keinen wesentlichen Unterschied zur traditionellen Wortartlehre, in der ja die gleichen Wortartkategorien angesetzt werden und die gleichen Probleme in der Beschreibung des Hawaiischen auftreten.

Hengeveld (1992: 51) sieht nun aber in diskurspragmatischen Funktionen ein besseres sprachübergreifendes *tertium comparationis* als in den tradierten Wortartkriterien, vgl.: „The functions to which reference is made in these definitions may be assumed to be universally recognizable, which makes this approach particularly useful for typological research.“ Die zugrunde liegende Vermutung ist hier, dass es sich bei diskurspragmatischen Funktionen um Universalien handelt, da jede Sprache zwischen diesen für die Kommunikation wesentlichen Funktionen unterscheiden muss (vgl. auch Hopper und Thompson 1984).

Auch wenn man diese Vermutung mit Hengeveld (1992) teilt und diskurspragmatische Funktionen als Universalien in der menschlichen Kommunikation ansieht, stellen sie damit noch kein besseres *tertium comparationis* für Wortartkategorien dar. Es ist etwas anderes, Sprachen darin zu vergleichen, wie diskurspragmatische Funktionen kodiert werden oder Sprachen darin zu vergleichen, ob sie Wortarten haben. Diskurspragmatische Funktionen sind nicht die

⁸Genau genommen zeigt das Beispiel *John is my best friend*, dass sich im Unterschied zu morphosyntaktischen Kategorien Verb und Nomen die Funktionen Referenz und Prädikation nicht gegenseitig ausschließen, weil das Prädikat *my best friend* auch referentiell ist.

alleinigen Konstitutionskriterien für Wortartkategorien. Der Kern von Wortartkategorien ist ja gerade ihr hybrider Charakter, das heißt, dass formal und semantisch differenzierbare Klassen prädisponiert sind für den Gebrauch in syntaktischem Gebrauch und pragmatischen Funktionen. Die Frage der Zuordnung von Lexemklassen zu pragmatischen Funktionen enthält die ausdrucksseitige Differenzierung der tradierten Wortartunterscheidung und setzt die Unterscheidung von Lexemklassen, die selbst nicht über die pragmatischen Funktionen definiert werden können, voraus. Hengeveld (1992) spricht also entweder über Wortarten – dann muss er wie in der tradierten Wortarttheorie gleichermaßen formale Kriterien berücksichtigen – oder er spricht universell über diskurspragmatische Funktionen, die aber nichts mit Wortarten zu tun haben müssen.

Hengeveld (1992: 62–71) versucht dann aufbauend auf seiner Definition Sprachen typologisch danach zu klassifizieren, ob Lexeme für den Gebrauch in syntaktisch-funktionalen Phrasen prädisponiert sind, und zwar als Sprachen mit flexiblem, differenzierendem oder rigidem „Part-of-Speech-System“ (Hengeveld 1992: 62, vgl. auch Rijkhoff 2002, Rijkhoff 2003, Hengeveld, Rijkhoff und Siewierska 2004, Hengeveld 2013, Van Lier und Rijkhoff 2013, Van Lier 2016, Luuk 2010).

Das Englische wäre demnach eine Sprache mit differenzierendem Part-of-Speech-System, in dem eine Eins-zu-Eins-Zuordnung von Lexemklasse und Funktion gegeben ist. Das Hawaiische müsste in Analogie zu anderen polynesischen Sprachen wie dem Tongaischen (vgl. Hengeveld 1992: 66) als Sprache mit flexiblen Part-of-Speech-System klassifiziert werden, in dem wie bei *hana* in den Sätzen (1a) und (1b) (S. 1) keine Zuordnung von Lexemklasse und Funktion möglich ist. In rigiden Sprachen wie dem Quechua würden wenige Lexemklassen für mehrere Funktionen gebraucht (vgl. Hengeveld 1992: 64). Unklar ist dabei allerdings, ob immer Sprachen als Ganzes so klassifiziert werden müssten (vgl. z.B. Barner und Bale 2002 zum Englischen).

Die so definierten Part-of-Speech-Systeme hält Hengeveld (2013: 31) für eine „basic typological determinant“, weil sie mit bestimmten Eigenschaften einer Sprache wie der Wortstellung (vgl. auch Hengeveld, Rijkhoff und Siewierska 2004, Hengeveld 2013: 50) korrelieren. Die Gültigkeit der postulierten Korrelation zwischen Part-of-Speech-Systemen und anderen Beschreibungsdimensionen kann hier für die einzelnen Sprachen nicht beurteilt werden, ist aber auch nicht von besonderer Relevanz für die Wortarttheorie. Denn Hengeveld (1992) definiert mit seinen Part-of-Speech-Systemen überhaupt keine Wortartsysteme.

me. Wenn einer Sprache wie dem Hawaiischen im Gegensatz zu einer Sprache wie dem Englischen die für Wortartkategorien konstituierenden Merkmale fehlen, hat sie keine Wortarten und gehört damit auch keinem Wortartsystem an. Nicht existierende Kategorien können nicht mit existierenden Kategorien verglichen werden. Wenn die Unterscheidung von Part-of-Speech-Systemen mehr sein soll als die Unterscheidung, ob eine Sprache Wortarten hat oder nicht, müsste Parts-of-Speech als syntaktische Kategorie ein eigenständiger kategorieller Status zugesprochen werden, wobei dies, wie bereits besprochen, nicht problemlos möglich ist (vgl. Kap. 2.2.2, S. 14–16). Ich halte es geringstenfalls für terminologisch irreführend, von flexiblen Parts-of-Speech oder gar flexiblen Wortarten⁹ (vgl. z.B. Van Lier und Rijkhoff 2013) oder auch „contentives“ (Van Lier und Rijkhoff 2013: 4, vgl. auch Hengeveld 2013: 36) als eigenständiger Wortartkategorie zu sprechen.

Zusammenfassend kann man fragen, ob die Hinwendung auf pragmatische Funktionen als konstituierendes Kriterium tatsächlich einen wesentlichen Beitrag zur Wortarttheorie liefert. Zumindest für die Beschreibung des Hawaiischen mag man damit nicht über die schon von Chamisso (1837) gemachten Feststellungen hinauszukommen. Und Hengeveld (1992) selbst zieht ja den Schluss, dass es keine Wortarten in den polynesischen Sprachen gibt. Nur braucht man dafür seine ganze Theorie nicht.

Der Ansatz von Broschart (1997)

Während Hengeveld (1992: 63–67) das Tongaische als Sprache mit flexiblem Part-of-Speech-System klassifiziert, versucht Broschart (1997) unter Bezugnahme auf Hengevelds (1992: 58) Definition in dem viel zitierten Artikel *Why Tongan does it differently* eine andere Art der Klassifizierung, in der Broschart (1997: 124 und passim) explizit Wortarten als Begriffe aufgibt und vorschlägt, stattdessen zwischen „Types“ und „Tokens“ zu unterscheiden. Broschart (1997) entlehnt die Begriffe *Type* und *Token* Jackendoff (1983) – dabei seien Types identisch mit Prädikaten im Sinne der Prädikatenlogik, Tokens hingegen stellen die „linguistic instantiations“ (Broschart 1997: 156) der Types dar. Types bzw. Prädikate seien dann Frege (1892) folgend als referentiell ungesättigt, Tokens als referentiell gesättigt zu betrachten. Tokens wiederum

⁹Man müsste genauer von flexiblen Worten sprechen, wobei auch damit nichts anderes gesagt würde, als dass keine Wortarten unterscheidbar sind.

wären entweder prädizierbar, wenn sie in Tempus-Aspekt-Modus-Syntagma auftreten, oder nicht prädizierbar, wenn sie in Artikel-Syntagma auftreten, vgl.:

- (2) [na'e kata]_{TAM} [e tangatá]_{ART}
 TAM laugh ART man.DEF
 'The man laughed.'
 (Broschart 1997:134)

Während in Type-Token-Sprachen wie dem Tongaischen eine Unterteilung zwischen nicht-referentiellen Types, z.B. *na'e* oder *tangata*, und referentiellen Tokens, z.B. *na'e* in Tempus-Aspekt-Modus-Syntagma in Satz (2) oder *tangata* in Artikel-Syntagma in Satz (2), gegeben sei, würde in Nomen-Verb-Sprachen wie dem Englischen zwischen nicht-prädizierbaren Nomen wie *man* in (2) und prädizierbaren Verben wie *laugh* in (2) unterschieden (vgl. Broschart 1997: 157) – auf der Definition von Hengeveld (1992: 58) beruhend, in der Verben nur in prädikativer Phrase auftreten können, genüge für die Unterscheidung von Nomen und Verben das Kriterium [+/- prädizierbar] (vgl. Broschart 1997: 155). Letztlich versteht Broschart (1997) unter Types nichts anderes als lexikalische Einheiten (vgl. Broschart 1997: 155) und unter Tokens Wörter in ihrem aktuellen Auftreten. Mir scheint die von Broschart (1997) vorgenommene Gegenüberstellung von Type-Token-Sprache und Nomen-Verb-Sprache im Kern auf einem auf verschiedenen Beschreibungsebenen stattfindenden Vergleich zu basieren: Erstens werden zwar in beiden Sprachtypen Wörter in ihrem aktuellen Auftreten bzw. Tokens als sprachliche Kategorien, in Type-Token-Sprachen im Gegensatz zu Nomen-Verb-Sprachen aber lexikalische Einheiten bzw. Types als rein logische Kategorien analysiert. Zweitens werden auf Type-Token-Sprachen die Begriffe Referenz und Prädikation in unterschiedlicher Weise angewandt als auf Nomen-Verb-Sprachen. So lässt sich unterscheiden zwischen Referenz und Prädikation als diskurspragmatische Funktionen – Referenz als Bezug auf einen Gegenstand, um ihn zum Satzgegenstand zu machen im Gegensatz zu Prädikation als Aussage über diesen – und Referenz im Sinne von aktueller Bezeichnung im Satz (vgl. Mumm 1995: 422–423), bei der man „statt zu fragen, was ein Sprecher bezeichnet, um darüber eine Aussage zu machen, sich auf die Frage beschränkt, was dieser Sprecher überhaupt in dem Satz bezeichnet [...]“ (Mumm 1995: 423). Im Sinne der aktuellen Bezeichnung im Satz referieren aber nicht nur Tokens in Type-Token-Sprachen, sondern auch

Nomen und Verben in Nomen-Verb-Sprachen. So referiert das Nomen *man* in Satz (2) (S. 23) auf den Mann und das Verb *laugh* auf die Tätigkeit des Mannes, wobei gilt: „**Die** [Hervorhebung im Original] Bezeichnung eines Nomens‘ existiert so strenggenommen nicht, sondern nur die mithilfe der nominalen Bedeutung von Sprechern jeweils aktuell vorgenommene Bezeichnung.“ (Mumm 1995: 221). Insofern hat Broschart (1997) damit recht, dass anders als z.B. bei *laugh* und *laughter* im Englischen bei *kata* im Tongaischen die jeweilige aktuelle Bezeichnung nicht mithilfe der verbalen oder nominalen Bedeutung vorgenommen wird, sondern erst im Auftreten im Syntagma, *kata* in Tempus-Aspekt-Modus-Syntagma, wie *na‘e kata* in Satz (2) (S. 23), versus *kata* in Artikel-Syntagma, wie *e kata* „the laughing, the laughter“ (vgl. Broschart 1997: 134). Das bedeutet aber nicht, dass im diskurspragmatischen Sinne von Referenz in Nomen-Verb-Sprachen Wörter an sich referieren – das können Wörter gar nicht – und dagegen in Type-Token-Sprachen Wörter an sich nicht referieren. Im diskurspragmatischen Sinne werden sowohl *tangatá* als auch *man* in Satz (2) (S. 23) im Auftreten in Artikel-Syntagma zum Satzgegenstand gemacht, über den mit *kata* bzw. *laugh* im Auftreten in Tempus-Aspekt-Modus-Syntagma etwas ausgesagt wird. Broschart (1997) stellt also die Trennung von Type und Token anhand des Kriteriums der Referenz im Sinne der aktuellen Bezeichnung der Trennung von Verb und Nomen anhand des Kriteriums der Referenz und Prädikation im Sinne der pragmatischen Funktion gegenüber und analysiert in Sprachen wie dem Tongaischen Types als logische Prädikate, um Wörtern als Tokens auf syntaktischer Ebene einen eigenständigen kategorialen Status unabhängig von der Differenzierung im Lexikon zuzuweisen.

Den entscheidenden Vergleich sehe ich hier aber wie bei Hengeveld (1992) und in der tradierten Wortarttheorie in der Frage gegeben, ob Wörter für den Gebrauch im Satz prädisponiert sind.

Im Gegensatz zu den Arbeiten von Hengeveld (1992) und Broschart (1997), die im Rahmen der typologischen Klassifizierung der polynesischen Sprachen als flexible Sprachen oder Type-Token-Sprachen Wortarten ablehnen, stehen Ansätze, die eine protoypische Abdeckung semantischer Klassen auf syntaktisch-funktionale Slots annehmen. Diese Ansätze sollen im Folgenden genauer betrachtet werden.

2.2.4 Prototypische Abdeckung semantischer Klassen auf syntaktisch-funktionale Slots

Wie zuvor gezeigt (vgl. Kap. 2.2.1, S. 11–14), wird in vielen Grammatiken einzelner polynesischer Sprachen zumindest implizit die Annahme vertreten, Wörter seien entsprechend ihrer Zugehörigkeit zu einer semantischen Klasse für das Auftreten in syntaktischen Slots prädisponiert. In diesen Grammatiken spielen weder diskurspragmatische Funktionen eine Rolle noch wird sich Wörtern und ihrer Bedeutung, die in mehr als einem syntaktischen Slot auftreten, ausführlich theoretisch gewidmet. Aus einer typologischen Perspektive hat dies Croft (2001) versucht, in dessen Folge eine Diskussion um Polysemie und Monosemie in den polynesischen Sprachen entstanden ist. Im diesem Kapitel soll zunächst Crofts (2001) Ansatz ausführlicher behandelt werden, um danach zwei weitere Ansätze (Don und Van Lier 2013 und Völkel 2017) vorzustellen und schließlich in der Auseinandersetzung um Polysemie und Monosemie zu vergleichen.

Der Ansatz von Croft (2001)

Ähnlich wie Hengeveld (1992) setzt Croft (2001) bei der universellen Unterscheidung der diskurspragmatischen Funktionen Referenz, Prädikation und Modifikation an, kritisiert aber in seiner *Radical Construction Grammar* (Croft 2001) Hengeveld (1992) dafür, dass dieser die Semantik von Wörtern in entsprechenden Phrasen nicht berücksichtigt (vgl. Croft 2001: 67 und passim).

Auch Croft (2001) unterscheidet wie Hengeveld (1992) nicht zwischen den syntaktischen Einheiten, über die pragmatischen Funktionen ausgedrückt werden und den pragmatischen Funktionen selbst (vgl. Kapitel 2.2.3, S. 17).

Croft (2001: 67 und passim) zufolge übersieht Hengeveld (1992), dass die konzeptuell-semantischen Klassen OBJEKT, EREIGNIS¹⁰ und EIGENSCHAFT prototypisch auf die diskurspragmatischen Funktionen Referenz, Prädikation und Modifikation abgebildet werden (vgl. auch Croft 2005: 438 und passim).

Das Kriterium der Prototypikalität liegt in dem begründet, was Croft (2001) nach eigenen Angaben angelehnt an Greenberg (1966) *Typologische Markiertheit* nennt (vgl. Croft 2001: 89), vergleichbar mit den „further measures“ bei

¹⁰Croft (2001) spricht im Englischen mehrheitlich von „action“. Da Croft (2001) darunter aber nicht nur Handlungen versteht, sondern auch Ereignisse ohne Agens, habe ich dies hier entsprechend als Ereignis (vgl. engl. „event“) übersetzt.

Hengeveld (1992), wobei gelten soll: „Typological markedness represents a cross-linguistic universal about the encoding of function by linguistic form.“ (Croft 2001: 91). Typologische Markiertheit unterteilt Croft (2001: 66 und passim) in die Kriterien des „structural coding“ und „behavioral potential“. Soweit ich Croft (2001: 66) richtig verstehe, umfasst „structural coding“ neben Fällen wie *the detective who is singing*, die Hengeveld (1992) als Beispiele für „further measures“ nennt (vgl. Kapitel 2.2.3, S. 17) vor allem Derivationsmorphologie, während sich „behavioral potential“ auf die Flexionsmorphologie bezieht (vgl. Croft 2001: 91 und passim).¹¹

Entscheidend ist hier die universelle Gegenüberstellung von markiertem und unmarkiertem Fall in Kombination mit der Zuweisung semantischer Klassen und dem Auftreten in syntaktisch-funktionaler Konstruktion, vgl. Tabelle 3:

semantische Klasse	pragmatische Funktion		
	Referenz	Prädikation	Modifikation
OBJEKT	unmarkiertes Nomen	Prädikatsnomen, Kopula	Genitiv, Adjektivisierung
EREIGNIS	Nomen Actionis, Invinitiv, Gerundium	unmarkiertes Verb	Partizip Relativsätze
EIGENSCHAFT	deadjektivisches Nomen	Prädikatsadjektiv, Kopula	unmarkiertes Adjektiv

Tabelle 3: prototypische Abdeckung semantischer Klassen auf pragmatische Funktionen (vgl. Croft 2001: 88)

Der universell prototypische Fall soll also daran erkennbar sein, dass z.B. ein Wort, das in einer referentiellen Phrase auftritt und dessen Bedeutung als EREIGNIS klassifiziert ist, mehr markiert ist relativ zu einem Wort, das als Kopf einer referentiellen Phrase auftritt und dessen Bedeutung als OBJEKT klassifiziert ist, wobei Croft (2001) etwas unklar bleibt in der (formalen) Bestimmung der Einheit, an der typologische Markiertheit gemessen und der die entsprechende Semantik zugewiesen wird. So spricht er einmal von „item“ (Croft 2001: 18) oder „lexical item“ (Croft 2001: 71), von „stem“ (Croft 2001: 66), „word“ (Croft 2001: 67) oder „lexeme“ (Croft 2001: 67).

¹¹Verwirrend ist, dass Croft (2001: 66) dann „structural coding“ im Gegensatz zu „behavioral potential“ als „function indicating morphosyntax“ bezeichnet.

Die konzeptuell-semantischen Klassen OBJEKT, EREIGNIS und EIGENSCHAFT könnten Croft (2001: 87) zufolge anhand der vier semantischen Merkmale [+/- relational], [+/- stativ], [+/- transitorisch] und [+/- graduierbar] bestimmt werden, vgl. Tabelle 4:

semantische Klasse	semantisches Merkmal			
	relational	stativ	transitorisch	graduierbar
OBJEKT	-	+	-	-
EREIGNIS	+	-	+	-
EIGENSCHAFT	+	+	-	+

Tabelle 4: Definition semantischer Klassen durch semantische Merkmale (vgl. Croft 2001: 88)

Die Bedeutung eines Wortes wie *detective* könnte so als OBJEKT klassifiziert werden, weil sie durch die Eigenschaften [- relational], [- transitorisch] und [- graduierbar], aber [+ stativ] erfassbar ist.

Auf den ersten Blick mögen sich die Ansätze von Hengeveld (1992) und Croft (2001) nicht wesentlich voneinander unterscheiden, da Croft (2001) mit der Semantik nur einen weiteren Aspekt in Hengevelds (1992) Modell einbringt. So analysiert Croft (2001: 66) die Beispiele aus Hengeveld (1992: 58) (vgl. Kapitel 2.2.3, S. 17) in ähnlicher Weise, nur dass z.B. *intelligent* in *the intelligent detective* semantisch als EIGENSCHAFT klassifiziert wird.

Der Unterschied offenbart sich erst im Detail, und zwar in dem, was Hengeveld (1992) als „further measures“ und was Croft (2001) als typologisch markiert gelten lässt und welche Konsequenzen Croft (2001) daraus zieht. Dies sei an dem Beispiel *happy* versus *happi-ness* von Croft (2001: 88) gezeigt:

Während Hengeveld (1992) *happy*, da es ohne „further measures“ in modifizierender Phrase auftritt, als Adjektiv und *happiness*, da es ohne „further measures“ in referentieller Phrase auftritt, als Nomen analysieren müsste, stellt für Croft *happy* den unmarkierten, i.e. prototypischen Fall in der Abdeckung der semantischen Klasse EIGENSCHAFT auf eine modifizierende Phrase dar in Gegenüberstellung zu *happi-ness* als nicht-prototypischer Fall in der Abdeckung der semantischen Klasse EIGENSCHAFT auf eine referentielle Phrase.

Entscheidend ist nun, dass Croft (2001) die von ihm definierten prototypischen Fälle als „typological prototypes“ (Croft 2001: 87) versteht und für sich reklamiert, damit eine universalistisch-typologische Theorie von Parts-of-Speech geschaffen zu haben, die im Gegensatz zu einer auf einzelsprachlicher Beschreibung basierenden Theorie zu Wortarten steht (vgl. Croft 2001: 63, auch Croft

2005: 437). Croft (2001, 2005) zufolge müssten Wortarten als rein einzelsprachliche Kategorien betrachtet werden, die sich einer sprachübergreifenden Analyse entziehen. Die von Croft (2001) definierten Parts-of-Speech seien hingegen universelle Kategorien, da sie auf externen, funktionalen Kriterien basieren, vgl.: „These constructions are defined on external grounds and therefore can be identified and hence compared across languages.“ (Croft 2005: 437).

Croft (2001: 89) bemerkt schließlich, ohne einen möglichen Zusammenhang zu einzelsprachlichen Wortartkategorien zu erläutern, dass die von ihm definierten „typological prototypes“ das sind, was man universell als *Verb*, *Nomen* und *Adjektiv* bezeichnet kann. Haspelmath (2012) hat später der Forderung Crofts (2001), Wortarten als einzelsprachliche Kategorien aufzugeben, beigepröflichtet und vorgeschlagen, an deren Stelle „comparative concepts“ (Haspelmath 2012: 114) zu setzen. Darunter versteht er ähnlöh wie Croft (2001) die semantischen Klassen OBJEKT, EREIGNIS und EIGENSCHAFT und bemerkt: „To be universally applicable, comparative concepts can be defined on the basis of meaning or sound, but not on the basis of meaning-sound combinations, because these are language-particular.“ (Haspelmath 2012: 114). Haspelmath (2012: 114 und passim) möchte „comparative concepts“ dann allerdings übereinzelsprachlich auf Wurzelebene – er spricht von „thing-root“, „action-root“ und „property-root“ (Haspelmath 2012: 115) – vergleichen und sieht es als unproblematisch an, für diese die tradierten Bezeichnungen *Nomen*, *Verb* und *Adjektiv* zu verwenden, solange man sich bewusst sei, eigentlich über „comparative concepts“ und nicht einzelsprachliche Wortarten zu sprechen (vgl. Haspelmath 2012: 115).

Ich sehe hier im Kern zwei Probleme: erstens als Voraussetzung die eindeutige Zuordnung der Bedeutung nicht nur eines Wortes, sondern über Wortgrenzen hinweg zu einer semantischen Klasse und zweitens das unklare Verhältnis von einzelsprachlichen Wortartkategorien und übereinzelsprachlichen Parts-of-Speech.

Betrachtet sei hier noch einmal das Beispiel von Croft (2001: 88) *happy* versus *happi-ness*: Der Analyse Crofts von *happy* als unmarkiertes, prototypisches Adjektiv in Gegenüberstellung zu *happines*, liegt die Voraussetzung zugrunde, trotz formaler Unterscheidbarkeit die semantische Identität von *happy* und *happiness* anzunehmen (vgl. S. 27). Abgesehen davon, dass es sich bei *happy* und *happiness* und zwei Wörter im Sinne von Lexemen bzw. morphologisch differenzierbaren Stämmen handelt, muss man *happy* und *happines* keinesfalls als semantisch identisch sehen. So schreibt Schmid (2016: 172) beispielsweise

zu englischen Wortbildungsprozessen: „[...] the function of deadjectival noun-forming suffixes is to reify qualities.“ Versteht man unter *Reifikation* (von lat. *res* „Ding“) Vergegenständlichung, bedeutet dies bei *happiness* die Konzeptualisierung einer Eigenschaft als Objekt (vgl. auch Langacker 1987b).

In ähnlicher Weise kann man auch mit Schulze (2012) *happiness* als Objektvorstellung analysieren. Wie Croft (2001) setzt Schulze (2012) konzeptuelle, universell in der menschlichen Kognition verankerte Kategorien OBJEKT und EREIGNIS an – Schulze (2012: 9) spricht von „Schematic Units of Event Images (SUEI)“ und unterscheidet zwischen Wortarten als Ergebnis wissenschaftlicher Klassifizierung und Parts-of-Speech (PoS) als „linguistic instantiations of SUEIs“ (Schulze 2012: 9), wobei gilt: „PoS are not grounded in language, but in cognition as ‘such’“ (Schulze 2012: 8). Entscheidend für den Umgang mit dem Ansatz von Croft (2001) und dem Beispiel *happiness* ist bei Schulze (2012) das Verhältnis von konzeptueller und sprachlicher Ebene, vgl. Abbildung 1:

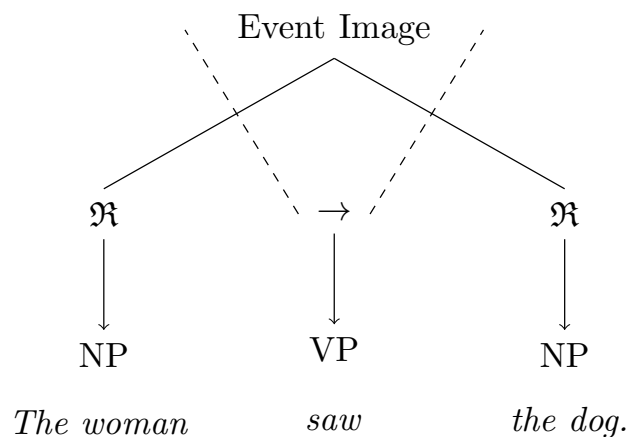


Abbildung 1: Schematisierung von Ereignisvorstellungen (vgl. Schulze 2012)

Auf einer konzeptuellen Ebene können Schulze (2012: 25 und passim) zufolge Ereignisvorstellungen (Event Images) als relationales Schema betrachtet werden, im Beispiel von Abbildung 1 aufgebaut aus zwei in einer Figure-Ground-Beziehung stehenden rereferentiellen Einheiten oder Objektvorstellungen (symbolisiert durch „ \mathfrak{R} “) und einer relationalen Einheit oder Ereignisvorstellung (symbolisiert durch „ \rightarrow “), die wiederum in einem meronymen Verhältnis zur gesamten Ereignisvorstellung steht (symbolisiert durch die gestrichelten Linien). Sprachlich erscheint dieses Schema als Satz mit zwei Nominalphrasen und

einer Verbalsphrase – der Kopf einer Nominalphrase profiliert als OBJEKT, der Kopf der Verbalphrase profiliert als EREIGNIS, wobei unklar ist, ob Schulze (2012) mit *Parts-of-Speech* Phrasen oder die Köpfe von Phrasen bezeichnet und warum Schulze (2012: 35 und passim) an universellen Begriffen *Nominal-* und *Verbalphrase* festhält, obwohl Schulze (2012: 33) die Wortartkategorien Nomen und Verb nicht als universell ansieht (vgl. auch die Kritik an Wortarten als syntaktische Kategorien in Kap. 2.2.2, S. 14–16). In jedem Fall müsste man *happiness* in Schulzens (2012) Modell aber im Auftreten in Nominalphrase konzeptuell als Objektvorstellung analysieren bzw. bei Croft (2001) semantisch als OBJEKT klassifizieren.

Lässt man Detailfragen dieses Modells beiseite, geht es mir bezogen auf den Ansatz von Croft (2001) noch um einen weiteren Aspekt, der die Frage um das Verhältnis von einzelsprachlichen Wortartkategorien und universellen Parts-of-Speech eröffnet: Wenn man mit Schulze (2012) und Schmid (2016) *happiness* semantisch als OBJEKT analysiert und davon ausgeht, dass der Kopf einer Nominalphrase semantisch immer als OBJEKT profiliert ist, dann muss man nicht wie Croft (2001) *happy* als universelles Adjektiv in der Abdeckung der semantischen Klasse EIGENSCHAFT auf eine modifizierende Phrase und *happiness* als unprototypischen Fall in der Abdeckung der semantischen Klasse EIGENSCHAFT auf eine referentielle Phrase und damit nicht als universelles Nomen sehen, sondern kann sehr wohl *happy* als Wortart Adjektiv in der Prädisposition für das Auftreten in modifizierender Phrase und damit semantischer Klasse EIGENSCHAFT analysieren und *happiness* als Wortart Nomen in der Prädisposition für das Auftreten in referentieller Phrase und damit semantischer Klasse OBJEKT. Zwar mag man *happy* als semantisch prototypische EIGENSCHAFT betrachten und *happiness* nicht als semantisch prototypisches OBJEKT, als Nomen in Nominalphrase ist *happiness* semantisch jedoch EIGENSCHAFT als OBJEKT.

Ich sehe dann aber sogar unter Beibehaltung der semantischen Analyse von Croft (2001) in seinem Ansatz ein ähnliches Problem wie bei Hengeveld (1992): Zwar mag man universell semantische Klassen und funktionale Konstruktionen definieren können, aber auch um im Croftschen Sinne von Parts-of-Speech zu sprechen, wird durch die Gegenüberstellung von markiertem und unmarkiertem Fall wie in der tradierten Wortarthorie eine systematische Korrelation zur formalen Dimension postuliert.

Unklar ist dann, warum auf der einen Seite Wortartkategorien als reine einzelsprachliche Beschreibungen stehen und auf der anderen Seite die auf pragma-

tischen Funktionen und semantischen Klassen aufbauenden Parts-of-Speech als universelle Kategorien.

Wenn es sich bei Wortarten und Parts-of-Speech um zwei substantiell verschiedene Kategorien handelt, kann man sie erstens unabhängig voneinander beschreiben, freilich ohne *a priori* einen Zusammenhang auszuschließen und zweitens heißt das nicht zwangsläufig, dass Wortarten als solche für einen übereinzelsprachlichen Vergleich untauglich sind. Nach der schwachen Äquivalenzhypothese (vgl. Sasse 1993: 190–191, vgl. auch Kapitel 2.1, S. 6) ist ein Vergleich prinzipiell möglich, und: Wortarten überhaupt als Kategorien anzusetzen, wenn auch als partikulare, so wie es Croft (2001) und Haspelmath (2014) tun, schließt die Unvergleichbarkeit aus. Was sollen nicht vergleichbare Kategorien Wortarten in einzelnen Sprachen sein? In der Konsequenz dürfte gar nicht mehr von Wortarten als Kategorien gesprochen werden, wenn es sich nicht bloß um inhaltsleere Etiketten handeln sollte.

In Crofts (2001) Theorie bleibt dann vor allem unklar, welcher kategorielle Status den nicht-prototypischen Fällen zugewiesen wird.

Ich sehe hier zwei Alternativen in der Gegenüberstellung von prototypischen und unprototypischen Fällen: Erstens nicht-prototypische Fälle werden in die einzelsprachliche Dimension verschoben, vgl. Tabelle 5:

prototypisch universell Parts-of-Speech		nicht-prototypisch einzelsprachlich? Wortarten?	
Nomen	Referenz auf OBJEKT	Nomen?	Prädikation durch OBJEKT Modifikation durch OBJEKT
Verb	Prädikation durch EREIGNIS	Verb?	Referenz auf EREIGNIS Modifikation durch EREIGNIS
Adjektiv	Modifikation durch EIGENSCHAFT	Adjektiv?	Referenz auf EIGENSCHAFT Prädikation durch EIGENSCHAFT

Tabelle 5: Gegenüberstellung von prototypischen und nicht-prototypischen Fällen bei Croft (2001)

Gegen die Einteilung, wie sie in Tabelle 5 vorgenommen wurde, spricht allerdings, dass das Modell von Croft (2001) im Gesamten prototypische und nicht-prototypische Fälle gleichermaßen umfasst, da erst mittels Gegenüberstellung zu nicht-prototypischen Fällen die prototypischen als solche definiert sind. Die unprototypischen Fälle können dann nicht rein einzelsprachliche Kategorien

darstellen, wie auch am Beispiel von *happy* und *happiness* zu sehen ist: *happiness* kann nicht bloß als einzelsprachliche Wortartkategorie Nomen und *happy* als einzelsprachliche Wortartkategorie Adjektiv analysiert werden, wenn anhand des Vergleichs gezeigt werden soll, dass *happy* der prototypische, unmarkierte Fall ist und damit universell als Adjektiv gelten soll.

Zweitens muss also für jede Sprache angenommen werden, dass die semantischen Klassen OBJEKT, EREIGNIS und EIGENSCHAFT auf jede der pragmatischen Funktion abgedeckt werden können (vgl. auch Croft 2001: 73).

Dann ist die Theorie Crofts (2001) zumindest teilweise vereinbar mit der Annahme von Wortarten Verb, Nomen und Adjektiv, wenn Croft (2001) mit seinen Parts-of-Speech allein semantisch prototypische Nomen, Verben und Adjektive definiert. Beispielsweise ist ein semantisch prototypisches Nomen dann ein nicht-deriviertes Nomen wie *man*, ein semantisch nicht-prototypisches Nomen z.B. ein Nomen Actionis wie *destruction* (vgl. Croft 2001: 88), das über kategorienverändernde Derivation von einer verbalen Basis gebildet wird oder ein Nomen Qualitatis wie *happiness*, das über kategorienverändernde Derivation von einer adjektivischen Basis gebildet wird, wobei Croft (2001) derivierte Nomen, die semantisch als OBJEKT klassifiziert werden, z.B. *detective*, gesondert behandeln müsste (vgl. dazu S. 34–35).

In allen Fällen sind jedoch Nomina über die Prädisposition für syntaktisch-pragmatische Konstruktionen definiert.

Das fällt zusammen mit der Annahme der tradierten Wortartlehre, dass Verben eher die semantische Klasse EREIGNIS, Nomen die semantische Klasse OBJEKT und Adjektive die semantische Klasse EIGENSCHAFT abbilden, ohne aber andere Fälle auszuschließen – Croft (2001: 73) selbst beansprucht, mit seiner Theorie, die semantische Begründung der Wortarten in Teilen bewiesen zu haben. Croft (2001) vermag diese Annahme durch die Gegenüberstellung formal markierter und unmarkierter Fälle präziser zu erfassen, kommt damit aber um eine einzelsprachliche, morphologische Analyse nicht herum.

Die von Croft (2001: 87 und passim) angesetzten „typological prototypes“ oder Parts-of-Speech, dass heißt der Zusammenhang von formaler Markierung, semantischer Klasse und pragmatischer Funktion, muss allemal für jede Einzelsprache auch anhand der spezifischen formalen Gegebenheiten nachgewiesen werden. Dabei bleibt die Frage offen, ob eine Sprache semantisch prototypische Nomen, Verben und Adjektive ohne Nomen, Verben und Adjektive aufweisen kann, oder anders ausgedrückt: Kann eine Sprache Parts-of-Speech haben, aber keine Wortarten? Zumindest im Hinblick auf die polynesischen Sprachen

scheint mir dies zweifelhaft.

Croft (2001: 68 und passim) selbst hat auf vergleichbare Fälle wie *hana* in (1a) und (1b) (S. 1) reagiert und zwar auf folgende von Hengeveld (1992: 66) diskutierte und von Tchekhoff (1981: 4) übernommene Beispiele aus dem Tongaischen, vgl.:

(3a) *na'e si'i 'ae akó*
PST **small** ABS **school:DEF**
'The school was small.'

(3b) *'i 'ene si'i*
in POSS.3SG **childhood:DEF**
'in his/her childhood'

(3c) *na'e ako 'ae tamasi'i si'i iate au*
PST **study** ABS child **little** LOC 1SG
'The little child studied at my house.'

Hengeveld (1992: 66–67) betrachtet *si'i* und *ako* als lexikalisch flexibel und meint damit, dass aufgrund fehlender „further measures“ in entsprechenden Funktionen im Tongaischen keine Wortartunterscheidung gegeben ist.

Croft (2001: 71) kritisiert nun Hengeveld (1992) dafür, dass er Wörter wie *si'i* und *ako* als semantisch vage analysieren müsste, damit aber die Bedeutungsunterschiede der beiden im Gebrauch in den Sätzen (3a) – (3c) nicht erklären könne (vgl. auch Evans und Osada 2005).

Genau in diesem Bedeutungsunterschied läge aber die universelle, prototypische Zuweisung von semantischer Klasse und diskurspragmatischer Funktion feststellbar. Denn ein Wort wie *si'i* trage zwei Bedeutungen, eine als EIGENSCHAFT wie in Satz (3a) und Satz (3c) und eine als OBJEKT wie in Satz (3b) klassifizierbar, und ein Wort wie *ako* trage ebenso zwei Bedeutungen, eine als OBJEKT wie in Satz (3a) und eine als EREIGNIS wie in Satz (3c) klassifizierbar; *si'i* und *ako* müssten folglich darin unterschieden werden, dass sie in ihren jeweiligen, identifizierbaren Bedeutungsklassen prototypisch auf die entsprechenden diskurspragmatischen Funktionen abgedeckt werden, vgl. Croft (2001: 72): „[...] the particular use represents a conventional meaning of the expression, that is, that there are two senses of the word in question, each of which may belong to a different part of speech.“

Damit könne dann von *si'i* in der Bedeutung „small“ als prototypisches Ad-

ektiv, in der Bedeutung von „childhood“ als prototypisches Nomen und *ako* in der Bedeutung von „school“ als prototypisches Nomen, in der Bedeutung von „study“ als prototypisches Verb gesprochen werden.

Ich sehe in dieser Analyse der polynesischen Sprachen von Croft (2001) mehrere Probleme: die unklare Unterscheidung von Polysemie und „semantic shift“ (Croft 2001: 70), dabei die eindeutige Identifikation der semantischen Klasse der Bedeutung eines Wortes ohne Entscheidbarkeit über die Richtung des Shifts und schließlich das Postulat der Prototypikalität von Parts-of-Speech ohne Anwendbarkeit des Kriteriums der typologischen Markiertheit.

Croft (2001) analysiert genau genommen die verschiedenen Bedeutungsweisen der Wörter *ako* und *si'i* sehr unterschiedlich, und zwar *ako* als Fall von semantischem Wechsel und *si'i* als Fall von Polysemie, wobei Croft (2001) diese Trennung nicht explizit macht. Zu semantischem Wechsel und *ako* schreibt Croft (2001: 74): „Action words used in referring expressions shift meaning to a person, place, or thing which is a typical salient participant in the action in its semantic frame (in the sense of Fillmore 1982), as with Tongan *ako* [...].“

Unter *ako* als „action word“ muss dann verstanden werden, dass eine Bedeutung, die Croft (2001: 74) als EREIGNIS klassifiziert, in der Form *ako* lexikalisiert ist und die im Auftreten der Form *ako* in referentieller Phrase zu einer Bedeutung, die Croft (2001: 74) als OBJEKT klassifiziert, wechselt. Im Gegensatz zu *ako* analysiert Croft (2001: 71) *si'i* als polysem. Das heißt, zwei Bedeutungen – Croft zufolge eine als EREIGNIS, eine als OBJEKT klassifizierbar – sind in einem Wort mit der Form *si'i* lexikalisiert und werden erst im Kontext disambiguiert. Unklar ist aber, warum *ako* als ein monosemes Wort mit einer lexikalisierten, klassifizierbaren Bedeutung und Wechsel dieser im Auftreten in syntaktisch-funktionaler Konstruktion und *si'i* als ein polysemes Wort mit zwei lexikalisierten, klassifizierbaren Bedeutungen, die prototypisch in entsprechender syntaktisch-funktionaler Konstruktion auftreten, betrachtet werden soll.

Problematisch sehe ich in der Annahme von semantischem Wechsel zudem ähnlich wie bei der Konversionsanalyse von Vonen (2000) und Don und Van Lier (2013) (vgl. Kap. 2.2.1, S. 13–14)¹², dass ein angenommener Wechsel immer auch eine Richtung desselben miteinschließt, die aber bei *ako* überhaupt nicht eindeutig bestimmbar ist, vgl. Croft (2001: 74): „Conversely, an object

¹²Croft (2001: 73) selbst hält entgegen der Konversionsanalyse die Polysemie-Analyse für die bessere Alternative.

word used in a predication construction shifts meaning to an action typically or saliently associated with the object in its semantic frame [...]“. Handelt es sich bei *ako* also um ein „action word“, das die Bedeutung zu einem salienten Teilnehmer seines Ereignis-Frames wechselt oder um ein „object word“, das die Bedeutung zu einem salienten Ereignis aus seinem Objekt-Frames wechselt? Beides ist denkbar.

In einer Polysemie-Analyse müsste Croft (2001) zumindest aus synchroner Perspektive überhaupt keinen Wechsel annehmen. Dieser wäre bei *ako* weder von der Bedeutung „study“ zu „school“ noch umgekehrt oder bei *si'i* weder von der Bedeutung „small“ zu „childhood“ noch umgekehrt gegeben.

Allerdings ist auch *si'i* nicht eindeutig als polysem zu analysieren. Völkel (2014: 460–461) hat darauf hingewiesen, dass *si'i* je nach Kontext auch so etwas wie „being small“ bedeuten kann, z.B. wenn ein Kind über eine Puppe spricht, vgl.:

- (4) *Ko 'ene si'isi'i 'oku tōatu.*
 PRES A-POSS:3SG small TAM:PRES cute
 ‘Its beeing small is cute.’
 (Völkel 2017:461)

Ebenso hat bereits Churchward (1953: 110) bemerkt, dass *si'i* je nach Kontext wortwörtlich mit so etwas wie „being small“ oder „smallness“ übersetzt werden muss, vgl.:

- (5) *Na'e hoko 'i he'eku kei si'i.*
 TAM:past happen while POSS:1SG still small
 ‘It happened while I was still a child [my beeing small].’
 (Churchward 1953: 110)

Croft (2001) müsste dann bei *si'i* als polysemes Wort unterscheiden zwischen einer lexikalisierten Eigenschaftsbedeutung, in der *si'i* prototypisch in modifizierender Phrase wie in Satz (3a) und Satz (3b) (S. 33) und nicht-prototypisch in referentieller Phrase wie in Satz (4) auftritt, von einer lexikalisierten Objektbedeutung, in der *si'i* prototypisch in referentieller Phrase wie in Satz (3b) (S. 33) und möglicherweise nicht-prototypisch in modifizierender Phrase auftritt. Mit Churchward (1953: 110) und Völkel (2017: 461), die bei *si'i* in der Bedeutung von „childhood“ von einer Spezifizierung einer lexikalisierten und

als EIGENSCHAFT klassifizierten Bedeutung ausgeht (vgl. Völkel 2017: 473), ließe sich einwenden, dass *si'i* als monosem betrachtet werden sollte.

Dem könnte mit Croft (2001) wiederum entgegengehalten werden, dass auch wenn im Tongaischen die Kindheit als Eigenschaft, klein zu sein, konzeptualisiert wird, dieser semantische Zusammenhang als metonym analysierbar ist, was für Croft (2001: 73) einen Fall von systematischer Polysemie darstellt.

Das wirft erstens die Frage auf, ob die Polysemie von *si'i* zwischen der als EIGENSCHAFT klassifizierbaren Bedeutung „small“ und der als OBJEKT klassifizierbaren Bedeutung „childhood“ besteht oder zwischen der Bedeutung von *si'i* im Auftreten in referentieller Phrase „being small“ oder „smallness“ und der Bedeutung von *si'i* im Auftreten in referentieller Phrase „childhood“ (vgl. dazu auch die Frage der semantischen Klasse von *happiness*, S. 28–30) und zweitens, ob *si'i* in der Bedeutung von „childhood“ semantisch als EIGENSCHAFT nämlich der Eigenschaft, klein zu sein, oder als OBJEKT nämlich der Zeit, in der man klein ist, klassifiziert werden muss.

Unabhängig von der Entscheidung über die genaue semantische Analyse ist der entscheidende Punkt für Croft (2001), zeigen zu können, dass bei *si'i* in der jeweiligen Lesart die prototypische Abdeckung von semantischer Klasse auf pragmatische Funktion gegeben ist und somit Parts-of-Speech im Tongaischen unterschieden werden müssen.

Genau darin liegt aber das Kernproblem der Analyse. Denn es ist gar nicht möglich, den prototypischen Fall zu bestimmen, wenn das dafür entwickelte Kriterium der *typologischen Markiertheit* überhaupt nicht anwendbar ist.

Wenn im Tongaischen *si'i* mit Eigenschaftsbedeutung sowohl in modifizierender Phrase als auch referentieller Phrase auftritt, wie soll dann ohne formale Unterscheidbarkeit der prototypische vom nicht-prototypischen Fall unterschieden werden? Gleiches gilt für den angenommenen semantischen Shift bei *ako*, wenn Croft (2001: 73) folgendes behauptet:

If there is a semantic shift in zero coding of an occurrence of a word (i.e. flexibility) in a part-of-speech construction, even if it is sporadic and irregular, it is always towards the semantic class PROTOTYPICALLY [Hervorhebung im Original] associated with the propositional act function, that is, the semantic classes for parts of speech [...].

Zwar weist Croft (2001: 90) darauf hin, dass der unprototypische Fall nur mindestens so viel markiert sein muss, wie der prototypische. So besteht die Möglichkeit, dass prototypischer und nicht-prototypischer Fall gleich markiert sind und ausgeschlossen ist nur, dass der prototypische mehr markiert ist als der nicht-prototypische (vgl. Croft 2001: 90 und *passim*). Croft (2001: 76)

resümiert schließlich: „The fundamental fact that is overlooked is that while difference of form entails difference in categorization, identity of form does not entail identity of categorization [...].“ Nur hilft diese Überlegung in der Beschreibung von Wortarten in Sprachen wie dem Tongaischen oder Hawaiischen keineswegs weiter: Es ist einfach nicht entscheidbar, was der prototypische, unmarkierte Fall ist, wenn alle Fälle gleich markiert bzw. gleichermaßen nicht markiert sind, wie bei *si‘i* im Vergleich in den Sätzen (3a) – (3c) (S. 33), (4) und (5) (S. 35) oder bei *hana* im Vergleich von Satz (1a) zu (1b) (S. 1). So könnte man auch ohne weiteres behaupten, dass *si‘i* mit Eigenschaftsbedeutung prototypisch in referentieller Phrase auftritt und nicht-prototypisch in modifizierender Phrase oder *hana* mit Ereignisbedeutung prototypisch in referentieller Phrase.

Lässt sich keine prototypische Abbildung semantischer Klassen auf pragmatische Funktionen in den polynesischen Sprachen nachweisen, muss man die Universalität von Parts-of-Speech im Sinne Crofts (2001) ablehnen. Möchte man gleichwohl versuchen, an einer universellen Kategorisierung festzuhalten, so bleibt nur zu postulieren, dass jede Sprache überhaupt semantische Klassen und pragmatische Funktionen scheidet. Darüber, ob auch Wörter und Wortbedeutungen in gleicher Weise geschieden sind, ist hiermit aber nichts gesagt.

Im Folgenden soll der Ansatz von Croft (2001) in der Gegenüberstellung der bisher betrachteten Beispiele aus dem Tongaischen, Hawaiischen, Englischen und Deutschen noch einmal zusammengefasst werden (vgl. Abbildung 2, S. 38), bevor auf Arbeiten eingegangen wird, die darauf aufbauen.

Im Kern der Theorie von Croft (2001) steht die Annahme universeller Kategorien als „typological prototypes“. Diese sind definiert als prototypische Abbildung semantischer Klassen auf pragmatische Funktionen – OBJEKT auf Referenz, EREIGNIS auf Prädikation und EIGENSCHAFT auf Modifikation. Als Identifikationskriterium solcher Prototypen dient die Gegenüberstellung von markiertem und nicht-markiertem Fall. Voraussetzung dafür ist die eindeutige Bestimmbarkeit (der Identität) einer semantischen Klasse. So soll *small* mit einer als EIGENSCHAFT klassifizierten Bedeutung und in modifizierender Phrase auftretend den unmarkierten Fall darstellen gegenüber *small-ness* mit einer ebenso als EIGENSCHAFT klassifizierten Bedeutung, aber in referentieller Phrase auftretend (vgl. Abbildung 2 Mitte rechts). Es handelt sich nun bei *small* und *smallness* um zwei unterscheidbare Stämme, die morphologisch zwei Lexeme, ein adjektivisches Lexem *small* und ein nominales Lexem *smallness* repräsentieren.

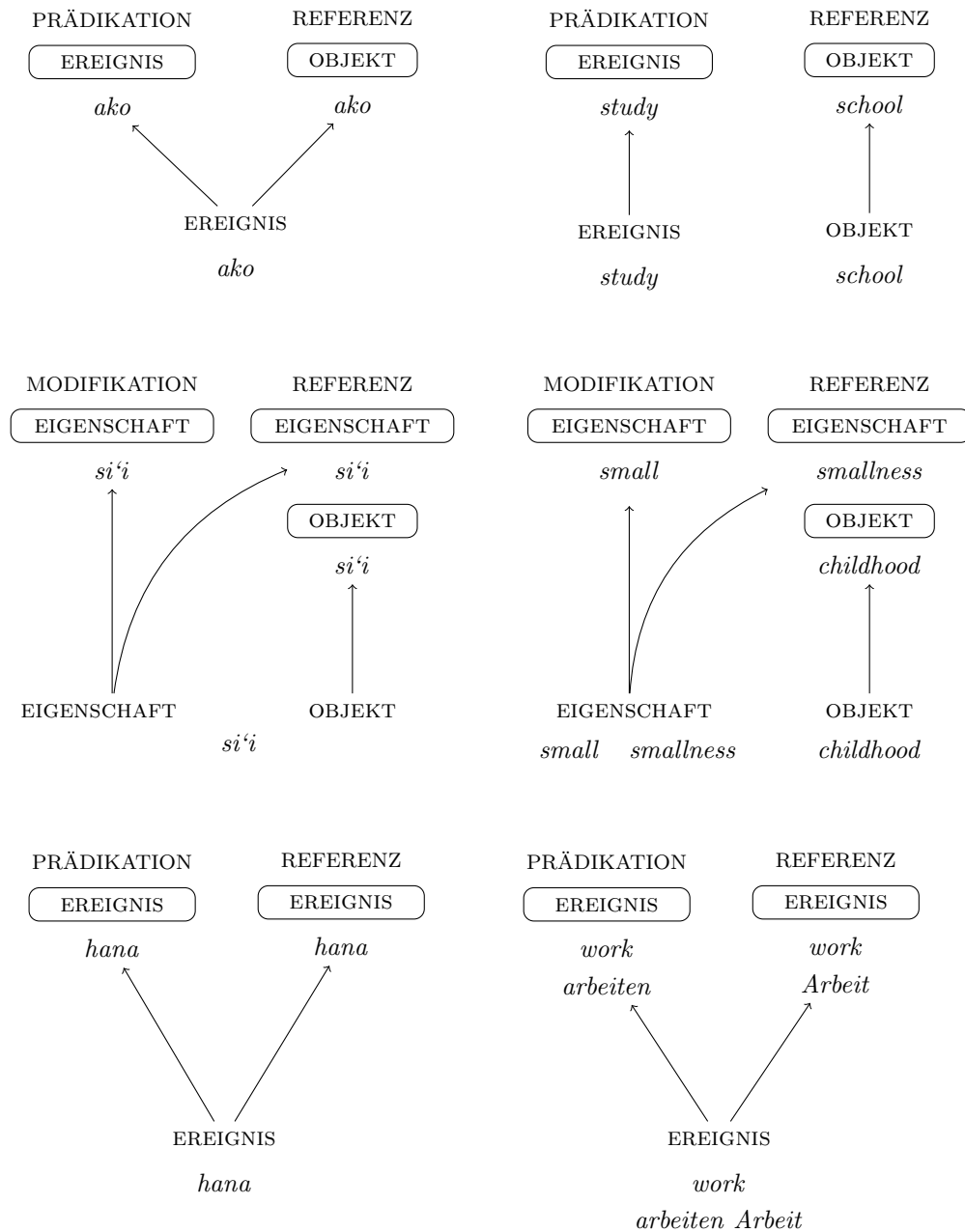


Abbildung 2: Zusammenfassung des Ansatzes von Croft (2001)

Das Adjektiv *small* ist prädisponiert für den Gebrauch in modifizierender Phrase, das Nomen *smallness* für den Gebrauch in referentieller Phrase, unabhängig von der Zuweisung zu einer semantischen Klasse. Der formale Zusammenhang besteht im Wortbildungsprozess der overten Derivation von der Basis *small* zum Derivat *smallness*. Aus funktionaler Perspektive kann man

das so analysieren, dass die semantische Klasse EIGENSCHAFT referenzfähig gemacht wird, sei *smallness* nun konzeptuell-semantisch als EIGENSCHAFT oder OBJEKT im Sinne einer Reifikation zu betrachten. Die von Croft (2001) angesetzten Prototypen können dann als semantisch prototypische Adjektive, Verben oder Nomen interpretiert werden.

Wenn der unprototypische Fall nur mindestens so markiert sein muss wie der prototypische (vgl. S. 36), können dann in gleicher Weise solche Fälle von Konversion, i.e. formale Identität der Stämme unterscheidbarer Lexeme, im Englischen und Deutschen behandelt werden, bei denen dieselbe semantische Klasse in unterschiedlichen pragmatischen Funktionen auftritt. So kann ein semantisch prototypisches Verb *work* von einem semantisch unprototypischen Nomen *work* und eine semantisch prototypisches Verb *arbeiten* von einem semantisch unprototypischen Nomen *Arbeit* unterschieden werden (vgl. Abbildung 2 unten rechts, S. 38).

Solche Fälle von Konversion, bei denen auch die semantischen Klassen in pragmatischen Funktionen unterschieden werden, stellen dann die prototypischen Fälle dar. So können die verbalen Lexeme *study*, *school* usw. als semantisch prototypische Verben, die nominalen Lexeme *study*, *school* usw. als semantisch prototypische Nomen behandelt werden.

Um an universellen Prototypen festhalten zu können, muss Croft (2001) diese aber von der einzelsprachlichen Klassifizierung von nominalen, verbalen und adjektivischen Lexemen abstrahieren. Dies tut er, indem er Fälle wie *work* als einen monosemen Stamm und Fälle wie *study* oder *school* als einen polysemen Stamm analysiert (vgl. Croft 2001: 70 passim) und diese mit der Unterscheidung von *small* und *smallness* identifiziert. Croft (2001) setzt da Konzepte gleich, wo er keine Wörter gleichsetzen kann. Denn er macht keinen Unterschied zwischen einer klassifizierbaren Bedeutung eines Stammes wie bei *work* und einer klassifizierbaren Bedeutung zweier Stämme wie bei *small* und *smallness*. Zudem hält Croft (2001) weiterhin an einzelsprachlich definierten Lexemen fest und spricht genau genommen von einem Stamm *work* und zwei Lexemen *work* (V) und *work* (N) in Analogie zu zwei Stämmen und Lexemen *small* versus *smallness*. Ob Stämme und Lexeme in dieser Weise unabhängig voneinander in universelle und einzelsprachliche Dimensionen verschoben werden können, mag bezweifelt werden, und es ist unklar, ob nun Stämme oder Lexeme prototypische Fälle darstellen.

In der Übertragung auf polynesischen Sprachen sollen dann Wörter wie *hana* als monosem betrachtet werden in Analogie zu Stämmen wie *work* oder *arbeit*

(vgl. Abbildung 2 unten, S. 38), Wörter wie *ako* monosem mit semantischem Wechsel im Auftreten in einer Phrase in Analogie entweder zu *study* (V) versus *study* (N) und *school* (V) versus *school* (N) oder *study* (V) versus *school* (N) (vgl. Abbildung 2 oben, S. 38) und *si'i* als polysem in Analogie zu *small* versus *childhood* (vgl. Abbildung 2 Mitte, S. 38). Es können bei *hana*, *ako* und *si'i* aber weder über die Zuweisung zu formalen Paradigmen Lexeme unterschieden werden, noch können prototypische Fälle anhand der Markierung im Auftreten in einer bestimmten Phrase analysiert werden. Feststellbar ist in den polynesischen Sprachen nur, dass jedes Wort und ebenso jede Bedeutungsklasse in jedem Phrasentyp auftritt, ohne Möglichkeit der formalen Differenzierung. Die Annahme identischer Kategorisierung in den polynesischen Sprachen und Sprachen wie dem Englischen oder Deutschen wäre nicht in der Zuweisung formaler Klassen zu semantischen Klassen begründet, sondern alleine dadurch, dass in der jeweiligen okkasionellen Bedeutung eines Wortes – und *Wort* kann in diesem Vergleich nicht in derselben Weise gemeint sein – überhaupt eine semantische Klassen identifiziert werden kann.

Der Ansatz von Don und Van Lier (2013)

In der Folge von Hengeveld (1992) und Croft (2001) ist vor allem die Diskussion um Monosemie und Polysemie in den polynesischen Sprachen weitergeführt worden. Dabei sind diverse Versuche unternommen worden, die Ansätze zu vereinen, meist in der Annahme einer lexikalisch monosemen Bestimmbarkeit von Wörtern und verschiedener Prozesse der Disambiguierung, ohne indessen eine Kategorisierung in der Art der prototypischen Abdeckung semantischer Klassen auf syntaktisch-funktionale Konstruktionen aufzugeben.

So nehmen Don und Van Lier (2013: 58) für flexible Sprachen im Sinne Hengevelds (1992) (vgl. Kapitel 2.2.1, S. 21) an, „that roots have inherent meanings [...] they have either a basic action-meaning or a basic object-meaning, though this does not imply that they are inherently categorized as either a verb or a noun.“ Die Kategorisierung als Verb, Nomen und Adjektiv wäre lediglich als „matter of syntactic potential“ (Don und Van Lier 2013: 58) zu verstehen, da Wurzeln zwar lexikalisch-semantisch als EREIGNIS, OBJEKT oder EIGENSCHAFT klassifiziert seien, jedoch gleichermaßen in prädikativer, referentieller und modifizierender Phrase auftreten würden. Hier stellt sich zuerst wieder die Frage, was genau unter syntaktischen Kategorien Verb, Nomen und Adjektiv zu verstehen ist (vgl. auch Kap. 2.2.2, S. 14–16).

Überdies ließe sich Don und Van Lier (2013: 58) zufolge in Analogie zu den von Croft (2001) angesetzten „typological prototypes“ (vgl. S. 27) von prototypischen Kategorien sprechen. *Prototypisch* heißt in diesem Fall nichts anderes, als dass die lexikalisierte semantische Klasse einer Wurzel mit der semantischen Klasse im Auftreten dieser Wurzel in einer Phrase übereinstimmt, und zwar in der spezifischen Weise, dass eine OBJEKT-klassifizierte Wurzel im Auftreten in referentieller Phrase OBJEKT-klassifiziert ist, eine EREIGNIS-klassifizierte Wurzel im Auftreten in prädikativer Phrase EREIGNIS-klassifiziert und eine EIGENSCHAFT-klassifizierte Wurzel im Auftreten in modifizierender Phrase EIGENSCHAFT-klassifiziert ist. *Nicht-prototypisch* wäre hingegen, wenn z.B. eine EREIGNIS-klassifizierte Wurzel in referentieller Phrase oder eine OBJEKT-klassifizierte Wurzel in prädikativer Phrase auftritt. In solchen Fällen analysieren Don und Van Lier (2013: 57) „syntactic derivation“ und meinen damit, dass die Bedeutung einer EREIGNIS-klassifizierten Wurzel in referentieller Phrase oder einer OBJEKT-klassifizierten Wurzel in prädikativer Phrase regelhaft über den Gebrauch in der Phrase und nicht über die Lexikalisierung erklärt werden kann (vgl. auch das Kriterium der Kompositionalität bei Evans und Osada 2005: 367–386).

Im Samoanischen sei beispielsweise die Wurzel $\sqrt{u\bar{o}}$ semantisch inhärent als OBJEKT mit der Bedeutung „friend“ klassifiziert und würde somit prototypisch in derselben Bedeutung bzw. Bedeutungsklasse in referentieller Phrase auftreten. Demgegenüber wäre im unprototypischen Fall die Bedeutung von $u\bar{o}$ mit „to be friends“ im Auftreten in prädikativer Phrase nicht lexikalisiert, sondern regelhaft über das Auftreten in der Phrase erklärt. Don und Van Lier (2013: 77) beziehen sich hier auf ein Beispiel von Mosel und Hovdhaugen (1992: 77), vgl.:¹³

- (6a) *E uō Tanielu ma Ionatana.*
 TAM friend Daniel and Jonathan
 ‘Daniel and Jonathan are friends.’
 (Mosel und Hovdhaugen 1992:77)

¹³Ich orientiere mich hier an den Glossen von Mosel und Hovdhaugen (1992: 77).

- (6b) *E alofa Tanielu i l=a=na uō.*
 TAM love Daniel DIR ART=POSS=3.SG friend
 ‘Daniel loves his friend.’
 (Mosel und Hovdhaugen 1992:77)

Zum Tragen käme hier die Regel [to be X] mit [X] als lexikalisierte Bedeutung der Wurzel. Bei einer EREIGNIS-klassifizierten Wurzel wie \sqrt{alu} „go“, die in dieser Bedeutung prototypisch in prädikativer Phrase auftritt, käme mit der Bedeutung „going“ im unprototypischen Auftreten in referentieller Phrase die Regel [X-ing] oder [the act of X-ing] zur Anwendung mit [X] als lexikalisierte Bedeutung, vgl.:¹⁴

- (7a) *E alu le pasi i Apia*
 TAM go ART bus DIR Apia
 ‘The bus goes to Apia’
 (Don und Van Lier 2013: 77)

- (7b) *le alu o le pasi i Apia*
 ART go POSS ART bus DIR Apia
 ‘the going of the bus to Apia’
 (Don und Van Lier 2013: 77)

Da mit diesen syntaktisch-semantischen Ableitungsregeln nicht alle Bedeutungsunterschiede im Auftreten von Wurzeln in den verschiedenen Phrasentypen erklärt werden können, nehmen Don und Van Lier (2003: 78–80) an, dass daneben in Fällen von „idiosyncratic semantic shift“ (Don und Van Lier 2013: 78) Nullderivation analysiert werden muss, so z.B. bei *tusi* mit der Bedeutung „write“ in prädikativer Phrase und der Bedeutung „letter“ in referentieller Phrase. Bei *tusi* würde es sich demzufolge um zwei Wurzeln handeln, eine Wurzel $\sqrt{tusi_1}$ „write“ mit lexikalisierte EREIGNIS-klassifizierte Bedeutung, die prototypisch in prädikativer Phrase auftritt und als Ergebnis von Nullderivation eine Wurzel $\sqrt{tusi_2}$ „letter“ mit lexikalisierte OBJEKT-klassifizierte

¹⁴Don und Van Lier (2013: 77) machen keine genauen Angaben, woher diese Beispiele stammen. Sie finden sich aber in Mosel und Hovdhaugen (1992: 73), mit dem Unterschied der Übersetzung von (7b) mit „the fact that the bus goes to Apia“ und der späteren Feststellung von Mosel und Hovdhaugen (1992: 77): „Until now we have not, for instance, found *alu* ‘go’ in a nominal function [...]“. Ich habe in diesem Fall die Glossen von Don und Van Lier (2013) beibehalten.

Bedeutung, die prototypisch in referentieller Phrase auftritt (vgl. Don und Van Lier 2013: 79).

Für die Beschreibung polynesischer Sprachen sehe ich in der Analyse von Don und Van Lier (2013) keine wesentliche Veränderung zum Ansatz von Croft (2001). Letztlich geht es beiden darum, die prototypische Abdeckung semantischer Klassen auf pragmatische Funktionen zu zeigen und dabei die Festlegung auf einzelsprachliche Wortartkategorien zu umgehen, nur dass die Zusammenhänge von prototypischen und nicht-prototypischen Fall teils unterschiedlich analysiert werden. Während Don und Van Lier (2013) Fälle wie *uō* und *alu* als semantisch-syntaktische Ableitung behandeln, würde Croft (2001) bei *uō* entweder Polysemie oder semantischen Wechsel im Auftreten in prädikativer Phrase ansetzen, wobei ich keinen Unterschied zwischen „syntactic derivation“ bei Don und Van Lier (2013: 57) und „semantic shift“ bei Croft (2001: 73 und passim) erkennen kann; *alu* wäre Croft (2001) folgend sowohl im Auftreten in referentieller als auch prädikativer Phrase als EREIGNIS-klassifiziert. Während Don und Van Lier (2013) in Fällen wie *tusi* Nullderivation ansetzen, analysiert Croft (2001) auch solche als Polysemie oder semantischen Wechsel.

Erstens stehen Don und Van Lier (2013) damit ähnlich wie Croft (2001) in der Analyse von semantischem Wechsel oder Vonen (2000) in der Konversionsanalyse (vgl. Kap. 2.2.1, S. 13–14) vor dem Problem, bei jeder Wurzel, die sie als deriviert betrachten, eindeutig die Derivationsrichtung angeben zu können. So wären z.B. bei *fana* „shoot“ und „gun“ oder bei *gaoi* „steal“ und „thief“ (vgl. Don und Van Lier 2013: 79) beide Richtungen vorstellbar.

Zweitens erscheint die Trennung zwischen regelhafter syntaktischer Ableitung und idiosynkratischer Nullderivation willkürlich. Das wird deutlich am Beispiel von *tusi*. Denn gerade den analogen Fall im Tongaischen *tohi* mit der Bedeutung „write“ in prädikativer Phrase und „book“ oder „letter“ in referentieller Phrase sieht Völkel (2014: 461) als Beleg für die Annahme, *tohi* sei lexikalisch-semantisch EREIGNIS-klassifiziert und würde die Bedeutung regelhaft im Auftreten in referentieller Phrase ändern, etwa in der Regel [product of X] oder [result of X] mit [X] als lexikalisierte Bedeutung. Völkel (2014) erkennt also nicht alleinig die Zusammenhänge $[X]_{\text{OBJEKT}} \rightarrow [\text{to be X}]$ und $[X]_{\text{EREIGNIS}} \rightarrow [X\text{-ing}]$ als regelhaft an, sondern auch solche, die Don und Van Lier (2013) als idiosynkratisch interpretieren, wie *tusi* oder auch *gaoi* „steal“ und *gaoi* „thief“ = [the one who steals].

Ich sehe hier das Hauptproblem in der Trennung einer als universell angesetzten semantischen Ebene von einzelsprachlichen Mitteln bei gleichzeitiger Ver-

mischung der semantisch-konzeptuellen Interpretation durch Paraphrasierung mit sprachlichen Mitteln einer Sprache wie dem Englischen mit den tatsächlich gegebenen sprachlichen Mitteln des Samoanischen, Tongaischen oder Hawaiischen.

Wenn der semantische Wechsel von *alu* in referentieller Phrase in (7b) (S. 42) sprachlich die identische Strategie darstellt wie Nominalisierung von Verben im Englischen mit *-ing* oder Nominalisierung von Verben im Deutschen mittels Infinitivform, dann könnte *tusi* in der Bedeutung „letter“ oder „book“ auch als „writing“ oder „Schreiben“ gelesen werden. Auch wenn es sich bei *working* und *Schreiben* um Nomen handelt und man vielleicht noch in Analogie bei *tusi* Nullderivation analysieren wollte, könnte man z.B. bei *hana* im Hawaiischen im Auftreten in referentieller Phrase in (1b) (S. 1) nicht mehr zwischen „working“ bzw. „Arbeiten“ als regelhafte syntaktische Ableitung und *work* bzw. *Arbeit* als Nullderivation unterscheiden.

Die semantischen Differenzierungen im Samoanischen oder Hawaiischen werden mit den formalen Mitteln anderer Sprachen nicht adäquat beschrieben.

Zwar mag man z.B. *uō* in referentieller Phrase als „friend“ und in prädikativer Phrase als „to be friends“ übersetzen können. Hält man sich aber an die sprachlichen Mittel des Samoanischen, ist damit bei *uō* gar keine semantische Regel $[X] \rightarrow [to\ be\ X]$ mit lexikalisch-semantisch als OBJEKT klassifiziertem $[X]$ gezeigt. Es ist einfach nicht dasselbe, den Zustand des Befreundetseins zwischen Daniel und Jonathan in Satz (6a) (S. 41) mit Nomen in Kopulakonstruktion auszudrücken wie im Englischen oder durch das Auftreten des Wortes *uō* in Tempus-Aspekt-Modus-Syntagma wie im Samoanischen. Im Englischen oder Deutschen gibt es auch die präfigierten Verben *befriend* und *befreunden* – entsprechend möglich wäre eine Übersetzung in Satz (6a) (S. 41) auch mit Partizipialkonstruktion „Daniel und Jonathan sind befreundet.“ Deswegen ist aber *uō* lexikalisch-semantisch nicht identisch mit dem deutschen Partizip *befreundet*. Genauso könnte man, und genauso falsch wäre dies, mit der Diachronie des Deutschen argumentieren, dass *uō* lexikalisch semantisch als EREIGNIS klassifiziert ist, weil es sich bei *Freund* etymologisch um eine Partizipialbildung zu germanisch **frij-ō-* „freundlich behandeln, umwerben“ (vgl. Kluge 2002: 316 [1883]) handelt.

Zusammenfassend darf bezweifelt werden, dass Don und Van Lier (2013) die Trennung zwischen regelhafter syntaktischer Ableitung und Nullderivation konsequent am Samoanischen aufrecht erhalten können. Bezeichnend dafür ist auch, dass sie annehmen, es müssten theoretisch Fälle gefunden werden, bei

denen eine derivierte Wurzel auch in regelhafter Ableitung vorkommt, also z.B. $\sqrt{\text{tusi}_1}$ „write“ regelhaft in referentieller Phrase mit der Bedeutung „writing“ oder die Wurzel $\sqrt{\text{tusi}_2}$ „letter“ regelhaft in prädikativer Phrase mit der Bedeutung „to be a letter“ (vgl. Don und Van Lier 2013: 80). Genau dafür liefern sie aber keinerlei empirischen Belege.

Im Kern gelingt Don und Van Lier (2013) genauso wenig wie Croft (2001), der das Kriterium der typologischen Markiertheit in den polynesischen Sprachen nicht anwenden kann, in klarer Weise prototypische von nicht-prototypischen Fällen abzugrenzen und schließlich mehr zu sagen, als dass alle Wörter in allen Bedeutungsklassen in allen syntaktisch-funktionalen Konstruktionen auftreten. Mit der Interpretation des Zusammenhangs zwischen den semantischen Klassen der Bedeutungen einzelner Wurzeln wird schließlich nichts über deren Auftreten im Satz ausgesagt.

Der Ansatz von Völkel (2017)

Einen etwas anderen Ausgangspunkt als Croft (2001) und Don und Van Lier (2013) wählt Völkel (2017) in ihrer Untersuchung zur Flexibilität der Wörter im Tongaischen. Zugrunde liegt die Frage, ob alle Wörter im Tongaischen gleichermaßen in allen syntaktischen Kontexten auftreten können (vgl. Völkel 2017: 457) – eine Frage, die in vielen Arbeiten unbeachtet bleibt – und damit verbunden: „Do Tongan native speakers mentally store different meanings of a particular item (i.e. polysemy) or one meaning that is specified by the syntactic or pragmatic context (i.e. monosemy)?“ (Völkel 2017: 461). Völkel (2014) lehnt die Polysemie-Analyse von Croft (2001) ab. Vielmehr seien Völkel (2017: 447 und passim) alle Wörter in ihrer Bedeutung lexikalisch monosem bestimmbar und entsprechend kategorisiert als „object-denoting terms“, die prototypisch in referentieller Phrase oder genauer in syntaktischem Slot mit Artikel auftreten, „event-denoting terms“, die prototypisch in prädikativer Phrase auftreten oder genauer in syntaktischem Slot mit Tempus-Aspekt-Modus-Marker (vgl. die Kritik Völkels (2017: 448 Anmerkung 4) an Hengeveld (1992), Kapitel 2.2.3, S. 19) und „property-denoting terms“, die prototypisch in modifizierender Phrase auftreten. Da die Bedeutung eines Wortes in nicht-prototypischer Phrase immer regelhaft aus dem syntaktischen Kontext erklärt werden könne, muss hier im Gegensatz zu Don und Van Lier (2013) nicht syntaktisch-semantische Ableitung von Nullderivation unterschieden werden (vgl. das Beispiel *tohi*, S. 43).

Völkel (2017) hat nun untersucht, ob sich je nach zugeschriebener semantischer Klasse der Bedeutung eines Wortes das Verhalten im Auftreten in Phrasen unterscheidet. Dazu hat sie in einer Feldstudie vier Muttersprachlern verschiedenen Alters (20 – 60), verschiedenen Geschlechts (zwei männlich, zwei weiblich) und in verschiedenen Distrikten der Insel Tongatapu lebend 84 Wörter vorgelegt, zu denen diese selbst, falls möglich, Sätze oder Phrasen (TAM X, ART X, TAM Y X, ART Y X) bilden sollten. In vereinzelt Fällen wurden den Informanten zusätzlich Beispiele aus der Studie von Broschart (1995) zur Beurteilung vorgelegt. Die Auswahl der Wörter war vorrangig von der Einbindung möglichst vieler semantischer Subklassen motiviert (vgl. Völkel 2017: 464). Völkel (2017: 464) gruppiert die Wörter schließlich nach ihrer semantischen Klasse (17 „property denoting terms“, 27 „event-denoting terms“ und 40 „object-denoting terms“) und vergleicht, in welchen der syntaktischen Konstruktionen (TAM X, ART X, TAM Y X, ART Y X) diese laut Informanten und vereinzelt angereichert durch Beispiele aus Churchward (1953, 1959) und Taumoefolau (1998) auftreten können. Diese Ergebnisse vergleicht sie zusätzlich mit Angaben in den Wörterbüchern von Churchward (1959) und Thompson und Thompson (1992) (vgl. Völkel 2017: 477–494 Appendix). In Tabelle 6 sind die Ergebnisse aus Völkels Studie in komprimierter Form zusammengefasst, vgl.:

semantische Klasse	syntaktischer Slot			
	ART X	TAM X	ART X Y	TAM X Y
OBJEKT	100% (40/40)	48% (19/40)	70% (28/40)	23% (9/40)
EREIGNIS	100% (27/27)	100% (27/27)	81% (22/27)	19% (5/27)
EIGENSCHAFT	100% (17/17)	100% (17/17)	100% (17/17)	35% (6/17)

Tabelle 6: Zusammenhang von semantischer Klasse und Flexibilität eines Wortes (vgl. Völkel 2017: 465)

Wie in Tabelle 6 zu sehen, treten zwar Wörter aller semantischen Klassen in allen syntaktischen Slots auf, allerdings nicht mit gleicher Häufigkeit. Dies interpretiert Völkel (2017: 465) als Zusammenhang zwischen semantischer Klasse und der Flexibilität eines Wortes, ausgedrückt in folgender Hierarchie von flexibel zu weniger flexibel: „property-denoting terms“ > „event-denoting terms“ > „object-denoting terms“. Das bedeutet, dass ein Wort, dessen lexikalisierte Bedeutung als EIGENSCHAFT klassifiziert ist, ein größeres Potential hat, nicht nur in seinem prototypischen, sondern auch nicht-prototypischen syntaktischen Slot aufzutreten als ein Wort, dessen lexikalisierte Bedeutung als

EREIGNIS oder OBJEKT klassifiziert ist.

Ich sehe die Untersuchung von Völkel (2014) aus zwei Gründen kritisch. Als Erstes stellt sich aus methodischer Sicht die Frage nach der Repräsentativität der Ergebnisse. Bei nur vier Sprechern, die die gesuchten syntaktischen Konstruktionen selbst produzieren sollten, mag an einer generalisierbaren Tendenz gezweifelt werden. Es ist dann auch wenig überraschend, dass die Ergebnisse aus der Befragung teilweise von denen von Churward (1953, 1959), Thompson und Thompson (1992), Broschart (1995) und Taumoefolau (1998) abweichen. Völkel (2017) ist sich dessen durchaus bewusst. So bleibt ihr nur zu vermuten, „[...] that potentially some lexemes are even more flexible than considered by the informants“ (Völkel 2017: 464) und „[...] that this list of possible occurrences may not yet be complete or may contain examples that are not accepted by all native speakers.“ (Völkel 2017: 465).

Als zweiten und viel schwerwiegenderen Grund zur Kritik betrachte ich die theoretischen Annahmen, die Völkel (2017) voraussetzt, aber mit ihrer Studie selbst nicht belegen kann. Eine Monosemie-Analyse weist im Kern dieselben Schwachstellen wie der Ansatz von Croft (2001) oder Don und Van Lier (2013) auf: die eindeutige Bestimmung der semantischen Klasse der Bedeutung eines Wortes. Völkel (2017: 460) selbst sieht das skeptisch: „[...] the idea of monosemous lexemes raises problems regarding the assignment of a lexical item to one specific semantic class.“

Völkel (2017: 478 Anmerkung 31) diskutiert zum Beispiel die Frage, ob *fefine* „woman“ und „female“ so wie von Churchward (1953: 18) als Nomen, i.e. semantisch OBJEKT, oder als „property-denoting term“ klassifiziert werden soll. Sie entscheidet sich für letzteres, da in Analogie zu anderen „property-denoting terms“ in Tempus-Aspekt-Modus-Syntagma [to be X] und in Artikel-Syntagma [the being X], [X-ness] oder [a X Person] analysiert werden könne und folgert: „Thus, in nominal constructions it means ‘a female person’ and that is a woman.“ (Völkel 2017: 478). Völkel (2017: 478) übersieht hier zunächst das englische Nomen *female* und die mögliche Übersetzung von *fefine* in Artikel-Syntagma „a female“, womit die Regel [X] → [a X Person] gar nicht anwendbar wäre. Letztlich ist man bei dem Beispiel von *fefine* genauso wie bei der Frage des semantischen Wechsels (vgl. Croft 2001: 73 und passim) oder der Null-Derivation (vgl. Don und Van Lier 2013: 73 und passim) vor das Problem gestellt, die Richtung anzugeben und dabei nicht dem Trugschluss zu unterliegen, anhand der sprachlichen Reformulierung im Englischen die Konzeptualisierung im Tongaischen nachweisen zu können (vgl. S. 43–44), wobei

dies sogar im Englischen mit *female* (A) und *female* (N) schwierig würde. So müsste man sich im Falle von *fefine* für die Analyse Churchwards (1953: 18) entscheiden, würde man das Deutsche als Metasprache verwenden. Dann wäre *fefine* in modifizierender Phrase mit „weiblich“ zu übersetzen und damit Derivat der Basis *Weib*. In gleicher Weise wie Völkel (2017: 478 Anmerkung 31) behauptet, *fefine* sei mit „female“ monosem als EIGENSCHAFT zu bestimmen, könnte es mit „Weib“ monosem als OBJEKT gesehen werden. Wenn *fefine* in Artikel-Syntagma eine Person bezeichnet und in modifizierender Phrase eine Eigenschaft und ein semantischer Zusammenhang erkennbar ist, kann ohne vorhandene sprachliche Mittel einfach nicht eindeutig entschieden werden, von welcher Bedeutung aus in welche Richtung ein Wechsel geht und was die monoseme, lexikalisierte Bedeutung des Wortes ist.

An dieser Stelle tritt das wesentliche Problem in Völkels (2017) Analyse zu Tage: der undeutliche Zusammenhang von Monosemie, Prototypikalität und regelhafter Ableitung. Völkel (2017: 460 Anmerkung 16) ist sich beispielsweise ähnlich zu *fefine* unsicher bei der Zuweisung der semantischen Klassen von *kata* und *'uha*. So könne *kata* in der Bedeutung „laugh“ als prototypisch EREIGNIS gelesen werden, in der Bedeutung „laughter“ allerdings als „less prototypical object (i.e. an abstract concept)“ und es müsse daraus geschlossen werden: „In this case, one would generally agree to classify this term as event-denoting according to its most prototypical meaning.“ (Völkel 2017: 460 Anmerkung 16). Hingegen würde *'uha* in der Bedeutung „to rain“ ein prototypisches Ereignis denotieren, in der Bedeutung „rain“ ein prototypisches Objekt und damit wäre die monoseme Bedeutung von *'uha* unklar (vgl. Völkel 2017: 460 Anmerkung 16). Völkel (2017) versteht Prototypikalität hier als die prototypische semantische Klasse eines Wortes im aktuellen Auftreten z.B. EREIGNIS bei *'uha* in Tempus-Aspekt-Modus-Syntagma oder OBJEKT bei *'uha* in Artikel-Syntagma und setzt diese mit der prototypischen lexikalisierten semantischen Klasse des Wortes und Monosemie gleich. Deswegen hat sie Schwierigkeiten bei der Bestimmung von *'uha*. Nur wäre dann *fefine* genauso wenig als monosem bestimmbar, da es in modifizierender Phrase als prototypisch EIGENSCHAFT und analog zu *'uha* in Artikel-Syntagma als prototypisch OBJEKT gesehen werden muss. Die Bedeutung „woman“ kann nicht als prototypisch EIGENSCHAFT betrachtet werden, auch nicht durch Paraphrasierung mit „the female person“. Die Anwendung der Regel [X] → [the X person] macht [the X person] nicht als prototypisch EIGENSCHAFT analysierbar. Man kann im Deutschen auch ein Nomen Agentis wie *Sprecher* nicht semantisch der Klasse EREIGNIS zuordnen,

weil eine Paraphrasierung der Bedeutung von *Sprecher* mit „die sprechende Person“ oder „derjenige, der spricht“ möglich ist. Man kann *fefine* auch als Reifikation (vgl. Schmid 2016: 172 und das Modell von Schulze 2012, S. 29–30) bzw. Personifizierung lesen und damit als OBJEKT analysieren. Im Sinne Crofts (2001) wäre *fefine* gerade ein Beispiel für den semantischen Wechsel von einer Eigenschafts- zu einer Objektbedeutung, und die Monosemie-Analyse Völckels (2017) sagt nur, dass die Eigenschaftsbedeutung lexikalisiert ist, während die Objektbedeutung regelhaft über den syntaktischen Gebrauch erklärt werden kann. Wird also die monoseme Bedeutung eines Wortes über seine (prototypische) Bedeutung in jeweiligem syntaktischen Slot analysiert, ist zwar *kata*, aber weder *fefine* noch *‘uha* monosem. Wird die monoseme Bedeutung eines Wortes über eine Regel und ihre Richtung analysiert, ist *fefine*, aber weder *kata* noch *‘uha* monosem, es sei denn, man würde auch bei *kata* und *‘uha* eine regelhafte Ableitung annehmen, etwa mit Don und Van Lier (2003: 77) [X] → [the X-ing], bei *‘uha* „rain“ → „the raining“, bei *kata* „laugh“ → „laughing“. Dies führt letztlich wieder zur Ausgangsfrage der Regelhaftigkeit und Richtung eines – falls überhaupt analysierbaren – semantischen Wechsels.

Prototypikalität soll bei Völkel (2017) nun im weiteren Sinne verstanden werden als die prototypische Abdeckung einer semantischen Klasse auf eine pragmatische Funktion bzw. einen syntaktischen Slot. Nach dem bisher Besprochenen kann dies in zweierlei Hinsicht verstanden werden. Wäre erstens Monosemie definiert durch die Zuweisung derselben semantischen Klasse der lexikalisierten Bedeutung eines Wortes in allen syntaktischen Slots, z.B. bei *kata* EREIGNIS, wäre der prototypische Fall *kata* in Tempus-Aspekt-Modus-Syntagma, der unprototypische *kata* in Artikel-Syntagma. Wäre zweitens Monosemie definiert durch die Anwendbarkeit einer Regel, z.B. bei *fefine* als EIGENSCHAFT mit regelhafter Ableitung zu OBJEKT, wäre der prototypische Fall das Auftreten von *fefine* in modifizierender Phrase, der nicht-prototypische Fall das Auftreten von *fefine* in Artikel-Syntagma. Wenn aber ein Wort unabhängig von seiner monosem klassifizierten Bedeutung in jeder Bedeutung in jedem syntaktischen Slot auftreten kann – wie auch die Ergebnisse aus der Untersuchung (vgl. Tabelle 6, S. 46) andeuten, die zeigen, dass z.B. alle „property-denoting terms“ in allen syntaktischen Slots und alle „event-denoting term“ in sowohl Tempus-Aspekt-Modus- als auch Artikel-Syntagma auftreten – entsteht bei der Unterscheidung von prototypischem und nicht-prototypischem Fall die Frage nach einem geeigneten Kriterium dafür. Völkel (2017: 474) bietet dann auch zwei Parameter zur Überprüfung von Prototypikalität an, indem sie folgendes

vermutet:

- (1) In accordance with its semantics, a flexible lexeme is most frequently used in its prototypical syntactic context.
- (2) In accordance with its semantics, the prototypical syntactic context of flexible lexemes is cognitively the most present one in the mind of its speakers.

Damit stellt Völkel (2017) aufbauend auf einer Monosemie-Analyse im Tongaischen eine (quantitative) Prädispositionshypothese auf und rückt in die Nähe von Ansätzen in einzelsprachlichen Grammatiken wie der von Elbert und Pukui (1979), in denen angenommen wird, dass ein Wort je nach semantischer Klassifizierung eine bestimmte Distribution aufweist (vgl. Kap. 2.2.1, S. 12–13). Um die Klassifizierung flexibel auftretender Wörter als monosem miteinzubeziehen, ist die Distributionshypothese dann aber unterteilt in erstens die Annahme, Wörter würden entsprechend der Zugehörigkeit zu einer semantischen Klasse nur in einem¹⁵ oder zumindest nicht allen syntaktischen Slots auftreten und zweitens in die Annahme, Wörter würden, auch wenn sie in mehr als einem syntaktischen Slot auftreten, entsprechend der Zugehörigkeit zu einer semantischen Klasse häufiger in dem mit der semantischen Klasse assoziierten Slot auftreten (vgl. Abbildung 3, S. 51).

In der Prädispositionshypothese von Völkel (2017) sollen Wörter, deren Bedeutung als OBJEKT klassifiziert werden kann, immer in Artikel-Syntagma, und wenn in mehr als einem Syntagma, häufiger in Artikel-Syntagma und eher auch in modifizierender Phrase als in Tempus-Aspekt-Modus-Syntagma auftreten (vgl. Abbildung 3 oben, S. 51). Wörter, deren Bedeutung als EREIGNIS klassifiziert werden kann, sollen hingegen immer in Tempus-Aspekt-Modus-Syntagma, und wenn in mehr als einem Syntagma, häufiger in Tempus-Aspekt-Modus-Syntagma und eher auch in Artikel-Syntagma als in modifizierender Phrase auftreten (vgl. Abbildung 3 Mitte, S. 51). Wörter schließlich, deren Bedeutung als EIGENSCHAFT klassifiziert werden kann, sollen gleichermaßen in allen Syntagmen auftreten, in modifizierender Phrase allerdings häufiger.

Da Völkel (2017) erstens die Repräsentativität ihrer Untersuchung selbst skeptisch sieht, bleibt ihr nur, auf zukünftige Korpusstudien zu verweisen (vgl. Völkel 2017: 465 und passim), um ein vollständiges Bild über die Distribution von Wörtern zu erhalten. Da Völkel (2017) zweitens die vorausgesetzte, aber nicht belegte Monosemie-Hypothese im Einzelnen Probleme bereitet, müsste in einer Korpusstudie zuerst untersucht werden, ob überhaupt eine

¹⁵Bei Völkel (2017: 477–494 Appendix) z.B. finden sich neun als „object-denoting terms“ klassifizierte Wörter, die ausschließlich in Artikel-Syntagma auftreten.

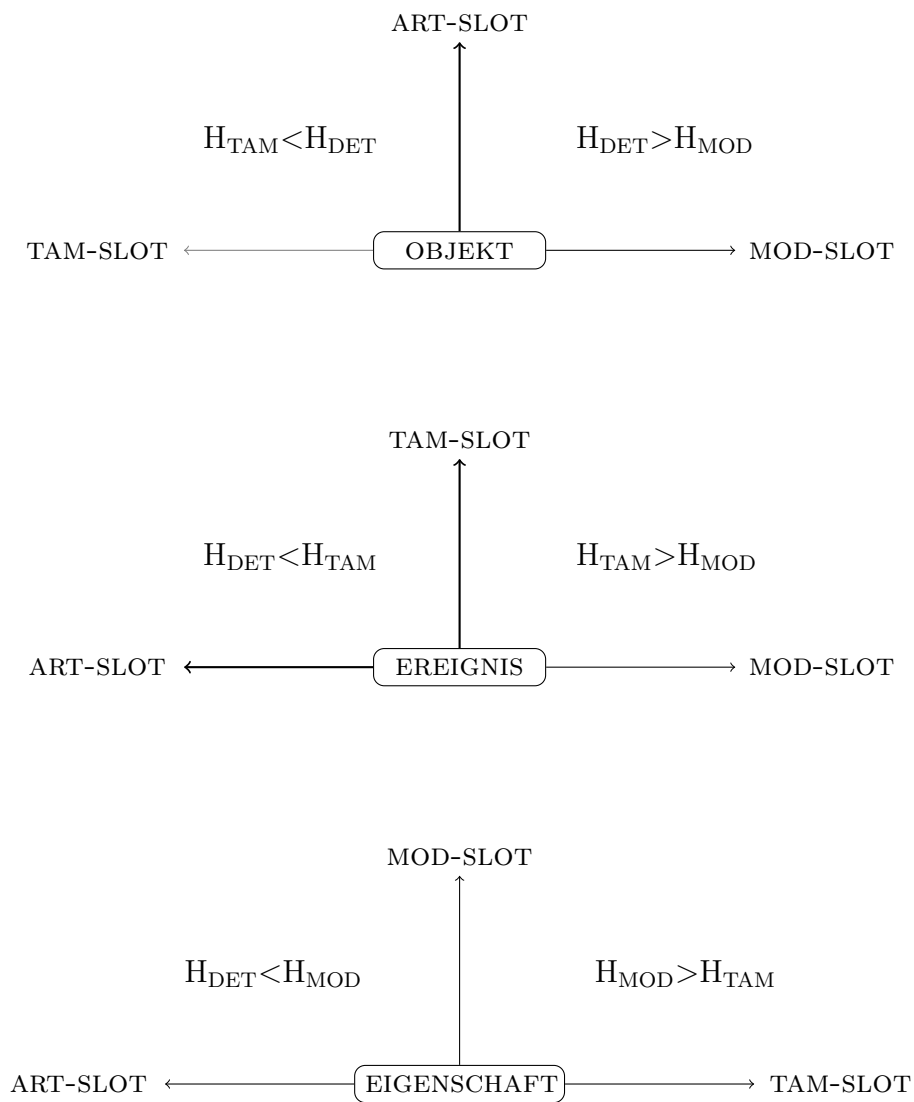


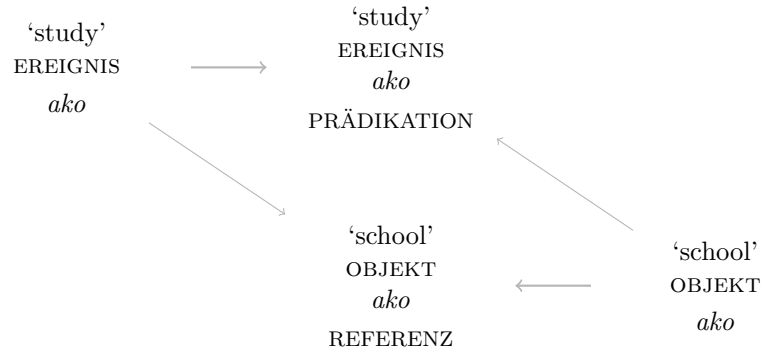
Abbildung 3: Prädisposition semantischer Klassen

allgemeine Tendenz in der Distribution von Wörtern, sei es im Auftreten in bestimmten Slots oder in Frequenzklassen, feststellbar ist, bevor Wörter semantisch klassifiziert werden. So steht bezogen auf eine Wortartkategorisierung noch vor der Frage, ob die Bedeutung von Wörtern im Auftreten in einem Syntagma aus ihrer lexikalisierten Bedeutung erklärt werden kann, die Frage, ob Wörter überhaupt prädisponiert sind für das Auftreten in einer syntaktisch-funktionalen Konstruktion, das heißt Distributionsklassen analysierbar sind. Das betrifft aber nicht nur die Untersuchung von Völkel (2017), sondern in gleicher Weise die Arbeiten von Croft (2001) und Don und Van Lier (2013). Deswegen sollen als nächstes die Ansätze von Croft (2001), Don und Van Lier (2013) und Völkel (2017) noch einmal unter dem Aspekt einer Prädispositionshypothese und der Polysemie-Monosemie-Diskussion in den polynesischen Sprachen zusammengefasst werden.

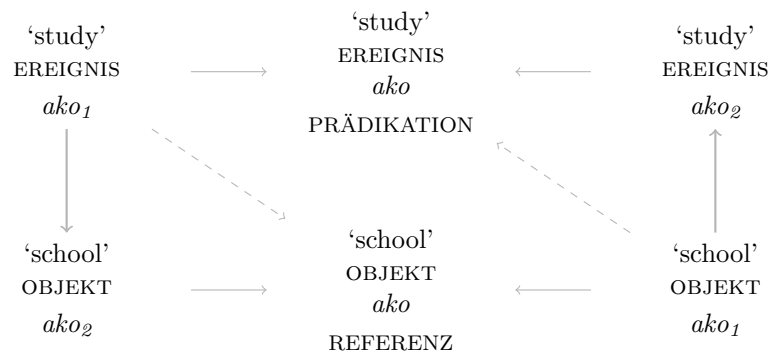
Vergleich von Croft (2001), Don und Van Lier (2013) und Völkel (2017)

Betrachtet man in der Frage der Prädisposition von Wörtern für syntaktisch-funktionale Konstruktionen vergleichend die Ansätze von Croft (2001), Don und Van Lier (2013) und Völkel (2017), so muss man fragen, ob die jeweiligen Versuche der semantischen Klassifizierung von Wörtern zu einer wesentlich verschiedenen Beschreibung polynesischer Sprachen kommen. Das soll am Beispiel von *ako* aus dem Tongaischen erläutert werden (vgl. Abbildung 4, S. 53). Croft (2001), Don und Van Lier (2013) und Völkel (2017) teilen als Voraussetzung die eindeutige Bestimmbarkeit der semantischen Klasse der Bedeutung eines Wortes im aktuellen Auftreten in einer pragmatisch-funktional definierten Phrase bzw. einem syntaktischen Slot. In einem Fall wie *ako* mit „study“ und „school“ mag dies weniger problematisch erscheinen. In anderen Fällen hingegen umso mehr. So ist unklar, ob *si'i* im Tongaischen in referentieller Phrase bzw. in Artikel-Syntagma semantisch als EIGENSCHAFT oder OBJEKT analysiert werden soll. In einer Monosemie-Analyse ist nicht zu entscheiden, welche der Bedeutungen aus „study“ und „school“ die lexikalisierte ist (vgl. Abbildung 4 Punkt I links und rechts, S. 53). Notwendig wäre dafür die Analysierbarkeit der Richtung eines semantischen Wechsels. Vorstellbar wäre sowohl die Schule als Ort, an dem gelernt wird, als auch das Lernen als Tätigkeit, die sich in der Schule ereignet. Ohne overte Markierung bleibt nur die Spekulation darüber. In der Folge kann dann weder gezeigt werden, dass der seman-

I. Monosemie



II. Monosemie + Null-Derivation



III. Semantischer Shift + Polysemie

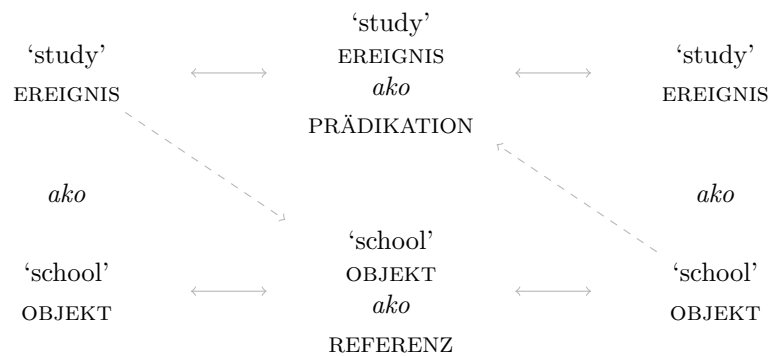


Abbildung 4: Vergleich von Völkel (2017), Don und Van Lier (2013) und Croft (2001)

tische Wechsel über eine Ableitungsregel syntaktisch zustande kommt, noch welche Bedeutungsklasse prototypisch auf welchen syntaktischen Slot abgebildet wird. Was bleibt, ist die eingangs gemachte Feststellung, *ako* nur im aktuellen Auftreten in verschiedenen syntaktischen Slots semantisch differenzieren zu können.

Don und Van Lier (2013) setzen ebenfalls Monosemie an, ebenfalls in zweifacher Weise: einmal analog zu Völkel (2017) mit regelhaftem syntaktischen Bedeutungswechsel und einmal analog zu Vonen (2000) in der Annahme zweier monosemer Wurzeln, genauer einer Basis und einer null-derivierten Wurzel (vgl. Abbildung 4 Punkt II, S. 53). Damit schaffen sich Don und Van Lier (2013) auch in zweifacher Weise Probleme. Erstens können Don und Van Lier (2013) genauso wenig wie Völkel (2017) die Richtung des semantischen Wechsels bestimmen, sei es nun mittels syntaktischer Ableitungsregeln oder über Null-Derivation. Zweitens können Don und Van Lier (2013) nicht immer eindeutig zwischen diesen beiden unterscheiden (vgl. Abbildung 4 Punkt II links und rechts, S. 53). Dann sind weder prototypische Fälle identifizierbar noch wird etwas über die Prädisposition von Wörtern für das Auftreten in einer syntaktisch-funktionalen Konstruktion gesagt.

Croft (2001) schließlich ist undeutlich in der Unterscheidung von semantischem Wechsel und Polysemie (vgl. Abbildung 4 Punkt III, S. 53). Die Polysemie-Analyse richtet sich gegen die Vagheits-Analyse von Hengeveld (1992) und damit gegen eine Monosemie-Annahme. Unklar ist aber, ob Croft (2001) *ako* als „event-word“ oder „object-word“ klassifiziert und „semantic shift“ in nicht-prototypischer Auftretensweise analysiert, oder ob Croft (2001) *ako* als polysemes Lexem betrachtet, das zwei unterschiedlich klassifizierbare Bedeutungen trägt. Im ersten Fall sehe ich keinen Unterschied zu den Analysen von Völkel (2017) und Don und Van Lier (2013), im zweiten Fall kann Croft (2001) keine prototypische Abdeckung semantischer Klassen auf pragmatische Funktionen zeigen, da das von ihm vorgeschlagene Kriterium der *typologischen Markiertheit* nicht anwendbar ist.

Völkel (2017), Don und Van Lier (2013) und Croft (2001) gleichermaßen müssen sich in ihrer Analyse zuerst entscheiden, ob sich ein Wort in verschiedenen pragmatischen Funktionen oder syntaktischen Slots in seiner Bedeutungsklasse unterscheidet. Falls ja, analysieren alle drei den Bedeutungswechsel eines Wortes eben im Satz, wobei dieser bei Völkel (2017) nicht klar von Vagheit, bei Croft (2001) nicht klar von Polysemie und bei Don und Van Lier (2013) nicht klar von Null-Derivation getrennt ist. Als Maßstab der Analyse kann in allen

Fällen ohne formale Differenzierung nicht einfach die Übersetzung in eine andere Sprache gelten, und weder scheint in allen Fällen die Unterscheidung eindeutig, noch gibt es überhaupt einen zwingenden Grund, in den polynesischen Sprachen ausschließlich z.B. Polysemie oder Vagheit anzunehmen. Am Ende kann in keinem der Ansätze die Prädisposition – und sei diese nur als prototypisch verstanden – von Wörtern für den Gebrauch in syntaktisch-funktionalen Konstruktionen gezeigt werden. Es bleibt die Frage, ob die Differenzierung semantischer Klassen relevant für mehr als eine reine Klassifizierung der Bedeutungen von Wörtern im aktuellen Auftreten ist.

Wenn man letztlich akzeptiert, dass der Kern der Wortarttheorie in der Annahme liegt, dass Wörter prädisponiert sind für den Gebrauch im Satz, sollte zunächst eine Distributionsanalyse darüber Aufschluss geben. Die wesentliche Frage ist dann nicht, ob semantische Klassen in jeder Sprache auch als Wortbedeutungen lexikalisiert sein müssen, sondern ob je nach Analyse damit eben etwas über die Distribution von Wörtern angenommen werden kann. Zumindest Völkel (2017) hat eine Prädispositionshypothese aufgestellt in der Annahme, dass der Flexibilitätsgrad und die Frequenz von Wörtern in syntaktischen Slots Rückschlüsse auf ihre lexikalisierte Bedeutung zulässt.

2.3 Zusammenfassung und These

Betrachtet man zusammenfassend die Theorie der Wortarten, lässt sich ein Weg nachzeichnen, der seinen Ausgang in der antiken Wortartbeschreibung nimmt und nach deren Tradierung seinen Scheidepunkt an der Auseinandersetzung mit nicht-indogermanischen Sprachen findet, (vgl. Abbildung 5 S. 56). Am Scheidepunkt in der Auseinandersetzung mit nicht-indogermanischen Sprachen stellt sich die Frage nach der Generalisierbarkeit des tradierten Wortartmodells und der Universalität von Wortartkategorien. Unter Beibehaltung des tradierten Wortartmodells wird entweder an seiner prinzipiellen Übertragbarkeit festgehalten oder nicht. Je nach Analyse kann das Urteil über das Vorhandensein von Wortartkategorien in einzelnen Sprachen, etwa den polynesischen Sprachen, dabei allerdings sehr unterschiedlich ausfallen (vgl. etwa Buschmann 1839 oder Vonen 2000). In neueren, typologisch orientierten Ansätzen wird der traditionellen Wortarttheorie mit diskurspragmatischen Funktionen und konzeptuell-semantischen Klassen vermeintlich ein für den übereinzelsprachlichen Vergleich besser geeignetes *tertium comparationis* entgegengestellt. Dabei wird genau genommen nur ein einzelnes, aber bereits in

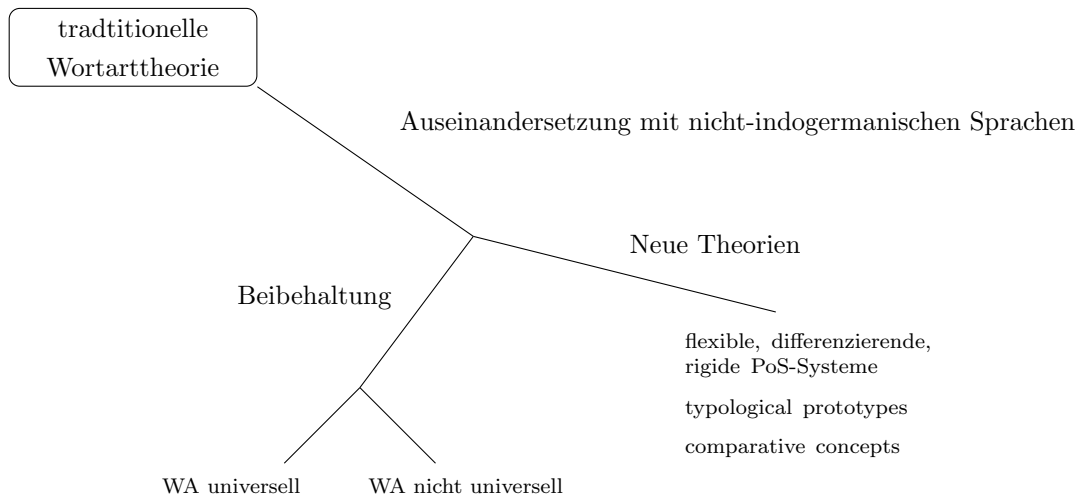


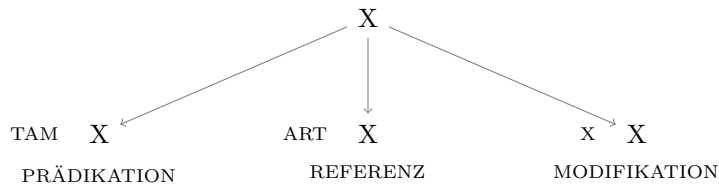
Abbildung 5: Entwicklung Wortarttheorien

der tradierten Wortarttheorie enthaltenes Kriterium hervorgehoben, ohne andere Kriterien wie die Unterscheidbarkeit formaler Klassen oder syntaktischer Slots und die gegenseitige Abdeckung dieser aufeinander tatsächlich zu überwinden. In der Folge müssen wie schon unter Beibehaltung des tradierten Wortartmodells in einzelnen Sprachen Wortarten abgelehnt werden und es ist unklar, worin genau der Unterschied zum tradierten Wortartmodell bestehen soll. Ob dann die Einführung neuer Begrifflichkeiten wie „flexible Parts-of-Speech“ (vgl. Hengeveld 1992) oder „flexible Wortarten“ (vgl. Van Lier und Rijkhoff 2013) einen wesentlichen Beitrag zum Verständnis von Wortartkategorien liefert, mag hier offen bleiben. Ansätze hingegen, die weiterhin an universellen Kategorien etwa „typological prototypes“ (vgl. Croft 2001) oder „comparative concepts“ (vgl. Haspelmath 2012) festhalten, verlassen damit aber entweder das Terrain der Wortarten und sprechen im Grunde über andere Kategorien oder versuchen sozusagen durch die Hintertür Wortarten universell vergleichbar zu machen, ohne so die Grenzen der tradierten Wortarttheorie tatsächlich hinter sich zu lassen.

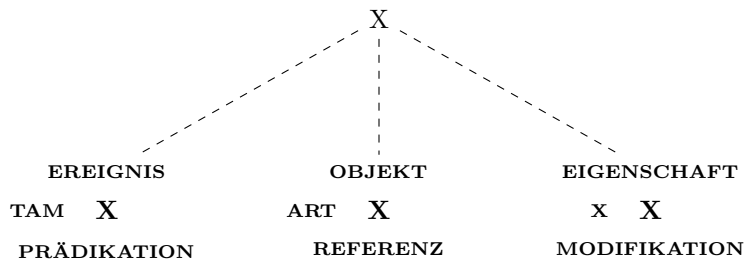
Die Beschreibungsversuche von Wortartkategorien in den polynesischen Sprachen stehen dem Problem gegenüber, dass anders als in flektierenden Sprachen Wörter nicht in Formklassen geschieden sind. Im Kern stellt sich hier die Frage, ob Wörter für den Gebrauch im Satz, sei es in unterscheidbaren Syntagmen oder in pragmatischen Funktionen, prädisponiert sind. Im Rahmen der

verschiedenen übergeordneten Theorien haben sich grundlegend drei Alternativen für den Umgang damit herausgebildet, vgl. Abbildung 6:

1. Keine Kategorisierung



2. Syntaktische Kategorien



3. Prädisposition semantisch-lexikalischer Klassen

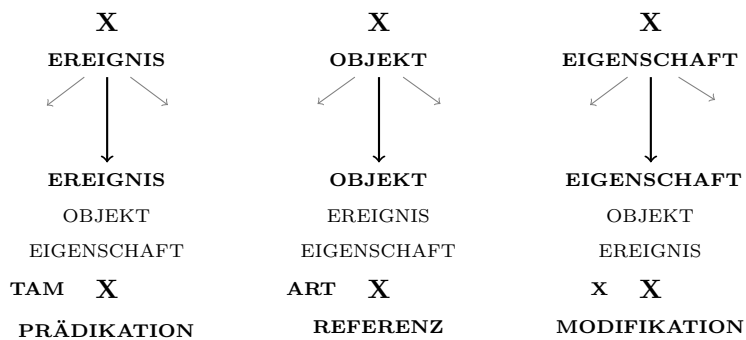


Abbildung 6: Alternative Ansätze zur Wortartkategorisierung in den polynesischen Sprachen

In der ersten Alternative (vgl. Abbildung 6 oben, S. 57) müssen Wortarten in den polynesischen Sprachen abgelehnt werden, da aufgrund der nicht möglichen Zuordnung von Wörtern zu morphologischen Paradigmen lexikalisch keine Formklassen differenziert werden können und Wörter damit auch nicht für den Gebrauch in Syntagmen oder diskurspragmatischen Funktionen prädisponiert sind (vgl. z.B. Schleicher 1865 oder Hengeveld 1992). Da die Semantik in dieser Analyse nichts über die Prädisposition von Wörtern aussagt, spielt sie nur eine untergeordnete Rolle.

In der zweiten Alternative (vgl. Abbildung 6 Mitte, S. 57) werden Wortarten als rein syntaktische Kategorien angesetzt: Trotz nicht vorhandener lexikalischer Klassen soll einem Wort im jeweiligen syntagmatischen Auftreten ein eigenständiger kategorialer Status zukommen (vgl. z.B. Mosel und Hovdhagen 1992). Dabei bleibt fraglich, was diese Kategorien genau sind und ob ohne klare Trennung von lexikalischer und syntaktischer Ebene solche Kategorien mit Wortarten gleichgesetzt werden können.

In der dritten Alternative (vgl. Abbildung 6 unten, S. 57) wird die Prädisposition von Wörtern für das Auftreten in Syntagmen oder pragmatischen Funktionen nicht durch Zuweisung zu Formklassen, sondern aufgrund der lexikalisierten Bedeutung, genauer der Einordnung dieser in ontologische Klassen, angenommen. Dabei ließe sich im einfachsten Fall die Bedeutung von Wörtern, die ausschließlich in einer syntaktisch-funktionalen Konstruktion auftreten, immer einer bestimmten semantischen Klasse zuordnen. So seien Wörter mit als EREIGNIS klassifizierter Bedeutung prädisponiert für das Auftreten in Tempus-Aspekt-Modus-Syntagma bzw. prädikativer Phrase, Wörter mit als OBJEKT klassifizierter Bedeutung für das Auftreten in Artikel-Syntagma bzw. referentieller Phrase und Wörter mit als EIGENSCHAFT klassifizierter Bedeutung für das Auftreten in modifizierender Phrase. Bei Wörtern, die in mehr als einer syntaktisch-funktionalen Konstruktion auftreten, stellt sich im Wesentlichen die Frage nach dem semantischen Zusammenhang zwischen den Auftretensweisen der identischen Wortformen. Dieser wird dann sehr unterschiedlich analysiert und die in der Folge angesetzten Kategorien werden mit verschiedenen Begriffen bezeichnet (vgl. Tabelle 7, S 59).

In einer Homonymie-Analyse (vgl. Dixon 2010) werden zwei identische Wortformen im Auftreten in unterschiedlichen syntaktisch-funktionalen Konstruktionen als zwei homonyme und damit kategorial verschiedene Wörter angenom-

Zusammenh.	Wortform	sem. Klasse	Syntagma	Funktion	Name
Homonymie	X	Ereignis	TAM	Prädikation	Verb
	X	Objekt	ART	Referenz	Nomen
Konversion	X ₁	Ereignis	TAM	Prädikation	Verb
	X ₂	Objekt	ART	Referenz	Nomen
	X ₃	Eigenschaft	x X	Modifikation	Adjektiv
syntakt. Abl. semantic shift	X	Ereignis	+TAM	+Prädikation	event-denoting term
			ART	Referenz	event-word event-root
			x X	Modifikation	(prototypisches) Verb
	X	Objekt	TAM	Prädikation	object-denoting term
			+ART	+Referenz	object-word object-root
			x X	Modifikation	(prototypisches) Nomen
	X	Eigenschaft	TAM	Prädikation	property-denoting term
			ART	Referenz	property-word property-root
			+ _x X	+Modifikation	(prototypisches) Adjektiv
Polysemie	X	Ereignis	+TAM	+Prädikation	prototypisches Verb
		Objekt	+ART	+Referenz	prototypisches Nomen
		Eigenschaft	+ _x X	+Modifikation	prototypisches Verb

Tabelle 7: Vergleich lexikalisch-semantischer Theorien

men und mit den Wortartkategorien Verb und Nomen¹⁶ gleichgesetzt. Da bei Wörtern wie *si'i* leicht ein semantischer Zusammenhang oder bei Wörtern wie *hana* je nach Sichtweise gar kein semantischer Unterschied analysierbar ist, überzeugt eine Homonymie-Analyse nicht.

In einer Konversions-Analyse wird der semantische Zusammenhang zwar berücksichtigt, aber kategorienverändernde Ableitung angesetzt (vgl. z.B. Vonen 2000, Don und Van Lier 2013), so dass ähnlich wie bei der Homonymie-Analyse

¹⁶Dixon (2010) betrachtet nur Verben und Nomen als universale Kategorien und äußert sich nicht zu Adjektiven in den polynesischen Sprachen.

Wortformen in (abgeleitete) kategorial verschiedene Lexeme oder Wurzeln differenziert werden. Problematisch ist hier insbesondere, die Richtung der Ableitung zu bestimmen. Dies gilt ebenso für die Analyse von syntaktischer Ableitung oder semantischem Wechsel (vgl. z.B. Croft 2001, Völkel 2017), bei der angenommen wird, dass die Klasse der lexikalisierten Bedeutung eines Wortes erst im Auftreten im Satz wechselt. In einer Polysemie-Analyse, das heißt der Annahme mehrerer (klassifizierter) Bedeutungen eines Wortes, würde die Bedeutung erst im Auftreten im Satz disambiguiert, wobei unklar ist, ob dann z.B. eine als EIGENSCHAFT-klassifizierte Bedeutung nur im Auftreten eines Wortes in modifizierender Phrase analysierbar sein sollte.

Bei der Bestimmung der semantischen Klasse der Bedeutung von Wörtern im Auftreten in syntaktisch-funktionalen Konstruktionen und der Frage des semantischen Zusammenhangs zwischen den Auftretensweisen geht es darum, trotz nicht unterscheidbarer Formklassen zumindest eine prototypische Abdeckung – und im Kern heißt das nichts anderes als die Prädisposition – von Wörtern für den Gebrauch in syntaktisch-funktionalen Konstruktionen nachzuweisen. Da das Kriterium der *prototypischen Markiertheit* von Croft (2001) für einen solchen Nachweis untauglich ist (vgl. Kapitel 2.2.4, S. 37), bleibt nur, mit Völkel (2017) zu überprüfen, ob Wörter unterschiedlich flexibel und unterschiedlich häufig in einzelnen Phrasen auftreten.

Damit lässt sich der zu Beginn in der Frage formulierte Forschungsgegenstand dieser Arbeit, ob Wortarten als Beschreibungsgrößen in flexionslosen Sprachen wie dem Hawaiischen abgelehnt werden müssen oder nicht, folgendermaßen präziser fassen: Sind Wortarten im Hawaiischen trotz fehlender Unterscheidbarkeit formaler Klassen definiert durch die Prädisposition von in ihren lexikalischen Bedeutungsklassen bestimmbar Wörtern für das Auftreten in syntaktisch-funktionalen Slots?

Meine These lautet: nein.

3 Methodik

Zur Untersuchung der Hypothese, die in ihrer allgemeinen Form besagt, dass Wörter im Hawaiischen prädisponiert sind für das Auftreten im Satz und in präzisierter Form, dass Wörter entsprechend ihrer lexikalisierten Bedeutungsklassen prädisponiert sind für bestimmte syntaktisch-funktionale Slots, sollen korpuslinguistische Verfahren einschließlich einer automatisierten Analyse eingesetzt werden. An dieser Stelle sei zunächst die damit vorausgesetzte Auffassung von Korpuslinguistik als Methode erwähnt. In der korpuslinguistischen Literatur findet sich eine Diskussion darum, ob der Korpuslinguistik der Status einer Theorie oder einer Methode zukommt, häufig verbunden mit den Etiketten *corpus-driven* versus *corpus-based* und einer idealisierten Gegenüberstellung von *bottom-up*- und *top-down*-Ansatz (vgl. Tognini-Bonelli 2001: 65–98, Gries 2010, McEnery und Hardie 2012: 147–162). Während bei ersterem versucht wird, Hypothesen aus einem Korpus zu generieren, sollen bei letzterem Hypothesen an Korpusdaten überprüft werden. Da die zu untersuchende Hypothese nicht aus Korpusdaten heraus entwickelt wurde, sondern aus der theoretischen Behandlung ausgewählter Einzelbeispiele (vgl. Kapitel 2, S. 4–60), betrachte ich eine korpuslinguistische Analyse als geeignetes Verfahren zur Überprüfung dieser Hypothese und verstehe Korpuslinguistik im Sinne von McEnery und Hardy (2012: 148–149) als Methode und nicht als Theorie, vgl.: „[...] corpus-as-method considers corpora and corpus techniques to be sources of empirical data that may be deployed in support or refutation of any explanatory theory about language – even a theory devised in whole or in part without reference to corpus data.“ Dabei definiere ich hier Stefanowitsch (2020: 1) folgend *Korpus* möglichst allgemein als „a large collection of authentic text (i.e., samples of language produced in genuine communicative situations).“

In den folgenden Abschnitten soll zunächst die Motivation für eine Korpusanalyse genauer dargestellt werden, um daraufhin einen Überblick über den Ablauf der Korpusanalyse zu geben und schließlich auf die einzelnen Schritte einzugehen.

Die Ausführungen zum Korpusdesign umfassen grundlegende Überlegungen bei der Erstellung eines Korpus, die Klärung der Datenlage im Hawaiischen und am Untersuchungsgegenstand orientierte Entscheidungen bei der Datenauswahl. Abschließend soll die technische Umsetzung des angewandten automatisierten Verfahrens erläutert werden.

3.1 Motivation für eine Korpusanalyse

Die methodische Entscheidung für eine Korpusanalyse liegt in der zu untersuchenden Hypothese und dem Vorgehen in bisherigen Arbeiten begründet, die Wahl eines automatisierten Verfahrens in der praktischen Umsetzung der Korpusanalyse.

Bei der zu untersuchenden Hypothese muss zunächst unterschieden werden zwischen der vorangestellten Distributionsannahme und der spezifischen Prädispositionsannahme (vgl. Kapitel 2.2.1, S. 12, Kapitel 2.2.4, S. 45–46). Die Distributionsannahme besagt erstens, dass nicht alle Wörter in allen syntaktisch-funktionalen Slots auftreten und zweitens, dass Wörter nicht in gleicher Häufigkeit in einem syntaktisch-funktionalen Slot auftreten. Die eigentliche Prädispositionsannahme besagt dann, dass die Distribution eines Wortes nicht als zufällig analysiert werden kann, sondern über die Zugehörigkeit zu einer semantischen Klasse bestimmt ist. Als ein weiterer Aspekt, der in den zuvor besprochenen theoretischen Ansätzen kaum Beachtung findet, kommt hier die Frage des Zusammenhangs der Wortdistribution zur Derivationsmorphologie hinzu. An gegebener Stelle sei ausführlich auf die Derivationsmorphologie im Hawaiischen eingegangen (siehe Kapitel 4.1.3, S. 131–150 und Kapitel 4.1.4, S. 167–175). An dieser Stelle soll ein kurzer Blick auf das Suffix *-na* genügen, um die Problemstellung beispielhaft herauszustellen. So bezeichnen Elbert und Pukui (1979: 80 und passim) *-na* als „nominalizing suffix“. Es ist aber unklar, ob mit *-na* suffigierte Derivate, z.B. *ha‘awi-na* „Lehrstunde“ mit Basis *hā‘awi* „geben“ (vgl. Elbert und Pukui 1979: 81), mehrheitlich oder gar ausschließlich in Artikel-Syntagma auftreten und inwieweit in solchen Fällen von *Nominalisierung* gesprochen werden kann.

Zusammenfassend gilt es also folgendes zu untersuchen:

1. Distribution
 - (a) Vorkommen von Wörtern in Syntagmen
 - (b) Vorkommenshäufigkeit von Wörtern in Syntagmen
2. Prädisposition
 - (a) Zusammenhang mit Derivationsmorphologie
 - (b) Zusammenhang mit (der Bestimmbarkeit) semantischer Klassen

In einer Korpusanalyse sehe ich nun insbesondere gegenüber den Arbeiten, die in Kapitel 2.2 (S. 8–60) besprochen wurden, den Vorteil, dass im Rahmen einer prämissenarmen Vorgehensweise eine größere Datenmenge systematisch untersucht werden kann. In diesem Zusammenhang bedeutet eine prämissenarme Vorgehensweise, dass die Bestimmbarkeit der semantischen Klassen der Wortbedeutungen nicht vorausgesetzt werden muss, wenn zuerst die Distribution von Wörtern untersucht wird. So bleibt dann erstens offen, ob eine semantische Klassifizierung von Wörtern überhaupt möglich ist und es entstehen anders als z.B. bei Völkel (2017) nicht schon vorab Probleme bei der Entscheidung darüber, welchen semantischen Klassen die Bedeutungen einzelner Wörter zugeordnet werden sollen (vgl. Kapitel 2.2.4, S. 47–48). Zweitens muss eine Distributionsanalyse einer semantischen Analyse vorausgehen, da bei fehlenden Distributionsklassen die Einordnung von Wortbedeutungen in semantische Klassen für die Wortartkategorisierung irrelevant wäre. Wenn Wörter im Hawaiischen nicht unterschiedlich (häufig) auf verschiedene syntaktisch-funktionale Slots verteilt sind, spielt auch die semantische Klassifizierung keine Rolle für die Prädisposition.

Zur Untersuchung der Wortdistribution bedarf es korpuslinguistischer Verfahren, dessen Ziel es im Allgemeinen ist „to describe and interpret generalizable patterns of language use“ (Biber und Jones 2009: 1287).

In Bezug auf die Vorkommenhäufigkeit von Wörtern in Syntagmen muss sich eine systematische Untersuchung der Wortdistribution dann an *corpus-based linguistics* im Gegensatz zu *corpus-illustrated linguistics* orientieren (vgl. Tummers, Heylen und Geeraerts 2005). Während in *corpus-illustrated linguistics* Korpora im Wesentlichen als empirische Basis für die „anecdotal selection of examples from a corpus“ (Tummers, Heylen und Geeraerts 2005: 238) mit dem Ziel verwendet werden, eine Hypothese zu stützen oder zu falsifizieren, gilt für *corpus-based linguistics* hingegen: „[...] to identify the relevant, i.e., significant, features of language use, descriptive and inferential statistical techniques have to be used.“ (Tummers, Heylen und Geeraerts 2005: 235 - 236). Dem Gegenstand der sich daraus im Einzelnen ergebenden, weiteren Untersuchungen entsprechend, folgt dann aber mehrheitlich eine Orientierung an *corpus-illustrated linguistics*.

Sowohl für das Hawaiische als auch für andere polynesischen Sprachen fehlen bisher systematische Untersuchungen zur Wortdistribution.

In typologisch ausgerichteten Ansätzen wie dem von Hengeveld (1992) oder Croft (2001) werden nur Einzelbeispiele aus den polynesischen Sprachen dis-

kutiert, um sie als Belege für die jeweils als universell gültig erachtete Theorie anzuführen (vgl. Kapitel 2.2., S. 16–40). Zwar können typologische Arbeiten selbst keine extensiven Beschreibungen einzelner Sprachen leisten. Wenn aber nur Einzelbeispiele aus Sprachen herangezogen werden, um diese in eine universelle Theorie von *Wortarten* oder *Parts-of-Speech* zu integrieren, besteht die Gefahr, dass relevante Beobachtungen zu einer Sprache nicht gemacht werden. Das hat sich beispielsweise in der Diskussion um Polysemie und Monosemie am Beispiel von *si'i* im Tongaischen gezeigt, in der nicht immer alle Bedeutungen und Auftretensweisen von *si'i* berücksichtigt wurden (vgl. Kapitel 2.2.4, S. 33–36). Zur Untersuchung der Distribution von Wörtern schließlich taugt die exemplarische Analyse einzelner Wörter gar nicht.

Aber auch in anderen Arbeiten wurde die Distribution von Wörtern nicht systematisch untersucht.

So bleibt in einzelsprachlichen Grammatiken wie der von Elbert und Pukui (1979) zum Hawaiischen unklar, wie viele Wörter sowohl in Artikel-Syntagma als auch Tempus-Aspekt-Modus-Syntagma auftreten. Ebenso finden sich bei Dixon (2010) keine empirischen Belege für die Vermutung, dass im Samoanischen die Mehrheit der Wörter allein in einem syntaktisch-funktionalen Slot auftritt (vgl. Kapitel 2.2.1, S. 13).

Völkel (2017) schließlich hat zwar die Verteilung von Wörtern auf syntaktisch-funktionale Slots im Tongaischen genauer zu analysieren versucht, allerdings ohne Distributions- und Prädispositionsannahme voneinander zu trennen. So hat Völkel (2017) allein die Distribution von zuvor semantisch klassifizierten Wörtern betrachtet und damit weder unabhängig die Distribution von Wörtern untersucht, noch die Prädisposition von Wörtern gezeigt.¹⁷ Da zudem nur wenige Daten aus Akzeptabilitätstests in der Feldforschung gewonnen wurden, gibt es begründete Zweifel an der Repräsentativität der Ergebnisse der Analyse und Völkel (2017) selbst verweist dann für zukünftige Untersuchungen auf die Notwendigkeit von Korpusstudien (vgl. Kapitel 2.2.4, S. 47–50). Zwar ist die Frage der Repräsentativität auch in Korpusuntersuchungen nicht ganz unproblematisch (siehe Kapitel 3.3.2, S. 68). Es ist aber zumindest möglich, mehr Daten systematisch zu untersuchen.

Bevor auf die Einzelheiten der hier durchgeführten Korpusanalyse eingegangen wird, lässt sich die Grundüberlegung, die die methodische Entscheidung dafür

¹⁷In ähnlicher Weise geht z.B. auch Van Lier (2016) vor, die in 36 Ozeanischen Sprachen erst nach der Einteilung von Wörtern in semantische Klassen die Distribution untersucht.

motiviert, folgendermaßen zusammenfassen: Viele der bisherigen Arbeiten zur Wortartthematik in den polynesischen Sprachen einschließlich des Hawaiischen – insbesondere solche, die von einer semantischen Kategorisierung ausgehen – stellen in dem Sinne *Top-Down*-Beschreibungen dar, dass sich darin auf wenige Daten bezogen wird, um diese entweder aus bestimmten theoretischen Prämissen heraus zu beurteilen oder exemplarisch zur Stützung vorangestellter Theorien heranzuziehen. Unklar bleibt dabei aber die Wortdistribution, die die eigentliche Voraussetzung für viele Theorien darstellt. Eine Korpusanalyse soll im Gegensatz zu den genannten Arbeiten und in der Orientierung an der formulierten Hypothese in einem ersten Schritt eine systematische Untersuchung der Wortdistribution im Hawaiischen ermöglichen, ohne dabei eine Kategorisierung vorauszusetzen. Um möglichst viele Daten untersuchen zu können, soll dafür auf ein automatisiertes Verfahren zurückgegriffen werden (siehe Kapitel 3.4, S. 78–104). Erst in einem zweiten Schritt ist es dann potentiell möglich, einen Zusammenhang mit der Klassifizierung von Wortbedeutungen herzustellen.

Im Folgenden soll zunächst ein Überblick über das Vorgehen in der Korpusanalyse gegeben werden, bevor auf die einzelnen Schritte genauer eingegangen wird.

3.2 Überblick Korpusanalyse

Der gesamte Ablauf der Korpusanalyse umfasst mehrere aufeinander aufbauende Schritte, die in Abbildung 7 kurz zusammengefasst sind, vgl.:

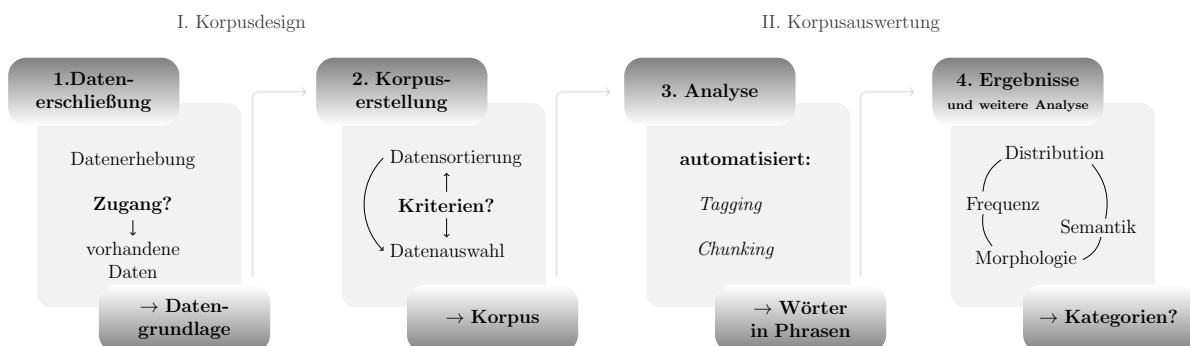


Abbildung 7: Überblick Schritte in der Korpusanalyse

Die Vorgehensweise in der Korpusuntersuchung lässt sich grob in zwei Teile gliedern: erstens das Korpusdesign und zweitens die Korpusauswertung. Im Korpusdesign geht es zunächst um die Datenerschließung, das heißt um die Frage, ob Daten selbst erhoben werden sollen oder vorhandene Daten zugänglich sind, wobei letzteres hier der Fall war. Die gesichteten Daten bilden dann die Grundlage für den nächsten Schritt: die Korpuserstellung. Für die Korpuserstellung müssen die Daten in der Grundlage nach an der Hypothese entwickelten Kriterien sortiert und für das Korpus ausgewählt werden. Die Auswertung des so erstellten Korpus beginnt mit einer automatisierten Analyse, die im Kern aus *Tagging* und *Chunking* besteht. Am Ende der automatisierten Analyse liegen erste Ergebnisse über die Distribution und Auftrenshäufigkeit von Wörtern in Syntagmen vor. Darauf müssen weitere Analysen folgen, um sich schließlich der Frage der Kategorisierung zu nähern und die zu untersuchende Hypothese zu prüfen.

3.3 Korpusdesign

Das Korpusdesign umfasst die Sichtung und Erschließung von Daten einschließlich der Entscheidung darüber, welche Daten als Grundlage zur Erstellung des Korpus dienen und welche Kriterien dabei ausschlaggebend sind. Bevor die Datenauswahl erläutert wird, soll im Folgenden die Datengrundlage geklärt werden.

3.3.1 Datenerschließung

Bei der Datenerschließung steht die Frage des Zugangs zu Daten im Hawaiischen im Vordergrund. Es geht genauer gesagt um die Frage, ob bereits Daten im Hawaiischen vorhanden sind, die für die Erstellung eines Korpus genutzt werden können oder Daten selbst erhoben werden müssen.

Was das Vorhandensein von Daten betrifft, nimmt das Hawaiische in gewisser Weise eine Sonderstellung unter den polynesischen Sprachen ein. Denn obwohl das Hawaiische weniger als 10.000 Sprecher hat (vgl. Eberhard, Gary und Fenning 2021) mit vermutlich deutlich weniger Muttersprachlern (vgl. den Exkurs zu Geschichte, Entwicklung und Dokumentation des Hawaiischen, S. 69–74), sind aufgrund der Geschichte der Verschriftlichung und insbesondere

der Bemühungen um eine Revitalisierung eine große Menge an schriftlichen Daten erhalten, von denen viele auch in digitaler Form vorliegen.

Die wohl umfangreichste, frei zugängliche, digitale Datensammlung des Hawaiischen stellt *Ulukau – the Hawaiian Electronic Library* (ulukau.org) dar, die vom *Hale Kuamo‘o*, dem *Hawaiian Language Center* innerhalb des *Ka Haka ‘Ula O Ke‘elikolani, College of Hawaiian Language* der *University of Hawai‘i at Hilo* gegründet wurde mit dem Ziel „to make these resources available for the use, teaching, and revitalization of the Hawaiian language and for a broader and deeper understanding of Hawai‘i.“ (ulukau.org). In *Ulukau* sind unzählige, diverse Daten in hawaiischer (und englischer) Sprache digital gespeichert. Dazu gehören neben Büchern, die im 19. Jahrhundert oder später im Rahmen der Revitalisierung des Hawaiischen seit den 1970er Jahren entstanden sind, verschiedene historische Sammlungen wie *Ho‘olaupa‘i – He ‘Ohina Nūpepa Ōlelo Hawai‘i*, *Hawaiian Nūpepa Collection* – eine Sammlung von Zeitungen, die zwischen 1834 und 1948 erschienen sind (nupepa.org), *Inoa ‘Āina Hawai‘i*, *Hawaiian Place Names* – eine Sammlung hawaiischer Ortsnamen oder *Papa Kuhikuhi Mo‘okū‘auhau Hawai‘i*, *Hawaiian Genealogy Indexes* – eine Sammlung, die unter anderem Heiratsurkunden und Einbürgerungsanträge aus dem 19. und frühen 20. Jahrhundert umfasst.

In Anbetracht der heutigen Situation des Hawaiischen als gesprochene und geschriebene Sprache (vgl. den Exkurs zu Geschichte, Entwicklung und Dokumentation des Hawaiischen, S. 69–74) und der Fülle an Daten, die durch *Ulukau* zur Verfügung stehen, habe ich darauf verzichtet, neue Daten zu erheben und stattdessen auf *Ulukau* als Datengrundlage aufgebaut. Zwar kann *Ulukau* selbst als digitales Korpus betrachtet werden. Verschiedene Überlegungen haben allerdings dazu geführt, *Ulukau* nicht im Ganzen als Korpus zu verwenden, sondern daraus ein neues für die Untersuchung adäquates Analysekorpus zu erstellen.

Diese Überlegungen und die genaue Erstellung des Analysekorpus sind Gegenstand des nächsten Kapitels.

3.3.2 Korpuserstellung

Bei der Korpuserstellung müssen verschiedene Überlegungen berücksichtigt werden. Diese betreffen einerseits grundsätzliche Fragen, die bei der Erstellung eines Korpus von Bedeutung sind, andererseits Eigenschaften der Datengrundlage, die mit der spezifischen Geschichte, Entwicklung und Dokumentation der hawaiischen Sprache zusammenhängen.

Das maßgebliche Ziel ist ein auf die Fragestellung ausgerichteter Analysekorpus. Dies allein schließt aus, bestimmte Daten aus *Ulukau* in das Korpus aufzunehmen. So werden z.B. Sammlungen von Ortsnamen oder Heiratsurkunden wenig Aufschluss über die Verteilung von Wörtern auf Artikel- oder Tempus-Aspekt-Modus-Syntagma geben können.

Der vielleicht wichtigste Aspekt bei der Erstellung eines Korpus ist die in der korpuslinguistischen Literatur viel diskutierte Repräsentativität (vgl. Biber 1993, Sinclair 2005, Evert 2006, Hunston 2008, Reppen 2010, Nelson 2010, Stefanowitsch 2020: 28–36), bei der es um die Frage geht, ob von einer Stichprobe, i.e. des Korpus, Rückschlüsse auf eine in Bezug auf die Forschungsfrage definierte Grundgesamtheit gezogen werden können. Da die Wortartkategorisierung im Hawaiischen zur Untersuchung steht, ist auch das Hawaiische als Grundgesamtheit definiert und das Korpus sollte repräsentativ für die hawaiische Sprache als Ganzes sein. Damit müsste das Korpus als repräsentative Stichprobe theoretisch die gesamte Architektur der Sprache einschließlich der diatopischen, diastratischen und diaphasischen Dimension abbilden. Dazu sei erstens angemerkt, dass Repräsentativität in der korpuslinguistischen Praxis vor allem als Ideal betrachtet wird, dem es sich durch Erstellung eines möglichst ausgewogenen Korpus, das heißt eines Korpus, das z.B. viele Texte, verschiedene Textsorten und Texte mehrerer Verfasser enthält, anzunähern gilt. Zweitens sei folgende Annahme betrachtet: Wenn die Hypothese zutrifft und Wörter über die Zugehörigkeit zu semantischen Klassen prädisponiert sind für den Gebrauch in bestimmten syntaktisch-funktionalen Slots, kann theoretisch jede Stichprobe, die groß genug ist, als repräsentativ betrachtet werden, weil die Prädisposition von Wörtern nicht von Faktoren wie Textsorte, Autor oder Varietät beeinflusst sein sollte und die (Häufigkeits-)Verteilung von Wörtern relativ konstant bleiben müsste. Bezogen auf die Größe des Korpus spielen genau genommen eher praktische Überlegungen eine Rolle und weniger die Frage der Repräsentativität.¹⁸ So sollte einerseits die Distribution möglichst vieler Wörter untersucht werden und andererseits müssen Wörter zumindest häufig genug vorkommen, um überhaupt ihre (Häufigkeits-)Verteilung auf verschiedene syntaktische-funktionale Slots untersuchen zu können (vgl. dazu allerdings die Ausführungen zum Zipfschen Gesetz in Kapitel 4.2.2, S. 186–189).

¹⁸ Auch wenn in der korpuslinguistischen Literatur Einstimmigkeit darin herrscht, dass die Größe eines Korpus mit seiner Repräsentativität korreliert, existiert kein eindeutiges Maß dafür, wann ein Korpus groß genug ist, um als repräsentativ zu gelten (vgl. Stefanowitsch 2020: 37–38).

Die grundlegenden Kriterien der Orientierung am Untersuchungsgegenstand, der Repräsentativität und der Korpusgröße unterliegen in der tatsächlichen Umsetzung der Korpuserstellung Einschränkungen, die durch spezifische Gegebenheiten der in der Datengrundlage enthaltenen Daten vorgegeben sind. Um diese besser zu verstehen und weitere Kriterien für die Sortierung der Daten in *Ulukau* und der Auswahl für das Analysekorpus zu entwickeln, ist es notwendig, einen Blick auf die historische Entwicklung Hawai‘is und den damit verbundenen Entstehungshintergrund der Daten in *Ulukau* zu werfen. Dazu folgt an dieser Stelle ein Exkurs zur Geschichte, Entwicklung und Dokumentation des Hawaiischen, bevor die Datenauswahl präzisiert werden kann.

Exkurs zu Geschichte, Entwicklung und Dokumentation des Hawaiischen

Eine ausführliche Abhandlung über die Geschichte Hawai‘is kann nicht Gegenstand dieses Exkurses sein. Für eine differenzierte Darstellung der tiefgreifenden Veränderungen in der hawaiischen Gesellschaft seit dem frühen 19. Jahrhundert, geprägt durch demographisch-ethnische Entwicklungen sowie eine immer größer werdende Einflussnahme von außen und damit verbunden des Zusammenbruchs der traditionellen sozialen Ordnung des *Tabu*-Systems, des Verlusts traditioneller Landrechte und der Durchsetzung eines politischen und wirtschaftlichen Systems nach westlichem Vorbild, sei auf die entsprechende Literatur verwiesen (vgl. Malo 1903, Kuykendall 1938, 1958, 1967, Kamakau 1992 [1866–1871], Trask 1993, Fornander 1996 [1880], Dibble 2005 [1838], Osorio 2002, Silva 2004b, Nogelmeier 2010).

Das Ziel ist es hingegen, anhand historischer Orientierungspunkte die Entwicklung und (schriftliche) Dokumentation des Hawaiischen nachzuzeichnen, um daraus Kriterien für die Sortierung der Daten in *Ulukau* und die Aufnahme in das Analysekorpus zu gewinnen. Tabelle 8 (S. 70) bietet einen Überblick über einschneidende Ereignisse in der Geschichte Hawai‘is, an denen beispielhaft das komplizierte Zusammenspiel zwischen politisch-gesellschaftlichen Veränderungen und der Entwicklung des Hawaiischen seit dem ersten Aufeinandertreffen von Hawaiianern mit Fremden¹⁹ veranschaulicht werden soll.

¹⁹Ich orientiere mich hier mit dem Begriffspaar „Hawaiier“ versus „Fremde“ an der Selbstbezeichnung *Kanaka Maoli* oder *Kanaka ‘Ōiwi*, von *kanaka* „Mensch, Person“ und *maoli* „wahr“ bzw. *‘ōiwi* „einheimisch“ (vgl. Blaisdell 1989, Silva 2004b: 12–13, Goodyear-Ka‘ōpua 2014: 2). *Kanaka Maoli* bezeichnet dann „the autochthonous people of the Hawaiian archipelago – the original people who emerged from this place“ (Goodyear-Ka‘ōpua 2014: 2) und „‘Kanaka’ by itself also means ‘Hawaiian’, especially when used in contrast with ‘haole’ when

James Cook erreicht Hawai‘i	1777	Oralität	
Eroberung des hawaiischen Archipels und Gründung des Königreich Hawai‘i durch Kamehameha I	1791-1810		
Ankunft der ersten protestantischen Missionare	1820	Literarität	
Erstes Schriftdokument in hawaiischer Sprache (<i>Spelling Book</i>)	1822		
Erste Zeitung in hawaiischer Sprache (<i>Ka Lama Hawaii</i>) unter Anleitung von Reverend Andrew Lorrins	1834		
Hawaiische Übersetzung der Bibel	1837-1839		
Erste von einem Hawaiianer (J. K. Kaunamano) gegründete Zeitung (<i>Ka Hoku o ka Pakipika</i>)	1861		
Umsturz der hawaiischen Monarchie	1893		
Höhepunkt erreicht mit 14 verschiedenen hawaiisch-sprachigen Zeitungen	1896		
Offizielle Annexion Hawai‘is durch die USA	1898		
Schließung der letzten hawaiisch-sprachigen Zeitung (<i>Hoku o Hawaii</i>)	1948	Niedergang des Hawaiischen	
Eingliederung Hawaiis in die USA als 50. Bundesstaat	1959		
Beginn der <i>Hawaiian Renaissance</i>	1970er	Revitalisierung	Entstehung des <i>Neo-Hawaiian</i>
hawaiisch-sprachige Radiosendung (<i>Ka Leo Hawai‘i</i>)	1972-1988		
hawaiisches Wörterbuch moderner Begriffe (<i>Māmaka Kaiiao</i>)	1996		
Zeitungsartikel in hawaiischer Sprache (<i>Kauakūkalahale</i>)	seit 2002		

Tabelle 8: Historische Entwicklung des Hawaiischen

Den sich zunächst im 19. Jahrhundert vollzogenen Umbruch fasst Chapin (1996: 15) folgendermaßen zusammen: „In Hawai‘i, an oral, memory-based

meant as ‘foreigner’.“ (Silva 2004b: 12).

traditional culture, a culture that had thrived in isolation [...], was rapidly displaced by codified laws, constitutions, and newspapers.“

Mit Ankunft der ersten protestantischen Missionare im Jahre 1820 wurde die Verschriftlichung des Hawaiischen angestoßen (vgl. Kuykendall 1938: 100–116, Mookini 1974: iv, Judd, Bell und Murdoch 1978: xiii–xvi). 1822 entstand mit *Piapa* (für „Alphabet“) – einer Art Buchstabiertafel – das erste Schriftdokument in hawaiischer Sprache (vgl. Kuykendall 1938: 105, Judd, Bell und Murdoch 1978: xiv). Darauf folgten verschiedene Lehrwerke, insbesondere zur religiösen Erziehung, aber auch zur Arithmetik oder zur menschlichen Anatomie, die 1839 abgeschlossene Übersetzung der Bibel und 1834 mit *Ka Lama Hawaii* unter Leitung von Reverend Lorrin Andrews die erste hawaiisch-sprachige Zeitung (Kuykendall 1938: 105 ff., Mookini 1974: iv ff.). Waren bis in die Mitte des 19. Jahrhunderts die Druckerzeugnisse noch wesentlich bestimmt durch die Erziehungsansprüche der Missionare (vgl. Chapin 1984: 52–53, 1996: 15–18, Schütz 1994: 339 ff.), hatte sich mit der Zeit das, was eine eigenständige hawaiische intellektuelle Tradition genannt werden kann (vgl. Chapin 1996: 59–62, Silva 2004b, 2008), emanzipiert, die spätestens 1861 in der ersten von einem Hawaier, namentlich J. K. Kaunamano, gegründeten Zeitung *Ka Hoku o Pakipika*, sichtbar wurde.

Insgesamt kann von einer Blütezeit der hawaiischen Presse gesprochen werden, die 1896 mit 14 verschiedenen hawaiisch-sprachigen Zeitungen ihren Höhepunkt erreichte (vgl. Mookini 1974: x) und spätestens 1948 mit der Schließung von *Hoku o Hawaii*, der letzten hawaiisch-sprachigen Zeitung ein Ende gefunden hat (vgl. Mookini 1974: xiv). In einem Klima der politisch-gesellschaftlichen und demographischen Veränderungen entstand eine heterogene Presselandschaft²⁰ mit ebenso religiös motivierten wie säkularen Zeitungen unterschiedlicher Interessensverfolgung und politischer Couleur, die sowohl von Muttersprachlern als auch Nicht-Muttersprachlern herausgegebenen worden sind (vgl. Mookini 1974, Chapin 1984, 1996, 2000, Silva 2004b: 54–86). Neben Zeitungen entstanden auch andere Veröffentlichungsformen wie Regierungsdokumente, Bücher und Lehrwerke, die jedoch häufig abschnittsweise in den Zeitungen abgedruckt wurden.

Insgesamt kann im 19. Jahrhundert von einer hohen Literarität in der hawaiischen Bevölkerung ausgegangen werden (vgl. Chapin 1984: 67, Silva 2004b:

²⁰Es wurden auch viele Zeitungen in anderen Sprachen herausgegeben (vgl. Chapin 2000). Insgesamt macht Chapin (2000: 1) 11 verschiedene Sprachen aus. Darunter neben Englisch und Hawaiisch z.B. Chinesisch, Japanisch oder Portugiesisch.

55) – Schmitt (1968: 12) geht beispielsweise für das Jahr 1878 von einer Literaritätsrate von fast 80 % aus – und die Vermutung liegt nahe, dass der Niedergang als Zeitungssprache auch den schleichenden Verlust des Hawaiischen als gesprochene Sprache gegenüber dem Englischen spiegelt (vgl. Mookini 1974: xiv, Chapin 1984, NeSmith 2002: 13, Silva 2004b: 54–86), oder wie es Schütz (1994: 341) ausdrückt: „In short, the spoken language was sustained and bolstered by the written language [...]“.

So sehr sich schon zum ausgehenden 19. Jahrhundert die Verschiebung vom Hawaiischen zum Englischen abzeichnen mag (vgl. Schütz 1994: 341 ff.), scheint es bis heute schwierig, verlässliche Angaben zur tatsächlichen Sprecherzahl des Hawaiischen zu machen. Im Zensus 2016 (vgl. DBEDT 2016: 9) sind dazu folgende Angaben zu finden: 18.610 Personen (5,7 % der Gesamtbevölkerung von Hawai‘i) sprechen zuhause neben Englisch Hawaiisch. Solche Angaben sind allerdings mit Vorsicht zu betrachten, weil sie erstens auf freiwilligen Angaben von Befragten in Umfragen beruhen (vgl. DBEDT 2016: 2) und zweitens wenig Einblick über den genauen Gebrauch oder gar die Muttersprachlichkeit geben. Ähnliche Probleme gibt es auch bei vorherigen Zensusdaten, wie Schütz (1994: 380, Anmerkung 41) in seiner Besprechung des Zensus von 1990 bemerkt, in der Schütz (1994: 365) eher von 2000 Sprechern ausgeht. Schütz (1994: 376) muss schließlich resümierend feststellen, „no one seems to know just how many native speakers there are, or how the language is being used.“ NeSmith (2002: 13) schätzt aufbauend auf Interviews mit Mitgliedern der Ni‘ihau-Gemeinde und Sprachaktivisten der *University of Hawai‘i at Mānoa* 800 – 1000 mehrheitlich ältere Muttersprachler, davon 500 – 600 Sprecher des Ni‘ihau-Dialekts. Für den Rückgang des Hawaiischen können diverse ineinandergreifende Faktoren vermutet werden (vgl. Schütz 1993: 340 ff.). Eine nicht zu unterschätzende Rolle dürften dabei die demographischen Veränderungen gespielt haben: Zum einen dezimiert sich – anfänglich ausgelöst durch eingeschleppte Krankheiten von außen – die hawaiische Bevölkerung drastisch (vgl. Silva 2004b: 24–27) und zum anderen steigt der Zuwachs anderer ethnischer Gruppen enorm an (vgl. Schütz 1994: 341–342).

Ähnlich wie bei den Sprecherzahlen ist es auch hier schwierig, genaue Zahlen zu nennen. Schätzungen zur Bevölkerungszahl in der Zeit vor dem ersten Kontakt mit Fremden schwanken zwischen 250.000 (Schmitt 1971: 240) und 800.000 (Stannard 1989). Beispielhaft für den Rückgang sei Lind (1967: 28) genannt, der für die Wende des 19. auf das 20. Jahrhundert 40.000 Hawaier mit einem Anteil von 24,4% der Gesamtbevölkerung schätzt. Dem Zensus 2018

(vgl. DBEDT 2018: 2) zufolge leben in Hawai‘i 299.451 *Native Hawaiians* (21,3 % der Gesamtbevölkerung), wobei auch diese Angaben auf der Selbstausskunft von Befragten basieren mit der Möglichkeit, mehreren ethnischen Gruppen zugehörig zu sein.

Die allmähliche Durchsetzung des Englischen kann ebenso auf machtpolitische Maßnahmen im Schulwesen zurückgeführt werden (vgl. Schütz 1994: 351-354, Benham und Heck 1994): Seit 1887 wurde in den von Missionaren kontrollierten Privatschulen nur noch in englischer Sprache unterrichtet, vermutlich aus Unzufriedenheit über den Missionierungserfolg und der dafür verantwortlich gemachten hawaiischen Sprache (vgl. Schütz 1994: 351). 1896, drei Jahre nach dem Umsturz der hawaiischen Monarchie, kam es dann zum offiziellen Verbot des Hawaiischen als Schulsprache, vermutlich in Verbindung mit der Unterdrückung von hawaiischen Souveränitätsansprüchen (vgl. Schütz 1994: 353, Chapin 1996: 29–31, Silva 2004b: 144).

Bis nach dem 2. Weltkrieg war die hawaiianische Sprache schließlich aus dem öffentlichen Leben Hawai‘is verschwunden (vgl. NeSmith 2002: 13).

In den 70er Jahre des 20. Jahrhunderts haben mit der sogenannten *Hawaiian Renaissance* Bemühungen allgemein um eine Revitalisierung der hawaiischen Kultur und spezifisch der hawaiischen Sprache begonnen (vgl. Kanahale 1979, 1982, 1986, Schütz 1994: 361 – 375). Diese zunächst auf private Initiativen zur Förderung von Immersionsprogrammen (*Pūnana Leo*) beschränkten Revitalisierungsversuche haben später auch ihren Niederschlag in der offiziellen Politik gefunden, so dass das Hawaiische seit 1986 als Unterrichtssprache an staatlichen Schulen zugelassen ist und mittlerweile von der Grundschule bis zur Universität unterrichtet wird (vgl. Schütz 1994: 369).

Zumindest vereinzelt hat das Hawaiische auch wieder mehr Raum in den Medien gefunden, z.B. in der von 1972 bis 1988 laufenden, mehrheitlich hawaiischsprachigen Radiosendung *Ka Leo Hawai‘i* (vgl. Kimura 1972–1988) oder der seit 2002 erscheinenden hawaiisch-sprachigen Zeitungskolumne *Kauakūkalahale* (vgl. Wong 2002).

Seit den 70er Jahren sind daneben zahlreiche Veröffentlichungen entstanden, wie das *Māmaka Kaiāo*, ein hawaiisches Wörterbuch moderner Begriffe (vgl. Kōmike Hua‘ōlelo 1996), Bücher in englisch- und hawaiisch-sprachiger Ausgabe (z.B. Stewart Williams 1993, Wong 1996), kommentierte Nachdrucke von Büchern aus dem 19. Jahrhundert (z.B. Hale‘ole 2001 [1863]) oder Lernergrammatiken (z.B. Kamanā und Wilson 2012).

Eine wichtige Veränderung zu früheren Schriftdokumenten haben all diese

Veröffentlichungen gemeinsam: eine neue Orthographie, in der im Gegensatz zur alten, in frühen Texten üblichen Orthographie der Glottal (haw. *‘okina* „Pause“) einheitlich graphemisch realisiert wird mit dem umgekehrten Apostrophen und lange von kurzen Vokalen mithilfe des Macron (haw. *kahakō*) unterschieden werden.

So wenig wie in der Frühphase der Verschriftlichung des Hawaiischen, kann aber seit den 1970er Jahren von einer Standardisierung des Hawaiischen ausgegangen werden. NeSmith (2003) spricht in Zusammenhang mit der Revitalisierung des Hawaiischen von der Entstehung des *Neo-Hawaiian*, einer neuen unter L2-Lernern verbreiteten Varietät, die sich weniger durch interne Homogenität auszeichnet, als durch die erheblichen Unterschiede zu dem von Muttersprachlern gesprochenen Hawaiischen. Zu den insgesamt 15 von Ne-Smith (2003: 70–72) als zentral herausgearbeiteten Unterschieden gehören neben Unterschieden in der Aussprache, der Gebrauch von aus dem Englischen entlehnten Wörtern und die eingeschränkte Verwendung von Tempus-Aspekt-Konstruktionen.

Betrachtet man zusammenfassend die Entwicklungen des Hawaiischen seit dem frühen 19. Jahrhundert, reichen diese von der anfänglich von Missionaren angestoßenen Verschriftlichung über eine sich daraus entwickelnde selbstständige hawaiische intellektuelle Tradition, die in einer Blütezeit der hawaiischsprachigen Presse gemündet ist, bis hin zum schleichenden Niedergang sowohl des geschriebenen als auch gesprochenen Hawaiischen und dem Beginn der Revitalisierungsbewegung, durch dessen Bemühungen eine neue Varietät entstanden ist.

Das bis hierin erarbeitete Hintergrundwissen zeigt, dass die in *Ulukau* überlieferten Daten das Resultat einer mehrfach gebrochenen Sprachgeschichte sind. Dies muss neben den grundsätzlichen Überlegungen zur Korpuserstellung bei der Auswahl von Daten für die Aufnahme in das Analysekorpus berücksichtigt werden.

Im Wesentlichen lassen sich zwei Kriterien festmachen, die die Qualität der Daten betreffen und für die Aufnahme in das Korpus relevant sind. Dies sind die Muttersprachlichkeit der Verfasser und die Orthographie der Texte.

Als erstes sollten dann soweit möglich nur von Muttersprachlern verfasste Texte in das Analysekorpus aufgenommen werden. Das hatte eine Verengung der Daten auf den Zeitraum zwischen 1834 und 1948 zur Folge, da somit das in der Revitalisierung entstandene *Neo-Hawaiian* ausgeschlossen werden konnte. Bei den Daten aus diesem Zeitraum wurde, soweit anhand verfügbarer Metadaten möglich (vgl. Siddall 1921, Mookini 1974, Chapin 1984, 2000, For-

bes 2000, Silva 2002a, 2002b, 2003, 2004a), zwischen muttersprachlichen und nicht-muttersprachlichen Autoren unterschieden. Insbesondere bei Zeitungstexten aus der Frühphase zwischen 1834 bis Mitte des 19. Jahrhunderts ist dies allerdings nicht ohne Weiteres möglich. Zu den Autoren der 1834 erschienenen Zeitung *Ka Lama Hawai'i* gehören beispielsweise sowohl missionarische Lehrer als auch hawaiische Schüler der von Reverend Lorrin Andrews geleiteten *Laihanaluna School* in Maui (vgl. Mookini 1974: iv, Chapin 1984: 51, Silva 2002a: 60–61). Es finden sich aber nicht in jedem Text Angaben zum Verfasser und es ist unklar, ob tatsächlich jedem Text nur ein alleiniger Verfasser zugewiesen werden kann. Ähnliches gilt auch für Lehrwerke, die innerhalb der *Laihanaluna School* entstanden sind, z.B. *Anatomia*, vermutlich eine partielle Übersetzung eines 1834 in Boston in englischer Sprache erschienenen Lehrbuchs zur menschlichen Anatomie von Jerome Smith, die dem amerikanischen Arzt und Missionar Gerrit Judd zugeschrieben wird (vgl. Judd 1838, Judd, Bell und Murdoch 1978: 59, Silva 2002a: 122–123).

Während dann Texte, dessen Verfasser eindeutig als muttersprachlich identifiziert werden konnten, in das Analysekorpus aufgenommen wurden, sind Texte, dessen Verfasser eindeutig als nicht-muttersprachlich identifiziert werden konnten, von einer Aufnahme in das Analysekorpus ausgeschlossen worden. Die Mehrheit der Texte ohne eindeutig identifizierbaren Verfasser ist ebenfalls in das Korpus aufgenommen worden. Dieser Entscheidung liegt erstens die Annahme zugrunde, dass zumindest in der *Laihanaluna School* kaum Texte ohne muttersprachliche Expertise entstanden sein dürften und zweitens, dass sich vermutlich in dem neu entstandenen, geschriebenen Hawaiischen genuin hawaiische Diskursstrategien mit nicht-hawaiischen vermischt haben (vgl. Chapin 1983: 49, Silva 2004a: 178–179, Nogelmeier 2005: xvii–xxi).

Neben der Muttersprachlichkeit der Autoren ist als zweites Kriterium für die Aufnahme von Texten in das Korpus die Orthographie herausgearbeitet worden. Die Orthographie spielt für die Disambiguierung von Wörtern, insbesondere in der späteren automatisierten Analyse (vgl. Kapitel 3.4, S. 78–104), eine wichtige Rolle. Als Beispiel sei hier <nana> in alter Orthographie genannt, das in neuer Orthographie mit <nana> als Bezeichnung für eine Taro-Varietät, <nāna> als Possessivpronomen, <nanā> mit der Bedeutung „provizieren, aggressiv“ und <nānā> mit der Bedeutung „schauen, beobachten“ realisiert wird.

Da Texte aus dem Zeitraum zwischen 1834 und 1948 in das Korpus aufgenommen werden sollten und diese in den Originalausgaben nur in alter Orthogra-

phie vorliegen, musste auf Nachdrucke in neuer Orthographie zurückgegriffen werden. Gerade für in diesem Zeitraum erschienene Texte in Buchform, sind Nachdrucke in neuer Orthographie herausgegeben worden (vgl. z.B. Hale‘ole 2001 [¹1863]). Für andere Texte waren vor allem die im Rahmen von *Papahana Ho‘oilina ‘Ōlelo Hawai‘i, The Hawaiian Language Legacy Program* entstandenen drei Ausgaben des *Legacy Journal* hilfreich (vgl. Silva 2002a, 2002b, 2003, 2004a), in dem Zeitungsartikel aus ausgewählten Zeiträumen, Regierungsdokumente und weitere „Legacy Materials“ in alter und neuer Orthographie, mit englischer Übersetzung und angereichert mit Metadaten herausgegeben wurden.

Zusammenfassend ist die Erstellung des Analysekorpus in mehreren Schritten erfolgt, in denen verschiedene Kriterien für die Aufnahme von Daten entwickelt und angewendet wurden. Diese Kriterien sind:

1. Grundlegend:
 - (a) Orientierung am Untersuchungsgegenstand
 - (b) Repräsentativität und Größe
2. Spezifisch:
 - (a) Muttersprachlichkeit
 - (b) Orthographie

Zunächst sollten keine Daten selbst erhoben werden, sondern *Ulu kau - the Hawaiian Electronic Library* (vgl. ukukau.org) als Datengrundlage dienen. Unter Berücksichtigung des spezifischen Entstehungshintergrundes, der durch die historische Entwicklung Hawai‘is und der hawaiischen Sprache bestimmt ist, wurden die Daten in *Uluka* sortiert und für die Aufnahme in das Analysekorpus weitere, neben den bei der Erstellung eines Korpus grundlegend zu beachtenden Kriterien, entwickelt.

Die Orientierung am Untersuchungsgegenstand hat letztendlich dazu geführt, Texte auszuschließen, in denen nur wenige Tempus-Aspekt-Modus-Syntagmen und Artikel-Syntagmen zu erwarten waren. Nach Überlegungen zur Repräsentativität des Korpus für das Hawaiische sollten zumindest möglichst viele Texte, Textsorten und Texte verschiedener Verfasser in das Korpus aufgenommen werden. Die Konzentration auf muttersprachliche Texte hat die Auswahl auf

Daten aus dem Zeitraum zwischen 1834 und 1948 verengt und Texte von eindeutig nicht-muttersprachlichen Verfassern ausgeschlossen.

Die Notwendigkeit von Texten in neuer Orthographie hat die Datenauswahl endgültig eingegrenzt.

Insgesamt konnte so ein 1.121.198 Tokens umfassendes Korpus erstellt werden, dass aus 101 Texten, die im Zeitraum von 1834 bis 1924 entstanden sind und von 49 identifizierbaren Autoren verfasst wurden, besteht.

Die Textsorten umfassen *Mo‘olelo* (Geschichten), *Mele* (Lyrik), Briefe, Werbeanzeigen, Traueranzeigen, Verfassungs- und Gesetzestexte, Regierungsmitteilungen, Kurznachrichten und nicht weiter klassifizierte Zeitungsartikel.²¹

Zur Repräsentativität sei noch folgendes angemerkt: Theoretisch wäre es denkbar, das Korpus als repräsentative Stichprobe für ein als finite Grundgesamtheit definiertes „Zeitungshawaiisch“ zu betrachten, das alle geschriebenen und mehrheitlich in Zeitungen veröffentlichten Texte in hawaiischer Sprache aus den Jahren 1834 bis 1948 einschließt. Dass die Texte allerdings nicht per Zufall in das Korpus aufgenommen worden sind, sondern anhand bestimmter Kriterien, wird in der Struktur deutlich. So überwiegen z.B. Texte mit vergleichsweise niedriger Tokenfrequenz – von den 101 Texten bestehen 80 Texte aus weniger als 1000 Tokens. Demgegenüber fallen fast 80% der Daten (888.399 von 1.121.198 Tokens) auf vier Texte von vier Autoren und einer Textsorte (*Mo‘olelo*). Ob das Korpus dann als repräsentativ für ein „Zeitungshawaiisch“ oder das Hawaiische als Ganzes gelten kann, ist ebenso wie die Frage nach der Relevanz der Zusammenstellung für die Distribution von Wörtern (vgl. S. 68) nicht endgültig geklärt.

Die Art des erstellten Analysekorpus kann schließlich als geschriebenes, synchrones, monolinguales, geschlossenes Referenzkorpus beschrieben werden (vgl. Bowker und Pearson 2002: 11–13), das die Grundlage für die Untersuchung der Distribution und Häufigkeitsverteilung von Wörtern auf syntaktisch-funktionale Slots bilden soll.

Um das über eine Millionen Tokens umfassende Korpus analysieren zu können, wurde ein automatisiertes Verfahren eingesetzt, dessen Realisierung Gegenstand des nächsten Kapitels ist.

²¹Die Einteilung in Textsorten basiert hier nicht auf einer differenzierten textlinguistischen Analyse, die besonders für die Textsortenspezifika des Hawaiischen gar nicht möglich war, sondern orientiert sich an den Angaben in den Metadaten (vgl. Silva 2002a, 2002b, 2003, 2004a).

3.4 Automatisiertes Verfahren

Das Ziel der Korpusanalyse ist es, die (Häufigkeits-)Verteilung von Wörtern auf bestimmte syntaktisch-funktionale Slots zu untersuchen. Um die Größe des Korpus bewältigen zu können, ist von einer manuellen Analyse abgesehen und stattdessen auf ein automatisiertes Verfahren zurückgegriffen worden, mithilfe dessen eine Liste erstellt werden soll, die alle im Korpus auftretenden Inhaltswörter und ihre Auftretenshäufigkeit in noch zu definierenden syntaktisch-funktionalen Slots enthält.

Im Folgenden sollen zunächst die Grundlagen geklärt und ein Überblick über die Verfahrensweise der Automatisierung gegeben werden, bevor genauere Ausführungen zu den einzelnen Schritten folgen.

3.4.1 Grundlagen und Überblick

Bei der technischen Umsetzung des automatisierten Verfahrens wurde auf die freie Programmiersprache *Python* (vgl. python.org) und das ebenfalls freie, quelloffene *Natural Language Toolkit (NLTK)*, eine Zusammenstellung von Bibliotheken und Programmen für computerlinguistische Anwendungen mit *Python* (vgl. nltk.org, Bird, Klein und Loper 2009, Perkins 2014), zurückgegriffen.

Um den Ablauf des automatisierten Verfahrens besser verständlich zu machen, soll folgendes abstraktes Beispiel betrachtet werden: Gegeben sei eine Zeichenkette *abcabc*. Vorab wird eine Handlungsanweisung formuliert, die es umzusetzen gilt:²² Zähle jedes Zeichen in der Zeichenkette *abcabc*, das nach dem Zeichen *a* und vor dem Zeichen *c* steht. In der Zeichenkette *abcabc* lässt sich dieser Handlungsanweisung folgend das Zeichen *b* zweimal zählen.

In einer vereinfachten Übertragung dieses Beispiels sei zunächst davon ausgegangen, das Korpus bestünde nur aus dem Satz *Ua hana ana 'o Herman me ka hale pauahi in Hilo*. Sei zudem vorausgesetzt, dass ein Tempus-Aspekt-Modus-Syntagma definiert ist als die Abfolge von *Ua* + Inhaltswort + *ana*. Dann würde die Handlungsanweisung zur Bestimmung der Auftretenshäufigkeit eines Wortes in Tempus-Aspekt-Modus-Syntagma lauten: Zähle jedes Wort im Korpus, das nach *Ua* und vor *ana* auftritt, mit dem Ergebnis, *hana* als einmal

²²vgl. hierzu auch den vor allem in der Informatik und Mathematik verwendeten Begriff des *Algorithmus* (vgl. Ziegenbalg 1996: 20–25, Cormen, Leiserson und Rivest 2000: 1–5, Klabunde 2010: 67).

in Tempus-Aspekt-Modus-Syntagma im Korpus auftretend zu zählen. Zwar gilt beim tatsächlich zu untersuchenden Korpus zu beachten, dass es aus mehr als einem Satz besteht, die Auftretenshäufigkeit von Wörtern nicht nur in Tempus-Aspekt-Modus-Syntagma gezählt werden soll und ein Syntagma im Hawaiischen aus mehreren, unterschiedlichen Elementen bestehen kann. Die grundlegende Verfahrenslogik bleibt aber die gleiche wie in den vorherigen Beispielen.

Die konkrete Umsetzung kann nun mit *Python* und *NLTK* über verschiedene Einzelschritte realisiert werden, vgl. Abbildung 8:

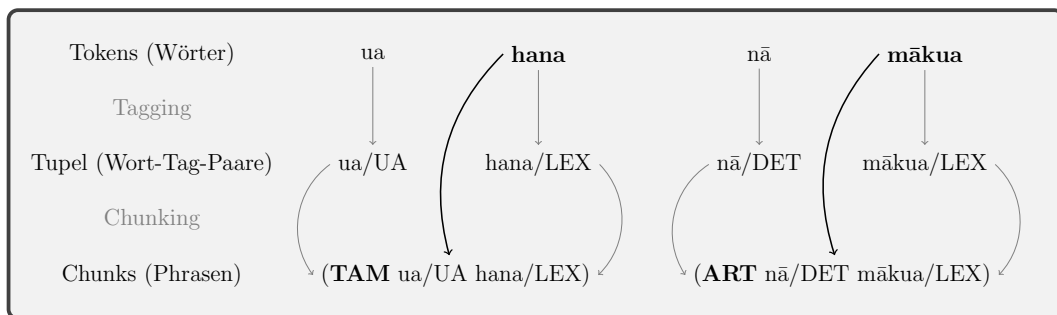


Abbildung 8: Einzelschritte im automatisierten Verfahren

Die beiden wesentlichen Schritte in der automatisierten Analyse sind das *Tagging*, bei dem jedem Wort im Korpus ein Tag zugewiesen wird und das *Chunking*, bei dem die aus dem *Tagging* entstandenen Tupel, i.e. geordneten Wort-Tag-Paare, sequentiell zu *Chunks*, i.e. größeren syntagmatischen Einheiten zusammengefasst werden. So lässt sich jedes Wortvorkommnis (Token) im Korpus eindeutig dem Auftreten in einem definierten Syntagma (*Chunk*) zuordnen. Eine Grundvoraussetzung für das *Tagging* und *Chunking* ist die *Tokenisierung*, das heißt die Segmentierung der Korpus-texte in Satz- und Worteinheiten. Im Beispiel aus Abbildung 8 wurde also nach der Segmentierung der Korpus-texte in Sätze und der Segmentierung des Satzes *ua hana nā mākua* in die Wörter *ua*, *hana*, *nā* und *mākua* jedem Wort ein Tag zugewiesen und schließlich die Abfolge der Wort-Tag-Paare *ua/UA hana/LEX* zu einem als Tempus-Aspekt-Modus-Syntagma definierten *Chunk* und die Abfolge der Wort-Tag-Paare *nā/DET mākua/LEX* zu einem als Artikel-Syntagma definierten *Chunk* zusammengefasst, so dass eine Instanz (ein Token) des Types *hana* in Tempus-Aspekt-Modus-Syntagma und eine Instanz (ein Token) des Types *mākua* in Artikel-Syntagma im Korpus erfasst bzw. gezählt werden konnte.

Zu den hier nur überblicksartig dargestellten Einzelschritten sei folgendes anzumerken, bevor in den nächsten Kapiteln im Detail darauf eingegangen wird: Die Umsetzung eines jeden Einzelschrittes umfasst Entscheidungen, die nicht-triviale theoretische Überlegungen erfordern. Diese betreffen unter anderem die Frage, was eine Worteinheit (im Hawaiischen) konstituiert, welche *Tags* verwendet werden können und wie Syntagmen (im Hawaiischen) strukturell definiert sind.

3.4.2 Normalisierung

Noch vor der *Tokenisierung*, dem *Tagging* und dem *Chunking* ist es notwendig, die in das Korpus aufgenommenen Texte zu normalisieren.

Unter *Normalisierung* lässt sich allgemein der Vorgang verstehen, alle in das Korpus aufgenommenen Texte für die nachfolgende Bearbeitung zu vereinheitlichen.

In einem ersten Schritt besteht die Normalisierung aus der Bereinigung der Korpustexte von systematischen Fehlern und der Entfernung von für die weitere Analyse nicht verwertbaren Material.

Die Bereinigung der Korpustexte hängt unter anderem mit der Entnahme der Buchtexte aus *Ulukau* und dem für die weitere Verarbeitung mit *Python* erforderlichen Format zusammen. In *Ulukau* liegen die digitalisierten Bücher als PDF-Download und als HTML-formatierter Webtext vor. Für die Aufnahme in das Korpus wurde jeweils der Webtext eines Buches als Textdatei gespeichert. Zu systematischen Fehlern, die bereinigt werden mussten, gehören dann solche, die typischerweise bei der Digitalisierung von Texten entstehen, wie die fehlerhafte Übertragung typographischer Eigenschaften, insbesondere der Schriftauszeichnung, der gedruckten Vorlage in der Umwandlung von PDF-Format auf HTML-formatierten Webtext, die sich auch in den gespeicherten Textdateien wiederfindet.

Beispielsweise ist in der gedruckten Version bzw. in der PDF-Vorlage des Buches *He Mo'olelo Ka'ao No Kekūhaupio* (Desha 1996 [1920,1924]) jeder Buchstabe zu Beginn eines neuen Kapitels als Versal, das heißt als vertikal über zwei Zeilen laufender Großbuchstabe, hervorgehoben. Falls der initiale Buchstabe auch ein Wort ist, kann es im Webtext bzw. der gespeicherten Textdatei dann vorkommen, dass sich dieses mit dem darauffolgenden Wort verbindet. So wird z.B. aus initial zwei Wörtern *I ka* in der Textdatei fälschlicherweise ein Wort *Ika*.

Derartige systematische Fehler wurden Text für Text in *Python* mit *Regulären*

*Ausdrücken*²³ korrigiert.

Auf die Bereinigung der Korpustexte folgte die Entfernung allen für die spätere Analyse nicht verwertbaren Materials. Dazu gehören insbesondere diejenigen Teile eines Textes, die nicht zum Textkörper zählen wie das Inhaltsverzeichnis, Tabellen, Fußnoten, Überschriften, Bildunterschriften und Seitenzahlen. Ebenso entfernt wurden Vorwort oder Danksagung eines Textes, wenn nicht vom Autor selbst, sondern vom Herausgeber verfasst, wie üblicherweise in Büchern zu finden, in denen Texte aus dem 19. Jahrhundert neu aufgelegt wurden (z.B. (Desha 1996 [¹1920,1924])).

Weitere notwendige Schritte der Normalisierung betreffen vor allem die Groß- und Kleinschreibung.

Dabei geht es darum, dass in einem Text z.B. *ua* am Satzanfang großgeschrieben ist wie in *Ua hana ana ‘o Herman me ka hale pauahi in Hilo*, nach einem Komma oder in anderen Positionen aber kleingeschrieben. Nun sollen im Hinblick auf das spätere *Tagging* (vgl. Kapitel 3.4.4, S. 83–90) und die Zählbarkeit jedoch *Ua* und *ua* nicht als Instanzen zweier Types, sondern Instanzen desselben Types *ua* erfasst werden. Dafür wurden mit *Regulären Ausdrücken* in *Python* in den Korpustexten alle Wörter, die auf ein Satzzeichen, genauer auf einen Punkt, einen Doppelpunkt, ein Ausrufezeichen oder ein Fragezeichen folgen von Groß- auf Kleinschreibung geändert, mit der Ausnahme von Abkürzungen nach Punkten, z.B. *S. D. Dole*, das zu *S D Dole* geändert wurde. Alle innerhalb eines Satzes großgeschriebenen Wörter wurden hingegen unverändert gelassen. Damit konnte der Umstand, dass im Hawaiischen nur Eigennamen und Wörter am Satzanfang großgeschrieben werden, beim späteren *Tagging* für die Erkennung von Eigennamen genutzt werden (vgl. Kapitel 3.4.4, S. 89).

Erst nachdem die Korpustexte wie beschrieben normalisiert wurden, konnten weitere Schritte folgen. Als erstes war die *Tokenisierung* der Texte notwendig, die im nächsten Abschnitt besprochen wird.

²³Etwas vereinfacht gesagt sind *Reguläre Ausdrücke* Zeichenketten, mit denen es möglich ist, Text zu filtern und zu manipulieren (vgl. Karttunen 2003, Stubblebein 2003, Jurafsky und Martin 2009, Klabunde 2010, Wintner 2010).

3.4.3 Tokenisierung

Mit *Tokenisierung* wird die Segmentierung einer Zeichenkette z.B. eines Textes in kleinere (sprachliche) Einheiten bezeichnet (vgl. Manning und Schütze 1999: 124-131, Mikheev 2003, Hagenbruch 2010: 264–271). Die *Tokenisierung* der Texte im Korpus in Sätze und Wörter gehört zu einem wesentlichen Schritt im Arbeitsablauf des *Tagging* und *Chunking* und soll hier nicht im Detail erläutert werden (vgl. Bird, Klein und Loper 2009 und Perkins 2014 zum genaueren Ablauf in *Python* mit *NLTK*).

Es folgen stattdessen ein paar Anmerkungen zu Aspekten der *Tokenisierung*, die mit Spezifika der Verschriftlichung des Hawaiischen und der vorangegangenen Normalisierung zusammenhängen.

Im Prinzip liegen die Korpustexte schon tokenisiert vor: Sätze sind durch Satzzeichen voneinander getrennt und Wörter durch Leerzeichen.

In der *Tokenisierung* mit *NLTK* wurden zusätzlich die Satzzeichen von Wörtern gelöst, so dass z.B. ein Satz wie *Ua hana ana ‘o Herman me ka hale pauahi in Hilo.* in die Tokens *ua, hana, ana, ‘o, Herman, me, ka, hale, pauahi, in, Hilo* und *.* segmentiert ist. Dadurch wird nicht fälschlicherweise *Hilo.* als ein Token erfasst und ist es später möglich, jedem Wort und Satzzeichen ein Tag zuzuweisen (vgl. Kapitel 3.3.4, S. 83–90).

Die Segmentierung in von Leerzeichen abgegrenzte Tokens hat nun aber Auswirkungen auf die Erfassung von Komposita, da diese in den Korpustexten nicht einheitlich geschrieben werden. So finden sich beispielsweise die Schreibweisen *meakanu* und *mea kanu* „Kulturpflanze“ mit *mea* „Sache, Ding, Person“ und *kanu* „pflanzen, Kulturpflanze“ oder *halepule* und *hale pule* „Gebetshaus, Kirche“ mit *hale* „Haus“ und *pule* „beten, Gebet“. In der Folge der hier gewählten *Tokenisierung* wird dann z.B. *meakanu* als ein Token eines Types, *mea kanu* hingegen als zwei Tokens zweier Types erfasst, auch wenn *meakanu* und *mea kanu* theoretisch zu einem Type gehören.

Dieser Umstand wurde hier bewusst in Kauf genommen. Denn erstens ist die Differenzierung von Komposition und Attribution im Hawaiischen nicht immer problemlos möglich (vgl. dazu die Vorschläge von Elbert und Pukui 1979: 123–130) und zweitens sollte das in den Texten vorliegende Material möglichst prämissenarm, so wie es gegeben ist, behandelt werden.

Ein ähnlicher Aspekt wie die *Tokenisierung* von Komposita betrifft den Umgang mit Konstruktionen, die nicht sinnvoll in ihre Bestandteile segmentiert werden können. Als Beispiel seien hier *a pau* mit der Bedeutung „ganz, völlig, vollständig“ oder *no laila* mit der Bedeutung „weil, folglich“ genannt. Der-

artige Konstruktionen sollten dann erst beim *Tagging* in besonderer Weise behandelt werden (vgl. Kapitel 3.4.4, S. 85–86).

Weniger problematisch ist die Frage, ob Teile von Namen und Abkürzungen zu einem Namen zusammengefasst werden sollen, das heißt z.B. S/ABR D/ABR Dole/NP als ein Name erfasst, da die Auftretensweise von Eigennamen nicht im primären Interesse der Korpusanalyse steht.

Die *Tokenisierung* des Korpus hat im Wesentlichen der Vorbereitung auf das *Tagging* gedient, das im Folgenden dargestellt wird.

3.4.4 Tagging

Allgemein ausgedrückt ist mit *Tagging* nichts anderes als die Annotation von Wörtern gemeint. Im genaueren lässt sich darunter die automatisierte Zuweisung einer Etikette (*Tag*) zu einem Token über die Bildung von Tupeln, i.e. geordneten Token-Tag-Paaren verstehen (vgl. Hagenbruch 2010: 271, Zinsmeister 2010: 483–484, Perkins 2014: 85).

Dazu sei anzumerken, dass ich bislang bewusst den in der Korpus- und Computerlinguistik üblichen Begriff des *Part-of-Speech-Tagging*, kurz *PoS-Tagging*, vermieden habe (vgl. Manning und Schütze 1999: 341–381, Bird, Klein und Loper 2009: 179–221, Perkins 2014: 85–123, Stefanowitsch 2020: 77–89), da ja gerade Wortartkategorien zur Untersuchung stehen. Im Kern dient hier das *Tagging* in seiner Voraussetzung für das *Chunking* nur als Mittel zum Zweck. Der Grundgedanke bei der Umsetzung ist, dass sich für das *Tagging* zumindest die Unterscheidung von Funktionswörtern und Inhaltswörtern im Hawaiischen nutzbar machen lässt – auch wenn diese nicht frei von Problemen ist (vgl. S. 85) – um schließlich über das *Chunking* Wörter in ihrem Auftreten in syntaktisch-funktionalen Slots erfassen und zählen zu können. Funktionswörter sollen dann mit der jeweiligen Wortart-Etikette, Inhaltswörter dagegen nur als solche ausgezeichnet werden. Damit handelt sich – dem Untersuchungsgegenstand geschuldet – sozusagen um ein „unsauberes“ *Part-of-Speech-Tagging*, weil genau genommen zwei unterschiedliche Ebenen ausgezeichnet werden.

Im Folgenden werden die einzelnen Schritte des *Tagging* erläutert.

Das Ziel ist es, einen *Tagger*, das heißt ein System zur Generierung von *Tags*, zu trainieren und schrittweise ein *Tagging*-Modell zu entwickeln, das zufriedenstellende Ergebnisse in der automatisierten Zuweisung von *Tags* an Wörter liefert.

In Abbildung 9 (S. 84) ist der Ablauf der einzelnen Schritte beim Trainieren eines *Taggers* zusammengefasst.

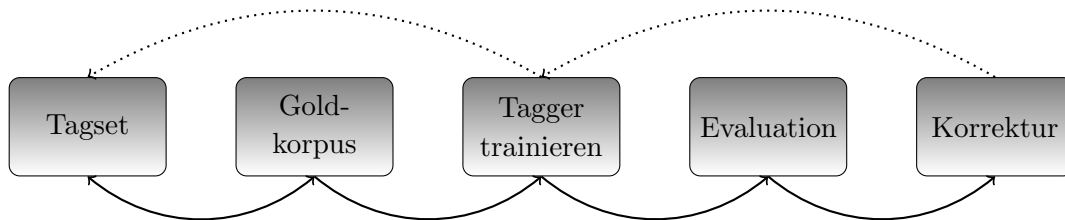


Abbildung 9: Schritte beim Trainieren eines Taggers

Der erste Schritt besteht in der Wahl eines *Tagset*, das heißt eines Inventars der verwendeten *Tags*. Darauf folgt die Erstellung eines *Goldkorpus*, das heißt eines manuell annotierten Korpus, das den Standard zur korrekten Zuweisung von *Tags* vorgibt und mit dem anfänglich ein *Tagger* trainiert und evaluiert werden kann. Durch kumulatives *Tagging* der einzelnen Korpus-texte, der Korrektur systematischer Fehler und des wiederholten Durchlaufs der vorangegangenen Schritte soll schließlich ein passendes *Tagging*-Modell entwickelt und das gesamte Korpus *getaggt* werden.

Die einzelnen Schritte seien im Folgenden genauer erläutert.

Eine wichtige Grundlage stellt zunächst das *Tagset* dar, in dem die beim *Tagging* verwendeten *Tags* und die Richtlinien der Verwendung dokumentiert sind. Bei der Wahl eines geeigneten *Tagset* kann nicht auf die Konventionen bekannter, für andere Sprachen entwickelter *Tagsets* wie des *Penn Treebank Tagset* für das Englische (vgl. Santorini 1991) oder des *Stuttgart-Tübingen-Tagset (STTS)* für das Deutsche (vgl. Schiller et al. 1999) zurückgegriffen werden, da die Spezifika des Hawaiischen und der zu annotierenden Ebenen berücksichtigt werden müssen. Für die Erstellung eines provisorischen *Tagset* war es zwar nützlich, sich an den Wortartdefinitionen in der Grammatik von Elbert und Pukui (1979) zu orientieren. Das endgültige *Tagset* stand jedoch erst nach dem gesamten *Tagging*- und *Chunking*-Prozesses fest, da es im Verlauf der einzelnen Schritte immer wieder modifiziert werden musste. In den Entscheidungen zu einzelnen *Tags* muss immer zwischen Einheitlichkeit und Granularität der Beschreibung und Erfolg des *Taggings* und *Chunkings* abgewägt werden, wobei letzteres als Priorität erachtet wurde.

An dieser Stelle sollen exemplarisch *Tags* vorgestellt werden, die im Besonderen von den Wortartdefinitionen in Elbert und Pukui (1979) abweichen.

Elbert und Pukui (1979: 43) teilen die Inhaltswörter des Hawaiischen in die Klassen „noun“ , „verb“ und „verb-noun“ (vgl. auch Kapitel 2.2.1, S. 12).

Dementgegen wurden hier Inhaltswörter nicht mit einer Wortartkategorie, sondern nur als solche annotiert, indem sie mit einem Stellvertreterlabel *LEX*, kurz für „Lexem“ oder „lexikalische Einheit“, versehen wurden.

Um beim *Chunking* den syntaktischen Slot eines Inhaltswortes differenziert zu bestimmen (vgl. Kapitel 3.4.5, S. 98), wurde bei einem attributiv auftretenden Inhaltswort eine Zahl hinzugefügt, die die Position des Inhaltswortes repräsentiert, also z.B. *mea kanu* als *mea/LEX kanu/LEX2* oder *ali'i wahine aloha* als *ali'i/LEX wahine/LEX2 aloha/LEX3* annotiert.

Bei Funktionswörtern unterscheiden Elbert und Pukui (1979: 43) die unter „Substitutes“ zusammengefassten Pronomina, Demonstrativa, Possessiva und Interrogativa von Präpositionen, Konjunktionen, Numeralia, Interjektionen und „Idioms“, das heißt „words not otherwise classifiable“ (Elbert und Pukui 1979: 43). Zusätzlich definieren Elbert und Pukui (1979: 44) Partikeln im Wesentlichen als in Nominal- und Verbalphrase auftretende Marker grammatischer Kategorien.²⁴

Da die Disambiguierung von formal identischen Funktionswörtern beim *Tagging* zu Problemen geführt hat, konnten diese mehrheitlich nicht mit einem spezifischen, sondern nur einem der Form entsprechenden *Tag* versehen werden. Das betrifft z.B. *ua* als Perfektivmarker (vgl. Elbert und Pukui 1979: 57–59) versus Demonstrativum (vgl. Elbert und Pukui 1979: 113), *i* als Perfektivmarker (vgl. Elbert und Pukui 1979: 59) versus Präposition (vgl. Elbert und Pukui 1979: 133–135) oder *ke* als Invitivmarker (vgl. Elbert und Pukui 1979: 62) versus Allomorph des definites Artikels (vgl. Elbert und Pukui 1979: 154–155). Entsprechend wurden *ua*, *i* und *ke* einfach als *ua/UA*, *i/I* und *ke/KE* annotiert, unabhängig von kategorialen Unterscheidungsmöglichkeiten. Es sei betont, dass die Entscheidung für diese Annotation keine theoretische Auseinandersetzung etwa mit der Frage der Multifunktionalität von Funktionswörtern im Hawaiischen spiegelt, sondern ausschließlich einen pragmatisch gewählten Ansatz bei der Umsetzung des *Tagging*.

Neben der formalen Identität ist ein weiterer problematischer Aspekt bei der Disambiguierung von Wörtern allein durch die übergeordnete Unterscheidung von Inhaltswörtern und Funktionswörtern gegeben. Denn diese Unterscheidung

²⁴Voraussetzung ist dafür freilich die von Elbert und Pukui (1979: 39–44 und passim) gemachte Trennung zwischen *verbalen* und *nominalen* Kategorien im Hawaiischen (vgl. auch Kapitel 2.2.1, S. 12).

ist nicht immer ohne weiteres möglich. Das betrifft beispielsweise Fälle von Grammatikalisierung wie *po‘e*, mit der lexikalischen Bedeutung „Menschen, Gruppe“ gegenüber dem Gebrauch als Pluralmarker (vgl. Elbert und Pukui 1979: 162) – *po‘e* und ähnliche Fälle mussten dann später aus dem Analysekorpus entfernt und gesondert behandelt werden (vgl. Kapitel 3.4.6, S. 104). Die Unterscheidung zwischen Funktions- und Inhaltswörtern betrifft ebenfalls Wörter, die Elbert und Pukui (1979: 120–123) über den Bezug auf die Raum- und Zeitdimension als „locative nouns“ klassifizieren. Viele dieser Wörter treten entweder in einem Artikel-Syntagma oder in bestimmten Konstruktionen auf, die nicht eindeutig einem syntaktischen Slot zugewiesen werden können. Dazu gehören z.B. *luna* in *ka luna* „Oberseite, Vorarbeiter“ und *i luna* „oben“, *hope* in *ka hope* „Letztgeborener“ und *ma hope* „nachdem, weil“ oder *kai* in *ke kai* „Meer“ und *i kai* „seewärts“. Um Schwierigkeiten beim *Chunking* zu umgehen (vgl. Kapitel 3.4.5, S. 90–98), mussten solche Wörter in Konstruktionen mit *LOC*, in definierten Syntagmen als *LEX* annotiert werden, also z.B. *i/I luna/LOC*, aber *ka/DEF luna/LEX*. In gleicher Weise wurde mit Konstruktionen wie *a laila* oder *a pau* verfahren (vgl. Kapitel 3.4.3, S. 82), das heißt *a laila* als *a/A laila/LOC* und *a pau* als *a/A pau/PM* annotiert.²⁵

Zusammenfassend ist im Laufe der Entwicklung des *Tagsets* keine Einheitlichkeit erreicht worden, da Funktionswörter und Inhaltswörter unterschieden wurden und aufgrund von pragmatischen Entscheidungen beim *Tagging* und *Chunking* verschiedene Ebenen ausgezeichnet wurden.

Eine weitere Grundlage für das Trainieren eines *Taggers* stellt neben dem *Tagset* das *Goldkorpus* dar, ein manuell annotiertes Korpus, mit dem der Standard für das korrekte *Tagging* gesetzt ist. Da im Hawaiischen anders als in anderen Sprachen wie dem Englischen oder Deutschen keine bereits annotierten Daten vorliegen, musste das *Goldkorpus* selbst erstellt werden. Das hat dazu geführt, dass zu Beginn des Trainings vergleichsweise wenig Daten zur Verfügung standen und manuelle Korrekturschritte in den Gesamtprozess eingebaut werden mussten.²⁶ Für die Erstellung des Goldkorpus wurden schrittweise Texte aus

²⁵Genau genommen ist es nicht korrekt, *pau* in der Konstruktion *a pau* als Modalpartikel (*PM*) zu taggen, da *pau* z.B. auch in Tempus-Aspekt-Modus-Syntagma mit der Bedeutung „beenden“ auftritt. Es ist aber eine praktikable Möglichkeit, über das *Tagging* Konstruktion und Syntagma auseinanderzuhalten.

²⁶Ebenso sollte das Goldkorpus, um eine möglichst hohe Qualität zu gewährleisten, idealerweise von mehreren Annotatoren blind bearbeitet und das *Inter-Annotator-Agreement* überprüft werden (vgl. Pustejovsky und Stubbs 2012: 28), was hier allerdings nicht möglich war.

dem Korpus per Zufall ausgewählt und manuell und – soweit möglich – mit *Regulären Ausdrücken* (vgl. S. 81, Anmerkung 23) annotiert. Das Goldkorpus enthält 11 Texte und umfasst 32.344 Tokens. Aufbauend auf dem Goldkorpus kann schließlich das eigentliche *Tagging* beginnen.

Die grundlegende Prozedur bei Trainieren eines *Taggers* besteht darin, die manuell annotierten Daten per Zufallsprinzip in ein Entwicklungskorpus und ein Testkorpus aufzuteilen, wobei das Entwicklungskorpus wiederum in ein Trainings-Set und ein Entwicklungs-Test-Set aufgespalten wird (vgl. Jurafsky und Martin 2009: 9–10, Pustejovsky und Stubbs 2012: 29). Das Trainings-Set enthält 80%, Entwicklungs-Test-Set und Testkorpus jeweils 10% der Daten des manuell annotierten Korpus (vgl. Jurafsky und Martin 2009: 10). Das Trainings-Set dient als Input für das Training des *Taggers*, das Entwicklungs-Test-Set zur Entwicklung und Verbesserung des *Tagging*-Modells, z.B. über das Variieren des *Tagset* oder veränderter Parameter des *Taggers*, und das Testset zur endgültigen Überprüfung des trainierten *Taggers*. So wird beispielsweise mit den Trainingsdaten ein *N-Gram-Tagger* trainiert, der die *Tags* der $n - 1$ vorangegangenen Tokens nutzt, um ein Token zu *taggen* (vgl. Manning und Schütze 1999: 192–196, Bird, Klein und Loper 2009: 203–205, Jurafsky und Martin 2009: 29–52, Perkins 2014: 94–96). In der Entwicklung eines *Tagging*-Modells kann z.B. der eingebundene Kontext in der Anzahl der n variiert und die Leistung eines *Unigram*-, *Bigram*- und *Trigram-Taggers* verglichen werden. Anhand des Entwicklungs-Test-Set kann dann über verschiedene Maße wie *Accuracy* oder *Precision and Recall* evaluiert werden, wie gut der jeweilige *Tagger* funktioniert (zu den einzelnen Maßen vgl. Bird, Klein und Loper 2009: 239–241, Jurafsky und Martin 2009: 65–67, Pustejovsky und Stubbs 2012: 170–175). Eine Möglichkeit, mit der sehr geringen Datenmenge im manuell annotierten Korpus umzugehen, liegt in der Anwendung der *zehnfachen Kreuzvalidierung*. Dabei wird die zufällige Teilung des Entwicklungskorpus in ein Trainings-Set und Entwicklungs-Test-Set 10-fach wiederholt, sodass möglichst viele Daten in das Training und Testen fließen und die Einzelevaluation der Wiederholungsdurchgänge, z.B. mit dem *Accuracy-Score*, zu einem Durchschnittswert zusammengefasst werden. Das entwickelte *Tagging*-Modell kann abschließend am zurückgehaltenen Testkorpus überprüft werden.

Nach der bis hierhin beschriebenen Prozedur wurden im *Trial-and-Error*-Verfahren mehrere *Tagger*, darunter *N-Gram-Tagger* (vgl. Bird, Klein und Loper 2009: 203–205, Perkins 2014: 94–96), *TnT-Tagger* (vgl. Brants 2002, Perkins 2014: 105–107), *RegExp-Tagger* (vgl. Bird, Klein und Loper 2009: 199, Per-

kins 2014: 99–100), *Brill-Tagger* (vgl. Brill 1995, Bird, Klein und Loper 2009: 208–210, Perkins 2014: 102–105), und – soweit möglich – Kombinationen aus diesen, ausprobiert. Das im Resultat entwickelte *Tagging*-Modell weist einen *Accuracy-Score* von 0,93 auf, das heißt, das 93% der mit dem Modell generierten *Tags* im Testkorpus korrekt den *Tags* des manuell annotierten Goldkorpus entsprechen. Aufgrund der geringen Menge an ursprünglichen Trainingsdaten musste allerdings im weiteren Verlauf, in dem schrittweise die restlichen Texte des gesamten Analysekorpus *getaggt* und manuell überprüft wurden, das *Tagging*-Modell immer wieder revidiert und systematische Fehler mit *Regulären Ausdrücken* korrigiert werden.

Das endgültige *Tagging*-Modell besteht aus einer Kombination aus drei verschiedenen *Taggern*, vgl. Abbildung 10:

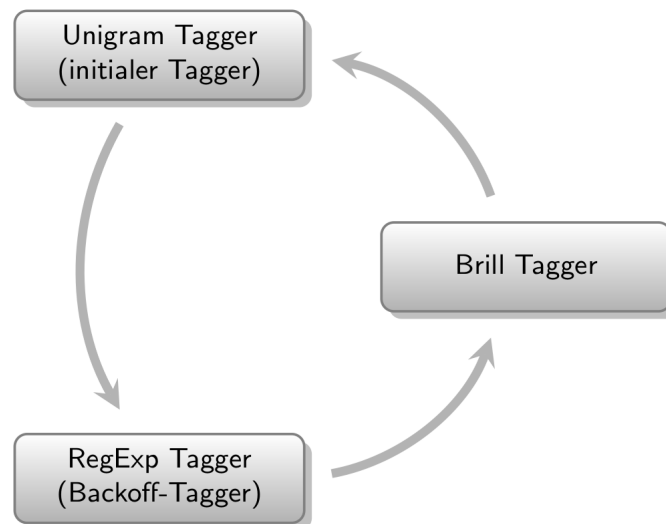


Abbildung 10: Tagging-Modell

Das *Tagging*-Modell besteht aus dem *Brill-Tagger* mit einem initial verwendeten *Unigram-Tagger* und einem als *Backoff-Tagger* verwendeten *RegExp-Tagger*. Ein *Unigram-Tagger* ist ein statistischer, sequentieller *Tagger*, der anhand der beobachteten Häufigkeiten in den Trainingsdaten jedes Token mit dem wahrscheinlichsten *Tag* versieht (vgl. Bird, Klein und Loper 2009: 202–203, Perkins 2014: 89–91). Zur Veranschaulichung der Funktionsweise sei angenommen, das automatisch zu annotierende Korpus bestünde nur aus dem

bislang schon häufiger zitierten Satz *Ua hana ana ‘o Herman me ka hale pauahi in Hilo*. Wäre beispielsweise im Trainings-Set *me* mehrheitlich als *PREP* annotiert, würde auch der *Unigram-Tagger* *me* in diesem Satz als *PREP taggen*. Eine wesentliche Einschränkung des *Unigram-Taggers* besteht darin, dass er nur aus den Trainingsdaten bekannte Tokens *taggen* kann. Würde z.B. *Herman* in den Trainingsdaten nicht auftreten, könnte der *Unigram-Tagger* *Herman* nicht als Eigenname identifizieren. Um auch unbekannte Tokens *taggen* zu können, kommt ein *Backoff-Tagger* (vgl. Bird, Klein und Loper 2009: 205–206, Perkins 2014: 92–93) ins Spiel, hier ein *RegExp-Tagger*, das heißt, ein *Tagger*, der auf *Regulären Ausdrücken* basiert (vgl. Bird, Klein und Loper 2009: 199, Perkins 2014: 99–100). Darin ist z.B. festgehalten, dass alle Wörter, die großgeschrieben sind, als Eigenname *getaggt* werden – Für diesem Fall wurden bei der Normalisierung ausschließlich Eigennamen in Großschreibung gelassen (vgl. Kapitel 3.4.2, S. 81). So würde *Herman* mithilfe des *Backoff-Taggers* als *Herman/NP* annotiert.

Der *RegExp-Tagger* hat vor allem bei zwei weiteren Problemen geholfen: Erstens wurden alle Tokens, die nicht vom *Unigram-Tagger* erfasst werden konnten, als *LEX* definiert. Zweitens konnten unbekannte Zahlwörter erfasst werden, da Wörter, die Zahlen größer als neun denotieren, im Hawaiischen präfigiert werden (vgl. Elbert und Pukui 1979: 158–161). So enthalten z.B. Zahlen zwischen 11 und 19 *‘umi* für 10 oder Zahlen zwischen 21 und 29 *iwakālua* für 20 und ein Token wie *iwakāluakūmākahi*²⁷ für 21 konnte problemlos als *NUM* *getaggt* werden, auch ohne in den Trainingsdaten aufzutauchen.

Schließlich dienen der *Unigram-Tagger* mit *RegExp-Tagger* als *Backoff-Tagger* dem *Brill-Tagger*. Der *Brill-Tagger* ist ein transformations-basierter *Tagger*, der Regeln entwickelt, um die falschen *Tags* des initialen *Taggers* zu korrigieren (vgl. Brill 1995, Bird, Klein und Loper 2009: 208–210, Perkins 2014: 102–105). Dabei berücksichtigt der *Brill-Tagger* anders als der *Unigram-Tagger* außerdem einen größeren Kontext, hier drei Tokens und drei *Tags* vor und nach dem Wort, das *getaggt* werden soll. Der Ablauf ist in Abbildung 11 (S. 90) beispielhaft am Satz *Ua hana ana ‘o Herman me ka hale pauahi in Hilo* festgehalten – die falschen *Tags* sind hervorgehoben.

²⁷Elbert und Pukui (1979: 159) schreiben solche Zahlwörter zwar auseinander oder mit Bindestrich, also z.B. *iwakālua kūmā-kahi*, im Korpus treten sie allerdings mehrheitlich zusammengeschrieben auf.

1. zu *taggender* Satz:
Ua hana ana ‘o Herman me ka hale pauahi in Hilo
2. initialer *Unigram-Tagger*:
Ua/UA hana/**NONE** ana/**PREP** ‘o/DEF Herman/**NONE**
me/PREP ka/DEF hale/LEX pauahi/**NONE** in/PREP Hilo/NP
./.
3. *RegExp-Tagger* als *Backoff-Tagger*:
Ua/UA hana/LEX ana/**PREP** ‘o/TOP Herman/NP me/PREP
ka/DEF hale/LEX pauahi/**LEX** in/PRE Hilo/NP ./.
4. *Brill-Tagger*:
Ua/UA hana/LEX ana/DEM ‘o/TOP Herman/NP me/PREP
ka/DEF hale/LEX pauahi/LEX2 in/PREP Hilo/NP ./.

Abbildung 11: Beispiel für kombiniertes Tagging-Modell

Nachdem das gesamte Analysekorpus vollständig *getaggt* wurde, konnten über das *Chunking* Wörter in ihrem Auftreten in syntaktischen Slots analysiert werden.

3.4.5 Chunking

Der Begriff *Chunking* oder *partielles Parsing* bezieht sich auf die Identifikation und Extraktion von Teilstrukturen in *getaggt*en Sätzen (vgl. Abney 1996, Bird, Klein und Loper 2009: 264, Hagenbruch 2010: 275-276, Perkins 2014: 123). Im Prinzip ist *Chunking* eine Form des *Tagging*, mit dem allerdings keine einzelnen Tokens, sondern spezifische Abfolgen von Tokens annotiert werden.

Die schrittweise Entwicklung bis zum *Chunking* der Sätze im Korpus ist in Abbildung 12 (S. 91) zusammengefasst.

Erst wenn die zu erfassenden Teilstrukturen möglichst genau definiert sind, kann ein geeignetes *Chunking*-Verfahren ausgewählt und an Teildaten des *getaggt*en Korpus getestet werden. Nach der Evaluation, Modifikation und Erweiterung des gewählten Verfahrens erfolgt schließlich das *Chunking* des gesamten Korpus.

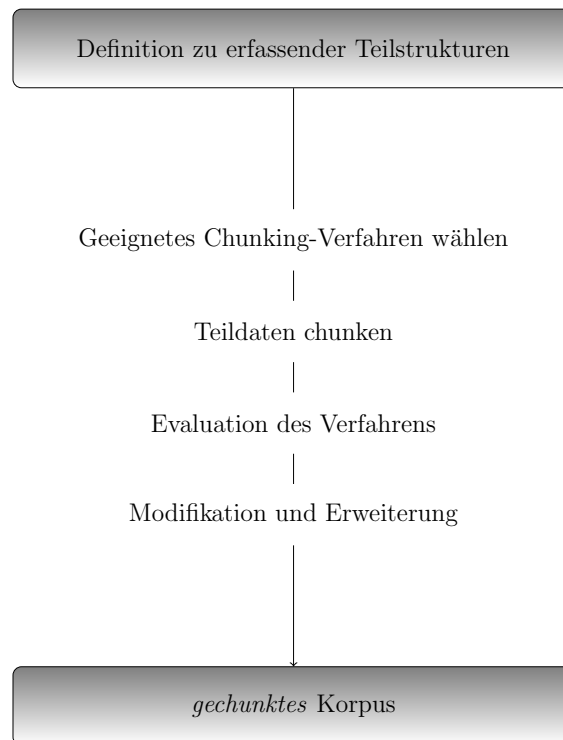


Abbildung 12: schrittweise Entwicklung zum Chunking

Da das Ziel der Korpusanalyse die Untersuchung der Prädisposition von Wörtern für das Auftreten in bestimmten syntaktisch-funktionalen Slots ist, müssen diese als die zu erfassenden Teilstrukturen definiert werden.

Auf der syntaktischen Ebene lassen sich im Hawaiischen zunächst Tempus-Aspekt-Modus-Syntagma und Artikel-Syntagma als komplementär zueinander verteilte strukturelle Einheiten unterscheiden (vgl. Kapitel 2.2.2, S. 15). Die funktionale Ebene ist in Korpusdaten nicht direkt erfassbar, sondern muss theoretisch über strukturelle Parameter operationalisiert werden. Dies stellt sich allerdings als problematisch heraus. Denn wie Völkel (2017: 448 Anmerkung 4) zurecht bemerkt hat, ist die z.B. von Hengeveld (1992) oder Croft (2001) vorgenommene Gleichsetzung von Tempus-Aspekt-Modus-Syntagma mit funktional definierter, prädikativer Phrase und von Artikel-Syntagma mit funktional definierter, referentieller Phrase nicht überzeugend, weil nicht nur Tempus-Aspekt-Modus-Syntagma, sondern auch Artikel-Syntagma in prädikativer Funk-

tion gebraucht wird (vgl. Kapitel 2.2.3, S. 19–20). Das bedeutet in der Konsequenz, dass hier nur strukturell definierte syntaktische Slots bzw. Syntagmen ohne Berücksichtigung der funktionalen Ebene untersucht werden können.

In einem Satz wie *Ua hana ana 'o Hermann me ka hale pauahi in Hilo* besetzt diesem Verständnis zufolge *hana* das Tempus-Aspekt-Modus-Syntagma *ua hana ana* und *hale pauahi* das Artikel-Syntagma *ka hale pauahi*. Zusätzlich kann ein weiterer Slot analysiert werden, den *pauahi* in *ka hale pauahi* besetzt und den ich mangels besserer Terminologie als „Attribut“ oder „attributiver Slot“ bezeichnen möchte, ohne dass sich der Begriff „Attribut“ hier auf eine bestimmte Wortart bezieht. Dazu muss gesagt werden, dass es mir einerseits unter der Annahme, nur syntaktische Slots untersuchen zu können, inkonsistent erscheint, Begriffe wie „Modifikation“ (z.B. Hengeveld 1992: 58 und passim, Croft 2001: 88 und passim) oder „Qualifier“ (Elbert und Pukui 1979: 41 und passim) zu verwenden, die sich auf die diskurspragmatische oder semantische Ebene beziehen.²⁸ Andererseits ist ein „Attribut“ genau genommen kein eigenständiger Slot, da es immer in einem Tempus-Aspekt-Modus-Syntagma oder einem Artikel-Syntagma enthalten ist.

Für das *Chunking* ist es dann notwendig, den elementaren Aufbau von Tempus-Aspekt-Modus-Syntagma und Artikel-Syntagma im Hawaiischen zu verstehen. Ich orientiere mich dafür im Wesentlichen an der Beschreibung der Phrasenstrukturierung von Elbert und Pukui (1979: 41), habe jedoch die Wortartlabels von Elbert und Pukui (1979: 41 und passim) mit den im *Tagset* verwendeten ersetzt.

Zunächst sei der Aufbau des Tempus-Aspekt-Modus-Syntagmas betrachtet (vgl. Abbildung 13, S. 93).

Die Konfiguration des Tempus-Aspekt-Syntagma kann in einen präponierten Teil, zentralen Teil und postponierten Teil gegliedert werden.

Der präponierte Teil wird von einem Tempus-Aspekt-Modus-Marker besetzt, z.B. dem Aspektmarker *ua*, der dem *Tagset* folgend als *UA* annotiert ist (vgl. Kapitel 3.4.4, S. 85). Das Zentrum wird besetzt von obligatorischem Inhaltswort und fakultativem Attribut, das wiederum aus einer Kombination einer oder mehrerer Inhalts- und Funktionswörter bestehen kann. Für das zentrale Inhaltswort vermeide ich hier den Begriff „Kopf“, da in der Rede von *Tempus-*

²⁸Offensichtlich hat auch Völkel (2017), die konsequent versucht, sich nur auf die syntaktische Ebene zu beziehen, Schwierigkeiten, einen passenden Terminus zu finden und spricht einfach von „TAM [X Y]“ und „ART [X Y]“ als „syntactic contexts“ (Völkel 2017: 463 und passim).

präponiert		zentral				postponiert								
E			attributiv											
I	+	LEX	+	LEX2	+	LEX3	+	PASS	+	DIR	+	ANAPH	+	PM
KE			-	LOC	-	LOC	-	PM	-		-	DEM	-	
UA				NUM		NUM								

Abbildung 13: Aufbau des TAM-Syntagmas im Hawaiischen

Aspekt-Modus-Syntagma genau genommen der entsprechende Marker als Kopf analysiert werden müsste (vgl. Kapitel 2.2.2, S. 15).

Im postponierten Teil können Passivmarker, Modalpartikeln, Direktionale, Demonstrativa und *ai*, das Elbert und Pukui (1979: 96–99) folgend als anaphorische Partikel mit *ANAPH* annotiert ist, auftreten. Funktionswörter im präponierten Teil sind immer auf feste Positionen verteilt. Das bedeutet, dass wenn beispielsweise der Passivmarker *‘ia* und das Demonstrativum *nei* im Syntagma auftreten, der Passivmarker vor dem Demonstrativum steht und das Demonstrativum nicht vor dem Passivmarker stehen kann.

Der Aufbau eines Artikel-Syntagmas²⁹ gestaltet sich etwas komplexer, da dieses deutlich mehr Elemente enthalten kann, (vgl. Abbildung 14, S. 94). Zwar lässt sich auch das Artikel-Syntagma in einen präponierten Teil, zentralen Teil mit fakultativem Attribut und einen postponierten Teil gliedern. Für die Besetzung des Zentrums ist allerdings kein Inhaltswort notwendig. Anstelle von Inhaltswörtern können im zentralen Teil eines Artikel-Syntagmas auch Eigennamen oder Funktionswörter wie Pronomina stehen.

²⁹Ich verstehe hier unter „Artikel“ nicht nur Artikel im engeren Sinne, also definiten und indefiniten Artikel, sondern analog zum Begriff „Determinierer“ (vgl. Bloomfield 1933: 203–206, Himmelmann 1997: 11–13) auch Demonstrativa und Possessiva, folge aber Elbert und Pukui (1979: 41 und passim) darin, Interrogativa und Quantoren bzw. Numeralia nicht zu den „Determinierern“ zu zählen.

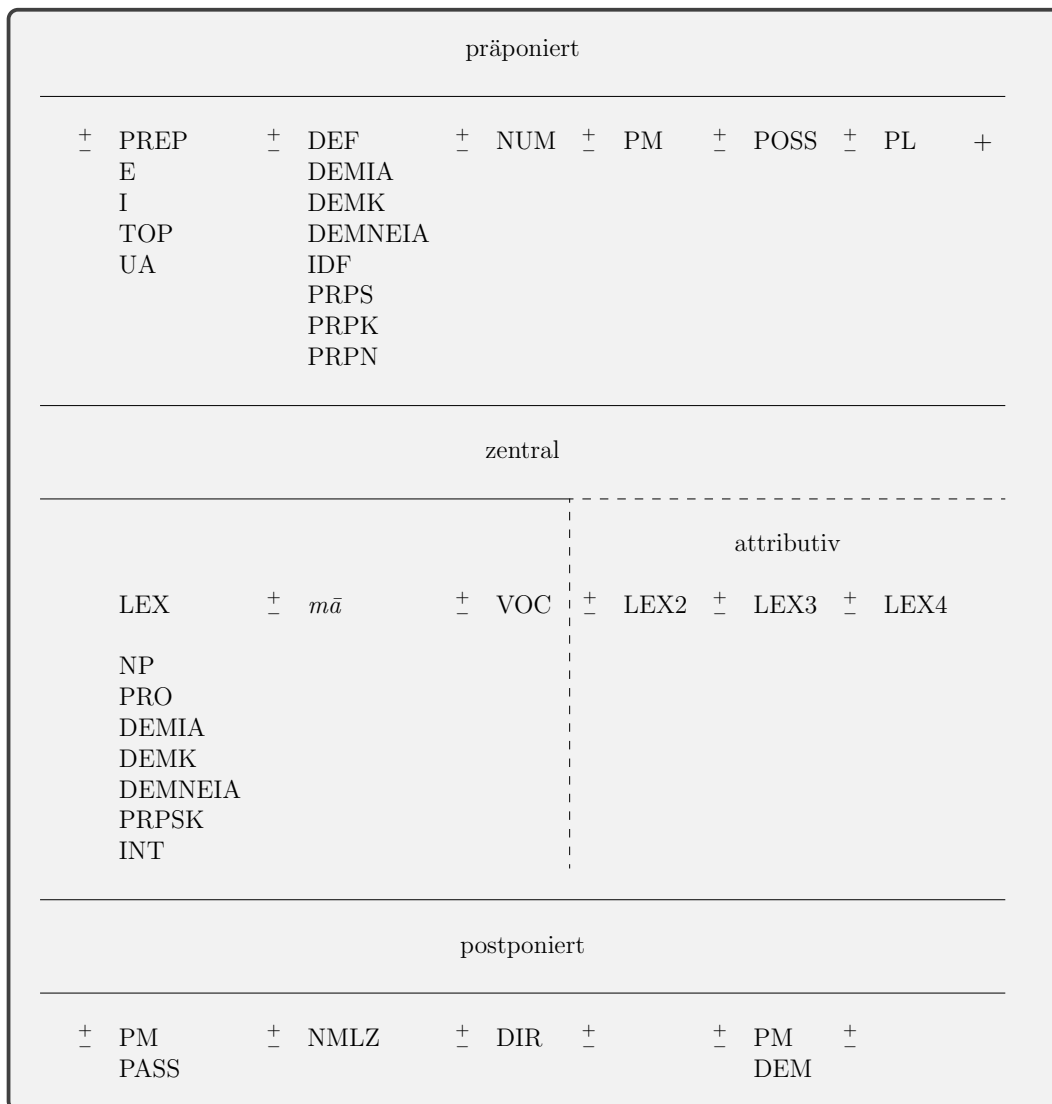


Abbildung 14: Aufbau des Artikel-Syntagmas im Hawaiischen

Der präponierte Teil bietet zudem im Vergleich zum Tempus-Aspekt-Modus-Syntagma deutlich mehr zu besetzende Stellen, die ebenfalls einer festen Ordnung folgen.

Im postponierten Teil besonders auffällig ist das von Elbert und Pukui (1979: 79 und passim) als „nominalizing particle“ bezeichnete und hier entsprechend mit *NMLZ* annotierte Funktionswort ‘*ana*. An dieser Stelle ist die Überlegung wichtig, dass Konstruktionen mit „Nominalisierungspartikel“ beim *Chunking* separat erfasst werden sollten. Denn falls die zu untersuchende Prädispositionsannahme zutrifft, müssten sich auch Unterschiede in der Verteilung von Wör-

tern auf diese Konstruktion zeigen. Die Problematisierung des Begriffs der „Nominalisierung“ im Hawaiischen erfolgt erst in der theoretischen Auseinandersetzung in Kapitel 4.1.3 (S. 131–167).

Nach einem genaueren Blick in den Aufbau von Tempus-Aspekt-Modus-Syntagma und Artikel-Syntagma im Hawaiischen kann nun festgehalten werden, dass folgende syntaktische Konstruktionen mit dem *Chunking* erfasst werden sollen:

1. Tempus-Aspekt-Modus-Syntagma
2. Artikel-Syntagma
3. „Nominalisierungskonstruktion“ mit *‘ana* (Artikel-Syntagma)
4. Attribut (in Tempus-Aspekt-Modus-Syntagma oder Artikel-Syntagma)

Zum Attribut sei hier noch folgendes anzumerken: Erstens unterscheide ich im Gegensatz zu Völkel (2017: 463 und passim) nicht zwischen Attribut in Tempus-Aspekt-Modus-Syntagma und in Artikel-Syntagma, weil die zu untersuchende Hypothese ausschließlich besagt, dass Wörter, dessen Bedeutung semantisch als EIGENSCHAFT klassifiziert ist, prädisponiert sind für den Gebrauch als Attribut in einem der beiden Syntagmen. Zweitens unterscheide ich, wie bereits in Kapitel 1.4.3 (S. 82) erklärt, anders als Elbert und Pukui (1979: 123–130) bei auseinander geschriebenen Wortfolgen nicht zwischen Attribut bzw. „qualifier“ und Kompositum.

Die Herausforderung beim *Chunking* besteht nun darin, die vier Konstruktionen im Korpus identifizieren und differenzieren zu können. Dafür müssen für jede Konstruktion die Kombinationsmöglichkeiten der Elemente erfasst werden.

Zur Veranschaulichung seien einmal nur die Kombinationsmöglichkeiten der präponierten Elemente in einem Artikel-Syntagma betrachtet: Geht man vereinfacht davon aus, dass auf den sechs Position jeweils nur ein bestimmtes Element stehen kann, z.B. auf der ersten Position nicht *PREP* oder *E* oder *I*, sondern nur *PREP* und dass mindestens ein Element steht, lassen sich mithilfe der summierten Binomialkoeffizienten $2^6 - 1 = 63$ Kombinationen berechnen (vgl. Abbildung 15, S. 96).

Positionen im präponierten Teil:

1	2	3	4	5	6
PREP	DEF	NUM	PM	POSS	PL

Anzahl der Kombinationen ohne Wiederholung
von k Elementen aus n Elementen:

$$\binom{n}{k} = \frac{n!}{(n-k)! \cdot k!}$$

Summe mit Binomialkoeffizienten:

$$\sum_{k=0}^n \binom{n}{k} = \binom{n}{0} + \binom{n}{1} + \dots + \binom{n}{k} = 2^n$$

da $\binom{n}{0} = 1$ ergeben sich $2^6 - 1 = 63$ Kombinationen

Abbildung 15: Berechnung der Kombinationsmöglichkeiten von präponierten Elementen

Berücksichtigt man nun alle präponierten, zentralen und postponierten Elemente einschließlich der potentiellen Besetzung einzelner Positionen durch unterschiedliche Elemente, wird ersichtlich, dass die tatsächliche Anzahl der Kombinationsmöglichkeiten in einem Artikel-Syntaxma deutlich höher ist als in diesem Beispiel.

Es sollte also ein *Chunking*-Verfahren gefunden werden, das es ermöglicht, alle Kombinationen zu erfassen. Ähnlich wie beim *Tagging* stehen mehrere Möglichkeiten bei der Realisierung des *Chunkings* zur Verfügung (vgl. Bird, Klein und Looper 2009: 266–270, Perkins 2014: 124–129). Eine gute Wahl schien hier, aufgrund des Aufbaus von Tempus-Aspekt-Modus-Syntaxma und Artikel-Syntaxma sowie der Kombinationsmöglichkeiten der verschiedenen Elemente einen *RegExp-Chunker* zu verwenden. Dessen Vorteil liegt vor allem darin, dass das Modell der Kombinationsmöglichkeiten direkt umgesetzt werden kann.

Die grundlegende Funktionsweise eines *RegExp-Chunkers* sieht in etwa folgendermaßen aus (vgl. Bird, Klein und Looper 2009: 264–281, Perkins 2014: 123–146):

1. Definiere eine „Grammatik“, die in Form von (modifizierten) *Regulären Ausdrücken* Regeln zum *Chunken* von Sätzen enthält.
2. Diese Regeln erfassen Abfolgen von *Tags*.
3. Die *Tags* der Korpussätze werden sequentiell abgerufen, bis eine in den Regeln festgelegte Abfolge erfasst ist, die dann zu einem *Chunk* gebündelt und mit einem Label versehen wird.

Der Ablauf des *Chunkings* mit dem *RegExpChunker* sei wieder vereinfacht an dem Beispielsatz *Ua hana ana 'o Hermann me ka hale pauahi in Hilo* verdeutlicht, vgl. Abbildung 16:

1. Satz im *getaggen* Korpus:
 Ua/UA hana/LEX ana/DEM 'o/TOP Hermann/NP me/PREP ka/DEF hale/LEX pauahi/LEX2 in/PREP Hilo/NP ./.
2. Regeln in der „Grammatik“:
 TAM: {<E|I|KE|UA>
 <LEX><LEX2|LOC|NUM>?
 <PASS|PM>?<DIR>?<ANAPH|DEM>?<PM>?}
 ART: {<PREP|E|I|TOP|>?<DEF|DEMIA|DEMK|DEMNEIA|IDF>?<PM|PL>?
 <LEX|NP|PRO><LEX2|LOC|NUM>?
 <PM|PASS>?<DIR|DIR>?<PM|DEM>?}
3. Repräsentation von *Chunks*:
 (TAM ua/UA hana/LEX ana/DEM)
 (ART 'o/TOP Herman/NP)
 (ART me/PREP ka/DEF hale/LEX pauahi/LEX2)
 (ART in/PREP Hilo/NP)
 ./.

Abbildung 16: Beispiel Chunking-Ablauf

Die Grundlage für das *Chunking* ist das Korpus mit *getaggen* Sätzen. In der „Grammatik“ ist festgelegt, dass eine bestimmte Abfolge von *Tags* als Artikel-Syntagma *gechunkt* und mit *ART* annotiert, eine bestimmte Abfolge von *Tags*

als Tempus-Aspekt-Modus-Syntagma *gechunkt* und mit *ART* annotiert werden soll. In der Syntax der *Chunking*-Regeln steht das Symbol „|“ für die ausschließende Disjunktion und das Symbol „?“ für Optionalität. Die Regel für das Tempus-Aspekt-Modus-Syntagma besagt beispielsweise, dass auf ein *E* oder *I* oder *KE* oder *UA* ein *LEX* folgen muss, auf das ein *LEX2* oder *LOC* oder *NUM* folgen kann, auf das wiederum ein *PASS* oder *PM* folgen kann usw. So kann die Abfolge *Ua/UA hana/LEX ana/DEM* erfasst und mit *TAM* annotiert werden.

Die verwendeten *Chunking*-Regeln wurden schrittweise in der Anwendung auf einzelne Texte entwickelt und mussten immer wieder modifiziert und erweitert werden. Die endgültig erstellte *Chunking*-Grammatik umfasst schließlich komplexere und deutlich mehr Regeln als im Beispiel in Abbildung 16, da es einige Besonderheiten zu beachten gilt, wie die Sequentialität des *Chunkers*, das Erfassen von eingebetteten Phrasen oder unerwartete Abweichungen des Phrasenstrukturmodells von Elbert und Pukui (1979: 41), das als Grundlage gedient hat. Solche Abweichungen betreffen z.B. die als *PM getaggt* Modalpartikeln, von denen häufig mehrere nacheinander und auch in fast allen Positionen in einem Syntagma auftreten können.

Die „Nominalisierungs konstruktion“ wurde analog zu Artikel-Syntagma und Tempus-Aspekt-Modus-Syntagma erfasst, indem das Auftreten des Markers *‘ana* bzw. des *NMLZ-Tags* als obligatorisch in der Regel angegeben wurde. Der attributive Slot muss nicht *gechunkt* werden, da Wörter darin bereits über das *Tagging*, z.B. als *LEX2* (vgl. Kapitel 3.4.4, S. 85), erfasst sind.

Mit dem angewendeten *Chunking*-Regeln konnten schließlich bestimmte Konstruktionen nicht als Bestandteile von einzelnen Syntagmen erschlossen werden, insbesondere solche wie *a pau*, *a laila* oder *a puni* (vgl. Kapitel 3.4.3, S. 82). Für die Distributionsanalyse ist das allerdings nicht weiter problematisch, weil diese Konstruktionen nicht aus Inhaltswörtern bestehen (vgl. Kapitel 3.4.4, S. 86).

Nach dem *Chunking* des gesamten Korpus sollten die Auftretenshäufigkeiten von Wörtern in den definierten Konstruktionen zusammengefasst und ausgewertet werden.

Vor der endgültigen Auswertung sind allerdings noch weitere Bearbeitungsschritte notwendig, die im nächsten Kapitel erläutert werden.

3.4.6 Weitere Schritte und Typisierung

Aus dem *gechunkten* Korpus konnte mithilfe von *Python* eine Liste erstellt werden, die alle Inhaltswörter und die jeweilige Auftretenshäufigkeit eines Inhaltswortes in den vier syntaktischen Konstruktionen enthält. Insgesamt umfasst diese Liste 3667 Wort-Types und 333.163 Wort-Tokens. Tabelle 9 zeigt einen Ausschnitt aus der Gesamttabelle, vgl.:

LEX	TAM	ART	ATTR	NMLZ	FREQ
\bar{a}	0	4	28	0	32
.
.
.
wohi	0	6	4	0	10

Tabelle 9: Ausschnitt Tabelle mit Frequenzen von Types in syntaktischen Konstruktionen

In Tabelle 9 ist zu sehen, dass das Wort \bar{a} im Korpus insgesamt 32 mal, davon 0 mal in Tempus-Aspekt-Modus-Syntagma, 4 mal in Artikel-Syntagma, 28 mal als Attribut und 0 mal in „Nominalisierungs-konstruktion“ und das Wort *wohi* insgesamt 10 mal, davon 0 mal in Tempus-Aspekt-Modus-Syntagma, 6 mal in Artikel-Syntagma, 4 mal als Attribut und 0 mal in „Nominalisierungs-konstruktion“ auftritt.

Bevor es nun möglich ist, die (Häufigkeits-)Verteilung der Wörter auf die syntaktischen Konstruktionen genauer auszuwerten, sind noch weitere Bearbeitungsschritte notwendig.

Dazu gehören eine erneute Fehlerkorrektur, die Entfernung niedrig-frequenter Wörter und bestimmter Problemfälle und das, was ich hier als „Typisierung“ bezeichne.

In einem ersten Schritt wurden beim Durcharbeiten der Liste weitere Fehler identifiziert und die betroffenen Formen manuell aus der Liste entfernt. Dazu gehören unsystematisch fehlerhafte Wortformen, die teilweise aus dem Quelltext übernommen und in der *Tokenisierung* nicht richtig erfasst wurden wie *‘*oia*‘*o* oder *‘*i*‘. Ebenso mussten bislang übersehene *Tag*-Fehler, z.B. **ko-lu*/*LEX* statt *kolu*/*NUM* oder **iō*/*LEX* statt *iō*/*PREP*, korrigiert und die Wörter entsprechend aus der Liste genommen werden.

Als nächstes folgten mehrere Einzelschritte, die ich in Analogie zur *Tokenisierung* (vgl. Kapitel 3.4.3, S. 82) und in Annäherung an den Begriff des *Stem-*

ming bzw. der *Lemmatisierung* (vgl. Manning und Schütze 1999: 132–134, Bird, Klein und Looper 2009: 107–108, Perkins 2014: 30–32) unter „Typisierung“ zusammenfasse. Zugrunde liegt hier die Vermutung, dass manche in die Liste aufgenommene Wörter aufgrund verschiedener Phänomene als Wortformen einzelner Wort-Types betrachtet werden müssen. Zu diesen Phänomenen gehören die graphische Realisierung von lautlicher Variation, der morphologisch markierte Plural und Homonymie, die im Folgenden genauer betrachtet werden sollen.

Da das Hawaiische nicht standardisiert ist (vgl. Kapitel 3.3.2, S. 74), finden sich häufig verschiedene Schreibweisen von Wörtern. Anders als im Beispiel von *mea kanu* und *meakanu*, die nicht zu einem Type zusammengefasst werden sollten (vgl. Kapitel 3.4.3, S. 82), geht es bei der „Typisierung“ in Bezug auf die Schreibweise um die Berücksichtigung der graphischen Realisierung von lautlicher Variation. Die lautliche Variation betrifft den Einschub eines Gleitlauts, die Verschmelzung von Vokalen, die Vokallänge, die Alternation von Konsonanten, Fremdwörter und verschiedene Formen der Reduplikation.

Zum Gleitlaut: In der Umgebung zwischen dem Vokal *u* oder *o* und darauffolgendem Vokal kann als Gleitlaut ein labialisierter stimmhafter velarer Approximant [w] eingeschoben werden, den Elbert und Pukui (1979: 13) als „phonetic and predictable glide“ bezeichnen und der in manchen Texten graphisch mit <w> realisiert wird, z.B. <kauwila> statt <kauila> (vgl. Elbert und Pukui 1979: 12). Im Korpus treten die Paare <uaki> / <waki> und <‘auae> / <‘auwae> auf.

Zur Vokalverschmelzung: Zwei aufeinander folgende Vokale werden häufig verschmolzen, z.B. wird *aloha ‘ia aku* „geliebt“ zu *aloha ‘iaku* (Elbert und Pukui 1979: 15; vgl. auch Elbert und Pukui 1979: 22–23). Im Korpus finden sich *hoaloha* / *hoaloha* „Freund“ aus *hoa* „begleiten, Begleiter“ und *aloha* „lieben, geliebt, Liebe“ und *luaahi* / *luahi* „Feuerstelle, Opfer“ mit *lua* „Stelle“ und *ahi* „brennen, Feuer“.

Zur Vokallänge: In manchen Fällen variiert die Vokallänge, ohne dass davon die phonologische Unterscheidung betroffen ist. Das tritt z.B. beim Kausalpräfix in den Formpaaren *ho-/hō-* und *ha-/hā-* auf. Andere Beispiele im Korpus sind *haumana/haumāna* „Schüler“ oder *papālina/pāpālina* „Kinn“, von PEP **pa(a)paalinga*.

Zur Alternation von Konsonanten: Eine vermutlich dialektale Variation auf den Inseln Moloka‘i und Lāna‘i (vgl. Elbert und Pukui 1979: 25–26) ist die Substitution von *l* durch *n*, z.B. in *hānana* für *hālana* (Elbert und Pukui 1979:

25), wobei die *l*-Formen mehrheitlich Reflexe der protopolynesischen Formen und die *n*-Formen hawaiische Innovationen darstellen (vgl. Elbert und Pukui 1979: 26). Im Korpus kommen vor: *kūnou* / *kūlou* von PPN **tulou* und *nanea* / *walea* von PNN **walea*, wobei die *n-w*-Alternation am Wortanfang unklar ist. Eine weitere, nicht-dialektale Variation ist die Alternation zwischen *l* oder *k* und dem Glottal, z.B. *heka*/*he'a* (vgl. Elbert und Pukui 1979: 12, 26–27), wobei vermutlich *k* die protopolynesische Form reflektiert, während der Glottal die hawaiische Innovation darstellt (vgl. die Überblickstabelle in Elbert und Pukui 1979: 36). Im Korpus treten neben *haki* / *ha'i* mit PNN **faki* und daraus gebildeten Stämmen wie der Reduplikationsform *hakahaki* / *ha'ihai* auch weitere nicht ohne weiteres aufschlüsselbare Variationen mit dem Glottal auf, wie *pōhai* / *pō'ai* mit PCP **pookai*.

Zu Fremdwörtern: Die Orthographie der mehrheitlich aus dem Englischen entlehnten Fremdwörter³⁰ weist große Uneinheitlichkeiten auf. Häufig orientiert sich die Schreibweise zwischen der englischen Schreibweise und der hawaiianischen Aussprache, z.B. *ti* versus *kī*, von engl. „tea“ (vgl. Elbert und Pukui 1979: 27). Im Korpus finden sich sechs solcher Variationen – am auffälligsten ist mit vier verschiedenen Schreibweisen *mikanele* / *mikionali* / *misionari* / *mikioneli* „Missionar, missionarisch“.

Zur Reduplikation: Es lässt sich unterscheiden zwischen vollständiger und partieller Reduplikation, die Elbert und Pukui (1979: 22) der „natural fast speech“ zuweisen, vgl. *makemake* versus *mamake* mit der Basis *make* (vgl. Elbert und Pukui 1979: 22). Im Korpus kommen insgesamt 23 Paare aus vollständiger und partieller Reduplikation vor, darunter *nahunahu* / *nanahu*, *huhui* / *huhui* oder *li'ili'i* / *lili'i*.

Im Korpus treten weitere lautliche Variationen auf, die keinen der zuvor betrachteten zugeordnet werden können, z.B. *napo'o* / *kapo'o* „versinken, Höhle“ oder *nīpu'u* / *hīpu'u* „binden, Knoten“. Desweiteren bestehen parallele Formen, bei denen der Zusammenhang unklar ist, z.B. *hahau* / *uhau* „schlagen, Schlag“ und *hahai* / *uhai* „folgen, Anhänger“.

Alle im Korpus auftretenden Paare lautlicher Variation wurden dann in der Liste jeweils zu einem Type zusammengefasst und die Frequenzen entsprechend

³⁰Auch auf das Griechische, Lateinische und Hebräische zurückführbare Wörter sind über das Englische in die hawaiische Sprache gekommen, besonders unter dem Einfluss der Bibelübersetzung (vgl. Elbert und Pukui 1979: 31–33); weitere wenige Fremdwörter stammen aus anderem indogermanischen Sprachen, dem Chinesischen oder dem Tahitischen (vgl. Elbert und Pukui 1979: 33–34).

addiert.

Ein weiterer Aspekt der „Typisierung“ betrifft den Umgang mit der nicht produktiven Pluralbildung durch Vokalverlängerung, z.B. bei *wahine* „weiblich, Frau“ und *wāhine* „weiblich, Frauen“. Elbert und Pukui (1979: 106) nennen insgesamt 10 Wörter mit dieser Pluralbildung, die sich auf Menschen beziehen und teilweise zur Verwandtschaftsterminologie gezählt werden können. In Tabelle 10 sind die von Elbert und Pukui (1979: 106) angeführten und die im Korpus gefundenen Formen zusammengestellt, vgl.:

Singularform	Pluralform	Bedeutung	EP	Korpus
‘aumakua	‘aumākua	Familiengötter	✓	✓
‘elemakule	‘elemākule	alte Männer	✓	✓
kahiko	kāhiko	alte Menschen	✓	✓
kahuna	kāhuna	Priester	✓	✓
kahunapule	kāhunapule	Prediger	✗	✓
kaikamahine	kaikamāhine	Mädchen	✓	✓
kanaka	kānaka	Menschen	✓	✓
kupuna	kūpuna	Großeltern	✓	✓
kupunakāne	kūpunakāne	Großväter	✗	✓
luahine	luāhine	alte Frauen	✓	✗
makua	mākua	Eltern	✓	✓
makuahine	mākuahine	Mütter	✗	✓
makuahōnōwai	mākuahōnōwai	Schwiegereltern	✗	✓
wahine	wāhine	Frauen	✓	✓

Tabelle 10: Vergleich der Angaben zu morphologisch markierten Pluralformen von Elbert und Pukui (1979: 106) und der im Korpus gefundenen Formen

Die einzige von Elbert und Pukui (1979: 106) angegebene Pluralform, die nicht im Korpus vorkommt, ist *luāhine*. Zusätzlich treten allerdings drei Wörter in zusammengeschiedener Variante auf: *kahunapule* / *kāhunapule* als Variante von *kahuna pule* / *kāhuna pule*, *kupunakāne* / *kūpunakāne* als Variante von *kupuna kāne* / *kūpuna kāne* und *makuahōnōwai* / *mākuahōnōwai* als Variante von *makua hōnōwai* / *mākuua hōnōwai* (vgl. dazu Kapitel 3.4.3, S. 82). Wie zuvor bei der lautlichen Variation wurde jedes der Singular-Plural-Paare zu einem Type zusammengefasst und der Eintrag der Häufigkeiten geändert.

Die größte Herausforderung bei der „Typisierung“ aber besteht im Umgang mit Polysemie und Homonymie. Dabei geht es um die Frage, ob in bestimmten Fällen trotz identischer Form zwei Lexeme voneinander unterschieden werden

können, unabhängig von der zuvor betrachteten Diskussion um die Analyse von z.B. *ako* als monosemes oder polysemes Lexem und der damit verbundenen Entscheidung über die Wortartkategorisierung (vgl. Kapitel 2.2.4, S. 52–55).

So erscheint es beispielsweise bei der Wortform *leka* plausibel, zwischen einem Lexem *leka*₁ mit der Bedeutung „klebrig, Schleim“ und einem Lexem *leka*₂ mit der Bedeutung „Brief, einen Brief schreiben“ zu unterscheiden, da letzteres eine Entlehnung aus dem Englischen (vgl. engl. *letter*) darstellt und unmittelbar kein Zusammenhang zwischen den Bedeutungen erkennbar ist. Ich orientiere mich hier an Bréal (1897) und den Ausführungen von Ullmann (1967: 99–128) und Lyons (1977: 550–570) zur Definition und Unterscheidung von Polysemie und Homonymie³¹, auch wenn mir bewusst ist, dass eine klare Trennung zwischen diesen beiden Probleme aufwerfen mag (vgl. Tuggy 1993, Geeraerts 1993). Zur Bestimmung und Unterscheidung von Polysemie und Homonymie können zwei Kriterien herangezogen werden: Nach dem diachronen Kriterium sind Wörter polysem, wenn sie auf eine Form und homonym, wenn sie auf verschiedene Formen zurückführbar sind. Nach dem synchronen Kriterium sind Wörter polysem, wenn ein Bedeutungszusammenhang erkennbar und homonym, wenn keiner erkennbar ist.

Ist die Anwendung des diachronen Kriteriums solange noch möglich, wie Informationen zu den Proto-Formen der einzelnen Wörter vorhanden sind, stellt sich die Entscheidung über den semantischen Zusammenhang beim synchronen Kriterium in vielen Fällen als schwierig heraus.

Es ist kaum entscheidbar, in welchen Fällen für den Hawaiiisch-Sprecher ein gemeinsamer Bedeutungskern vorliegt und in welchen nicht. Da aber die zu überprüfende Hypothese von semantisch klassifizierten Wörtern letztlich die Differenzierungsversuche der vergleichenden Semantik zwischen Polysemie und Homonymie mit einschließt, soll diesen hier gefolgt werden.

Zur Unterscheidung von polysemen und homonymen Wörtern habe ich mich schließlich an den Angaben in Wörterbüchern (vgl. Andrews 1865, Parker 1922, Pukui und Elbert 1971) und am *Polynesian Lexicon Project Online* (vgl. Greenhill und Clark 2011) orientiert.

Die als polysem analysierten Wörter wurden dann in der Liste beibehalten. Die als homonym analysierten Wörter mussten hingegen aus der Liste entfernt

³¹Da in einem schriftlichen Korpus die lautliche Ebene nicht berücksichtigt werden kann, handelt es sich genau genommen um *Homographie*, das heißt um die Identität der graphemischen Repräsentation.

werden, da es angesichts der Tokenanzahl von 333.163 nicht möglich war, manuell alle Auftretensweisen zu disambiguieren.

Nach den Schritten, die hier unter dem Begriff „Typisierung“ zusammengefasst wurden, folgte die Entfernung weiterer Wörter aus der Liste.

Davon betroffen sind zunächst Wörter, die nicht eindeutig der Kategorie Inhaltswort oder Funktionswort zugewiesen werden konnten, z.B. *po‘e* (vgl. Kapitel 3.4.4, S. 86) oder *mea*, das unter anderem in Artikel-Syntagma mit der Bedeutung „Ding, Person“, aber auch in vielen verschiedenen Konstruktionen, z.B. *no ka mea* „weil“ oder *he mea aha?* „warum?“, auftritt.

Ebenso mussten Wörter mit niedrigfrequentem Auftreten aus der Liste entfernt werden, weil diese für die weitere Analyse nicht brauchbar sind.

In der Korpus- und Computerlinguistik werden üblicherweise *Hapax Legomena*, von gr. ἅπαξ ⟨hapax⟩ „einmal“ und λεγόμενον ⟨legómenon⟩ „was gesagt wird“, i.e. nur einmal im Korpus vorkommene Tokens, aus dem Datensatz entfernt, da z.B. Probleme mit der Repräsentativität (vgl. Brezina 2018: 38–65) oder Probleme mit fehlerhaftem *Tagging* und der fehlerhaften *Tokenisierung* bei automatisierten Verfahren auftreten können (vgl. Jurafsky und Martin 2009: 204–314, Bird, Klein und Loper 2009: 17–20 und passim).

Da in der Distributionsanalyse untersucht werden sollte, ob und wie häufig Wörter in vier definierten syntaktischen Konstruktionen auftreten, wurden nicht nur *Hapax*, sondern zusätzlich *Dis* und *Tris Legomena*, also alle Wörter mit einer gesamten Auftretenshäufigkeit niedriger als vier, aus der Liste entfernt.

Nach der Korrektur von Fehlern, der „Typisierung“ und der Entfernung von Homonymen, nicht eindeutig kategorisierbaren Wörtern und *Hapax*, *Dis* und *Tris Legomena*, enthält die anfangs aus dem *gechunkten* Korpus entstandene Liste schließlich 2250 Types und 243.514 Tokens.

Vor einer ersten Auswertung der Liste soll im nächsten Kapitel eine Zusammenfassung aller bisherigen Verfahrensschritte folgen.

3.5 Zusammenfassung

In Abbildung 17 (S. 105) ist noch einmal der gesamte Ablauf der Korpusanalyse, von der Erstellung des Korpus bis zur automatisierten Analyse zusammengefasst.

Nachdem zu Beginn die Entscheidung getroffen wurde, keine eigenen Daten zu erheben, sondern *Ulukau – The Hawaiian Electronic Library* als Datengrund-

lage zu verwenden, sind die Daten darin sortiert und für die Aufnahme in das Analysekorpus ausgewählt worden.

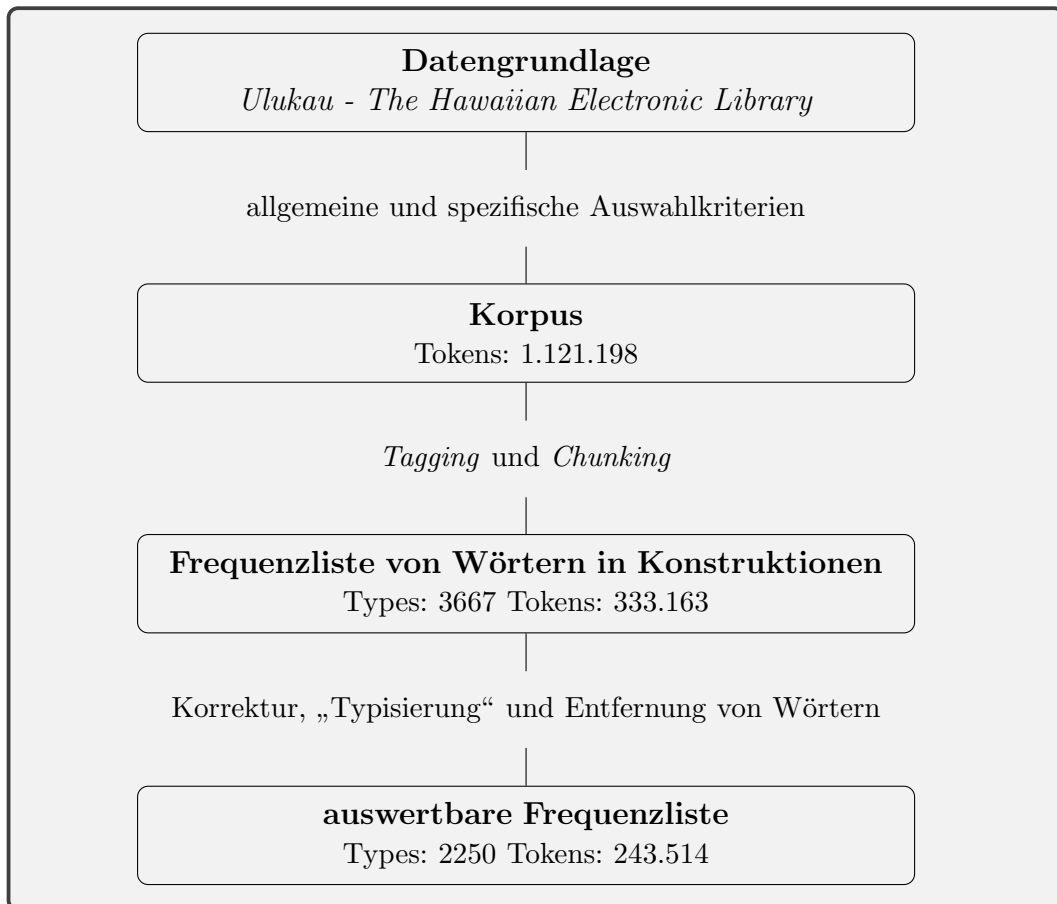


Abbildung 17: Zusammenfassung Korpusanalyse

Als Auswahlkriterien haben neben allgemeinen Kriterien, zu denen die Orientierung am Untersuchungsgegenstand und die Repräsentativität gehören, die am Entstehungshintergrund der Daten entwickelten Kriterien der Muttersprachlichkeit der Verfasser und der Orthographie der Texte gedient.

Das so erstellte Korpus umfasst 1.121.198 Tokens.

Um die Distribution und Auftretenshäufigkeit der Inhaltswörter in syntaktisch-funktionalen Slots im Korpus zu untersuchen, wurde schließlich auf ein automatisiertes Verfahren zurückgegriffen.

Durch das *Tagging* und *Chunking* des Korpus konnte eine Liste erstellt wer-

den, die alle im Korpus auftretenden Inhaltswörter (3667) und ihre Auftretenshäufigkeit in Tempus-Aspekt-Modus-Syntagma, Artikel-Syntagma, „Nominalisierungs-konstruktion“ und attributivem Gebrauch enthält.

Erst durch zusätzliche Bearbeitungsschritte dieser Liste, die eine weitere Fehlerkorrektur, die „Typisierung“, die Entfernung von nicht eindeutig als Inhaltswörter klassifizierbaren Wörtern und die Entfernung von Wörtern mit einer Gesamtfrequenz von niedriger als vier umfassen, konnte eine endültig für die Auswertung geeignete Liste erstellt werden, die 2250 Worttypes und 243.514 Worttokens umfasst.

Die Auswertung der Distribution und Auftretenshäufigkeit von Wörtern und sich daraus ergebende weitere Analysen sind Gegenstand des nächsten Kapitels.

4 Auswertung und Ergebnisse

Nachdem mithilfe eines automatisierten Verfahrens und darauf folgenden Bearbeitungsschritten eine Liste aus den Korpusdaten erstellt wurde, die die Auftrenshäufigkeit von 2250 Inhaltswörtern in Tempus-Aspekt-Modus-Syntagma, Artikel-Syntagma, „Nominalisierungs-konstruktion“ und attributivem Slot enthält, kann diese Liste nun ausgewertet werden.

Entsprechend der zu überprüfenden Hypothese, die besagt, dass im Hawaiischen Wörter nicht über Formklassen sondern über lexikalische Bedeutungsklassen für das Auftreten in bestimmten Syntagmen prädisponiert sind (vgl. Kapitel 2.3, S. 60), geht es bei der Auswertung der Liste zunächst um die Frage, ob die Wortverteilung auf Syntagmen im Korpus Hinweise auf Distributionsklassen im Hawaiischen liefert (Kapitel 4.1). Es wird gezeigt, dass dies nicht der Fall ist und daraus Probleme für die Konzeption von *Nominalisierung* und *verbaler* Morphologie im Hawaiischen entstehen.

Als nächstes folgt die Auswertung hinsichtlich potentieller Frequenzklassen (Kapitel 4.2). Auch hier kann die Annahme, die Prädisposition von Wörtern sei zumindest als prototypische Auftretensweise über die Gegenüberstellung der Auftrenshäufigkeiten von Wörtern in einzelnen Syntagmen nachweisbar (vgl. Kapitel 2.2.4, S. 50–52), nicht bestätigt werden.

Unter Berücksichtigung dieser Ergebnisse wird schließlich dargelegt, dass *Wortarten* im Hawaiischen nicht über lexikalisierte ontologische Klassen von Wortbedeutungen bestimmbar sind (Kapitel 4.3).

4.1 Wortverteilung und Distributionsklassen

Im Kern geht es bei der Untersuchung der Wortdistribution um die Frage, ob Wörter für das Auftreten in ganz bestimmten (Kombinationen aus) Syntagmen prädisponiert sind. Zu einer wesentlichen Motivation dieser Frage gehört auch die Unklarheit darüber, ob Wörter im Hawaiischen mehrheitlich in nur einem Syntagma auftreten und damit *Wortarten* angesetzt werden müssen (vgl. Kapitel 2.2.1, S. 11–14).

Im Folgenden werden die Ergebnisse aus der Untersuchung der Verteilung von Wörtern auf Tempus-Aspekt-Modus-Syntagma, Artikel-Syntagma, „Nominalisierungs-konstruktion“ und attributivem Slot im Korpus präsentiert.

Aus der Diskussion um die Bildung von Distributionsklassen anhand der Wortverteilung im Zusammenhang mit der Bestimmung von *Wortarten* geht die

Problematisierung der Konzeption von *Nominalisierung* und *verbaler Morphologie* im Hawaiischen hervor und schließlich die Ablehnung der Prädispositionsannahme.

4.1.1 Überblick der Wortverteilung im Korpus

Tabelle 11 bietet einen Überblick über das Auftreten von Wörtern in einzelnen Syntagmen und Kombinationen aus Syntagmen im Korpus, vgl.:

Syntagma	absolute Frequenz		relative Frequenz	
	Types	Tokens	Types	Tokens
TAM	54	448	0,02	0,00
ART	115	1415	0,05	0,00
ATTR	85	958	0,04	0,00
NMLZ	0	0	0,00	0,00
Σ	254	2821	0,11	0,01
TAM \wedge ART	109	437 + 542	0,05	0,00 + 0,00
TAM \wedge ATTR	100	386 + 528	0,04	0,00 + 0,00
TAM \wedge NMLZ	40	291 + 83	0,02	0,00 + 0,00
ART \wedge ATTR	453	13946 + 4371	0,20	0,05 + 0,02
ART \wedge NMLZ	6	95 + 9	0,00	0,00 + 0,00
ATTR \wedge NMLZ	4	18 + 5	0,00	0,00 + 0,00
Σ	712	20711	0,32	0,08
TAM \wedge ART \wedge ATTR	492	3587 + 29286 + 24529	0,22	0,02 + 0,12 + 0,1
TAM \wedge ART \wedge NMLZ	67	1094 + 223 + 187	0,03	0,00 + 0,00 + 0,00
TAM \wedge ATTR \wedge NMLZ	52	428 + 167 + 92	0,02	0,00 + 0,00 + 0,00
ART \wedge ATTR \wedge NMLZ	31	2550 + 606 + 45	0,01	0,01 + 0,00 + 0,00
Σ	642	62794	0,29	0,26
TAM \wedge ART \wedge ATTR \wedge NMLZ	642	57227 + 55693 + 34812 + 9456	0,29	+ 0,23 + 0,23 + 0,14 + 0,04
Σ	642	157188	0,29	0,64
Σ	2250	243514	1	1

Tabelle 11: Überblick Wortverteilung im Korpus

In der ersten Spalte von Tabelle 11 (S. 108) sind die Syntagmen bzw. die Kombinationen von Syntagmen festgehalten, in denen ein Wort auftritt, also z.B. in Zeile 4 Tempus-Aspekt-Modus-Syntagma (TAM), in Zeile 10 Tempus-Aspekt-Modus-Syntagma und Artikel-Syntagma (TAM \wedge ART) oder in Zeile 24 alle vier syntaktischen Konstruktionen (TAM \wedge ART \wedge ATTR \wedge NMLZ).

Zu jeder Kombination von Syntagmen sind in der jeweiligen Zeile die absoluten und relativen³² Frequenzen der Types und Tokens angegeben. In Zeile 4 lässt sich beispielsweise ablesen, dass 54 Wörter (Types) im Korpus ausschließlich in Tempus-Aspekt-Modus-Syntagma (TAM) auftreten und eine Gesamtfrequenz (Tokens) von 448 aufweisen. Die Spalte mit den Tokens gibt nur die Gesamtanzahl der Tokens aller ausschließlich in Tempus-Aspekt-Modus-Syntagma auftretenden Types an und differenziert nicht zwischen den Tokens der einzelnen Types. Die relative Frequenz ist die Häufigkeit in Relation zur Gesamthäufigkeit der Worttypes (2250) bzw. Worttokens (243.514) im Korpus. In Zeile 4 erkennt man z.B., dass 2 % aller im Korpus vorkommenden Inhaltswörter (Types) ausschließlich in Tempus-Aspekt-Modus-Syntagma auftreten, die mit 448 von 243.514 Tokens allerdings nicht einmal ein Prozent der Gesamttokenanzahl umfassen.

In den Fällen, in denen ein Wort in mehr als einem Syntagma auftritt, z.B. in Tempus-Aspekt-Modus-Syntagma und Artikelsyntagma (siehe Zeile 10), ist die Tokenanzahl für das jeweilige Syntagma einzeln angegeben. Der Grund dafür ist, dass ein Token als einmaliges Wortvorkommen nicht in zwei Syntagmen auftreten kann, sondern immer nur in einem Syntagma. Beispielsweise sind die Angaben in der dritten Spalte von Zeile 10 dann so zu verstehen, dass von den 109 Wörtern (Types), die sowohl in Tempus-Aspekt-Modus-Syntagma als auch Artikel-Syntagma vorkommen, insgesamt 437 Tokens in Tempus-Aspekt-Modus-Syntagma und 542 Tokens in Artikel-Syntagma auftreten.

Für einen möglichen Vergleich der Distribution sind in Tabelle 11 (S. 108) schließlich die verschiedenen Kombinationen von Syntagmen in Abschnitte getrennt und die summierten absoluten und relativen Häufigkeiten der jeweiligen Kombinationen am Ende eines Abschnittes angegeben. In den Summen der Kombinationen, die aus mehr als einem Syntagma bestehen, sind des besseren Überblicks wegen die Tokenfrequenzen der einzelnen Syntagmen addiert. In der dritten Spalte der 16. Zeile (grau eingefärbt) ist z.B. angegeben, dass 20.711

³²Die Zahlen wurden auf zwei Nachkommastellen aufgerundet, so dass ein Wert von 0,00 nicht als tatsächlich null Elemente zu lesen ist.

Wortinstanzen (Tokens), die im Korpus vorkommen, Instanzen (Tokens) von Wörtern (Types) sind, die in zwei Syntagmen vorkommen, ohne genaue Unterscheidung, in welchem jeweiligen Syntagma die Tokens auftreten.

So kann einerseits direkt verglichen werden, wie viele Wörter (Types) ausschließlich in einem Syntagma und wie viele Wörter (Types) in mehr als einem Syntagma auftreten und andererseits wie hoch die Gesamtfrequenz (Tokens) der jeweiligen Wörter (Types) ist, die in den Kombinationen von Syntagmen auftreten.

Tabelle 11 (S. 108) lässt sich dann entnehmen, dass von den 2250 Wörtern (Types) mit einer Gesamtfrequenz (Tokens) von 243.514, die im Korpus auftreten, 254 Wörter (Types) mit einer Gesamtfrequenz (Tokens) von 2821 ausschließlich in einem Syntagma, 712 Wörter (Types) mit einer Gesamtfrequenz (Tokens) von 20.711 in zwei Syntagmen, 642 Wörter (Types) mit einer Gesamtfrequenz (Tokens) von 62.794 in drei Syntagmen und 642 Wörter (Types) mit einer Gesamtfrequenz (Tokens) von 157.188 in allen Syntagmen auftreten.

Betrachtet man die Frequenzen der verschiedenen Kombinationen von Syntagmen genauer, werden deutliche Unterschiede erkennbar, besonders im Vergleich von flexiblen zu nicht-flexiblen Wörtern. So treten 11 % (254) aller im Korpus vorkommenden Wörter (Types) nicht-flexibel, das heißt ausschließlich in einem Syntagma auf. Demgegenüber treten 32% (712) aller im Korpus vorkommenden Wörter (Types) ausschließlich in zwei Syntagmen, 29% (642) ausschließlich in drei Syntagmen und ähnlich häufig 29% (642) ausschließlich in vier Syntagmen auf. Addiert man die Frequenzen derjenigen Wörter (Types), die ausschließlich in zwei Syntagmen (712), in drei Syntagmen (642) und in vier Syntagmen (642) zu finden sind, so treten knapp 89% (1996) der 2250 im Korpus vorkommenden Wörter (Types) flexibel, das heißt in mehr als einem Syntagma, auf.

Dieser Befund widerspricht der Annahme, dass in polynesischen Sprachen die Mehrheit der Wörter nicht flexibel auftreten, wie sie z.B. von Dixon (2010: 45) geäußert wurde (vgl. Kapitel 2.2.1, S. 13). Zumindest im untersuchten Korpus stellen Wörter, die ausschließlich in einem Syntagma auftreten, gegenüber Wörtern, die in mehr als einem Syntagma auftreten, eher den Sonderfall dar. Ein noch größerer Unterschied zwischen flexiblen und nicht-flexiblen Wörtern (Types) wird in der Gegenüberstellung der Gesamtfrequenzen (Tokens) deutlich. So umfassen die Wörter (Types), die ausschließlich in einem Syntagma auftreten, mit 2821 gerade etwas mehr als ein Prozent aller 243.514 im Korpus auftretenden Tokens. Addiert man die Gesamtfrequenzen (Tokens) der

Wörter (Types), die in zwei Syntagmen (29.711), in drei Syntagmen (62.794) und in vier Syntagmen (157.188) auftreten, so umfassen die flexiblen Wörter (Types) mit 240.693 Tokens knapp 99% aller im Korpus auftretenden Tokens. Beachtet man zudem, dass die Gesamtfrequenz (Tokens) der Wörter (Types) in einer Kombination aus Syntagmen mit der Anzahl der in der Kombination enthaltenen Syntagmen steigt, lässt das auf den ersten Blick einen Zusammenhang zwischen dem Flexibilitätsgrad eines Wortes und seiner Gesamtauftrittshäufigkeit vermuten: Je häufiger ein Wort im Korpus vorkommt, desto eher tritt es in mehr als einem Syntagma auf. Bevor dieser Zusammenhang noch genauer untersucht wird (vgl. Kapitel 4.2.2, S. 190), sollen zunächst folgende Ergebnisse festgehalten werden, die relevant für die Frage der Distributionsklassen im Hawaiischen sind: Zwar finden sich im Korpus Wörter, die ausschließlich in einem einzigen Syntagma auftreten, oder anders ausgedrückt: Nicht alle Wörter im Korpus treten flexibel auf. Dabei ist die „Nominalisierungs-konstruktion“ das einzige Syntagma, in dem kein Wort ausschließlich auftritt. Allerdings findet sich ebenso zu jeder der 11 möglichen Kombinationen aus mehreren Syntagmen mindestens ein Wort, das genau in dieser Kombination auftritt. Das bedeutet, dass im Korpus die Verteilung von Wörtern nicht auf ganz spezifische Kombinationen von Syntagmen festgelegt ist. Besonders auffällig ist hier, dass Wörter im Korpus ausschließlich in Artikel-Syntagma und „Nominalisierungs-konstruktion“ auftreten, auch wenn diese Wörter mit einer Frequenz von 6 im Vergleich zu Wörtern, die in anderen Kombinationen von Syntagmen auftreten, sehr wenige sind. Nur Wörter, die ausschließlich in attributivem Slot und „Nominalisierungs-konstruktion“ auftreten, kommen mit einer Frequenz von 4 noch seltener im Korpus vor. Das ist ein erster Hinweis darauf, dass die „Nominalisierungs-konstruktion“ an späterer Stelle noch genauer analysiert werden muss (vgl. Kapitel 4.1.3, S. 131–167).

Zunächst steht allerdings die Frage potentieller Distributionsklassen im Vordergrund. Aus rein kombinatorischen Überlegungen, das heißt unabhängig z.B. von der semantischen Analyse von Wörtern, sehe ich *ad hoc* drei Alternativen, Distributionsklassen im Korpus zu differenzieren: Erstens lässt sich jede der 14 spezifischen Kombinationen aus Syntagmen als eine Distributionsklasse ansetzen. Zweitens können die 14 spezifischen Kombinationen nach Flexibilitätsgrad, das heißt nach der Anzahl der darin enthaltenen Syntagmen, zu vier Distributionsklassen zusammengefasst werden, mit einer Distributionsklasse, zu der Wörter gehören, die ausschließlich in einem Syntagma auftreten, einer Distributionsklasse, zu der Wörter gehören, die ausschließlich in zwei

Syntagmen auftreten, einer Distributionsklasse, zu der Wörter gehören, die ausschließlich in drei Syntagmen auftreten und einer Distributionsklasse, zu der Wörter gehören, die ausschließlich in vier Syntagmen auftreten. Drittens ist es nach dem Kriterium des Flexibilitätsgrades ebenso möglich, nur zwei Distributionsklassen zu bilden, mit einer Distributionsklasse nicht-flexibler Wörter, die ausschließlich in einem Syntagma auftreten, und einer Distributionsklasse flexibler Wörter, die in mehr als einem Syntagma auftreten. Letztlich ist eine solche Einteilung von Wörtern in Distributionsklassen nach Flexibilitätsgrad jedoch willkürlich.

An dieser Stelle offenbart sich die Schwierigkeit, eindeutige Distributionsklassen nachzuweisen, wenn keine formalen Kriterien dafür existieren und Wörter nicht nur in spezifischen, sondern in allen möglichen Kombinationen aus Syntagmen auftreten, wie es im Korpus der Fall ist. Ohne die Bildung von Distributionsklassen können nun aber nicht ohne weiteres Wortartkategorien angesetzt werden.

Das ist insbesondere problematisch für die Übertragungsversuche von tradierten Wortartkategorien auf die polynesischen Sprachen, die in vielen einzelsprachlichen Grammatiken zu finden sind (vgl. Kapitel 2.2.1, S. 11–14).

Zur Veranschaulichung soll im Folgenden die Beschreibung von hawaiischen Wortartkategorien in der Grammatik von Elbert und Pukui (1979) betrachtet werden, um diese mit den Beobachtungen der Wortdistribution im Korpus zu vergleichen.

Elbert und Pukui (1979: 43 und passim) setzen im Hawaiischen die Wortartkategorien *Verb*, *Nomen* und *Nomen-Verb* („noun-verb“) an. Als *Verben* definieren Elbert und Pukui (1979: 43) Wörter, die ausschließlich in Tempus-Aspekt-Modus-Syntagma auftreten, als *Nomen* Wörter, die ausschließlich in Artikel-Syntagma auftreten und Personen, Orte oder Dinge denotieren und als *Nomen-Verben* schließlich „[...] verbs commonly used as nouns without the nominalizer ‘ana.“ (Elbert und Pukui 1979: 43). Als weitere Subklassen der *Verben* differenzieren Elbert und Pukui (1979: 46–56) *transitive*, *intransitive* und *stative Verben*, wobei vereinzelte Wörter auch zu mehr als einer dieser Subklassen gezählt werden müssen und ein als *Nomen-Verb* klassifiziertes Wort im Wörterbuch von Pukui und Elbert (1986) weiter z.B. als „stative noun-verb“ bezeichnet wird.

Eine eigenständige Kategorie *Adjektiv* oder *Adverb* setzen Elbert und Pukui (1979) nicht an, da *stative Verben* solche Wörter umfassen, die sowohl in Tempus-Aspekt-Modus-Syntagma als auch in attributiver Position auftreten.

ten (vgl. Elbert und Pukui 1979: 49).

In Abbildung 18 ist das von Elbert und Pukui (1979) für das Hawaiische ange-setzte Wortartsystem zusammengefasst, einschließlich der Angabe der im Korpus auftretenden Kombinationen von Syntagmen, die davon abgedeckt sind, vgl.:

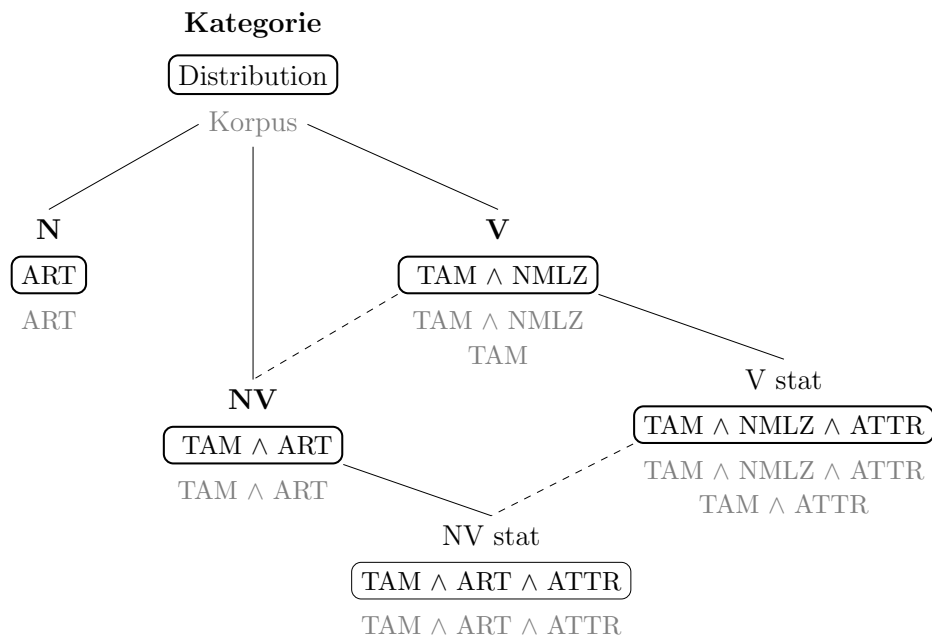


Abbildung 18: Wortartsystem von Elbert und Pukui (1979)

Unklar bleibt in diesem System zunächst, warum Elbert und Pukui (1979) auf der einen Seite drei Wortartkategorien *Verb*, *Nomen* und *Nomen-Verb* ansetzen, auf der anderen Seite *Nomen-Verben* aber als *Verben* bezeichnen, die als *Nomen* verwendet werden können. Genau genommen müssten Elbert und Pukui (1979) neben der Klasse der *Nomen* nur eine Klasse *Verben* ansetzen, die Wörter umfasst, die in Tempus-Aspekt-Modus-Syntagma auftreten und weiter differenziert werden kann in eine Subklasse von Wörtern, die in Tempus-Aspekt-Modus-Syntagma und potentiell in „Nominalisierungs-konstruktion“ auftreten und eine Subklasse von Wörtern, die sowohl in Tempus-Aspekt-Modus-Syntagma als auch Artikel-Syntagma, nicht aber „Nominalisierungs-konstruktion“ auftreten. Diese beiden *verbalen* Subklassen müssten jeweils noch in weitere Subklassen aufgeteilt werden, je nachdem, ob Wörter

neben Tempus-Aspekt-Modus-Syntagma und „Nominalisierungs-konstruktion“ oder Artikel-Syntagma zusätzlich in attributivem Slot auftreten. Alternativ ließen sich die beiden *verbalen* Subklassen auch einfach als zwei verschiedene Strategien der *Nominalisierung* von *Verben* betrachten.

Es fehlt dann aber eine Beschreibung des Unterschiedes zwischen *Verben* in „Nominalisierungs-konstruktion“, *Verben* in Artikel-Syntagma und *Nomen* in Artikel-Syntagma und es bleibt grundlegend eine Begründung von Elbert und Pukui (1979) aus, warum es sich bei *Nomen-Verben* um *Verben* handelt und nicht z.B. um *Nomen*, die als *Verben* auftreten können.

Elbert und Pukui (1979: 43) vermischen letztlich in der Definition von *Nomen-Verben* als *Verben* semantische und syntaktische Kriterien, weil *Verben*, zumindest implizit, semantisch klassifiziert werden (vgl. dazu auch Elbert und Pukui 1979: 46) und *Nomen-Verben* aber nach syntaktischen Kriterien als *Nomen* auftreten sollen.

Unabhängig von der Frage der semantischen Klassifizierung, tritt hier wieder das Problem einer rein syntaktischen Identifizierung von Wortarten auf (vgl. Kapitel 2.2.2, S. 14–16). So ist unklar, was es genau bedeuten soll, dass ein Wort als *Nomen* oder als *Verb* auftritt. Wenn ein *Nomen* als Wort definiert ist, das für das Auftreten in Artikel-Syntagma und ein *Verb* als Wort, das für das Auftreten in Tempus-Aspekt-Modus-Syntagma (und „Nominalisierungs-konstruktion“) prädisponiert ist, kann ein Wort, das für beide Syntagmen prädisponiert ist weder *Nomen* noch *Verb* sein. Zudem kann die „Nominalisierungs-konstruktion“ nicht Teil der Definition von *Verben* sein, da *Verben* und *Nomen* zunächst unabhängig von der „Nominalisierungs-konstruktion“ bestimmbar sein müssen.

Im Wörterbuch von Pukui und Elbert (1986) wird schließlich die Definition von *Nomen-Verben* als *Verben* aufgegeben und folgendermaßen umformuliert: „A base commonly used as both noun (without the nominalizer ‘*ana*’) and verb [...]“ (Pukui und Elbert 1986: ix). Freilich wird mit dieser Umformulierung von *Nomen-Verb* als „base“ statt *Verb* das Problem der syntaktischen Definition von Wortarten ebenso wenig gelöst.

Für einen Vergleich mit der Auftretensweise von Wörtern im Korpus lässt sich festhalten, dass Elbert und Pukui (1979: 43) aus rein distributionaler Perspektive folgende Klassen unterscheiden: *Nomen* als Wörter, die ausschließlich in Artikel-Syntagma auftreten, *Verben* als Wörter, die in Tempus-Aspekt-Modus-Syntagma und potentiell in „Nominalisierungs-konstruktion“ auftreten, *stative Verben* als *Verben*, die zusätzlich in attributivem Slot auftreten, *Nomen-*

Verben als Wörter, die sowohl in Tempus-Aspekt-Modus-Syntagma als auch Artikel-Syntagma, nicht aber in „Nominalisierungs-konstruktion“ auftreten und schließlich *stative Nomen-Verben* als *Nomen-Verben*, die auch in attributivem Slot auftreten.

Betrachtet man nun die Wortdistribution im Korpus (vgl. Tabelle 11, S. 108), kann gar nicht begründet werden, warum genau diese Auftretensweisen von Wörtern in Kombinationen aus Syntagmen Distributionsklassen als Wortart-kategorien bestimmen.

Zur besseren Übersicht sind in Abbildung 19 die möglichen Kombinationen aus Syntagmen in einem Venn-Diagramm dargestellt, vgl.:

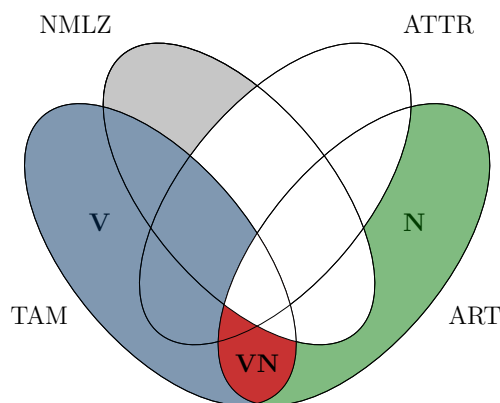


Abbildung 19: Venn-Diagramm Distributionsklassen

Das Venn-Diagramm in Abbildung 19 besteht aus vier Sets (Ellipsen). Ein Set (eine Ellipse) steht dabei für die Menge der in einem Syntagma auftretenden Wörter. Die Schnittmengen zeigen dann die Kombinationen aus Syntagmen, in denen Wörter auftreten.

Da im Korpus keine Wörter ausschließlich in „Nominalisierungs-konstruktion“ auftreten, ist der entsprechende Teil der Ellipse im Diagramm grau gefärbt.

Die Wörter, die Elbert und Pukui (1979) als *Verben* kategorisieren sind blau gefärbt, die Wörter, die Elbert und Pukui (1979) als *Nomen* kategorisieren

sind grün gefärbt und die *Nomen-Verben* sind rot gefärbt. Diejenigen Wörter, die zwar im Korpus auftreten, aber im System von Elbert und Pukui (1979) nicht erfasst werden, sind weiß gelassen.

Erstens wird in Abbildung 19 (S. 115) erneut die asymmetrische Kategorisierung von *Verb* und *Nomen* als Distributionsklassen bei Elbert und Pukui (1979) deutlich. Während *Verben* definiert sind als Wörter, die in Tempus-Aspekt-Modus-Syntagma und allen anderen Syntagmen außer Artikel-Syntagma oder der Kombination aus Artikel-Syntagma und einem weiteren Syntagma auftreten, sind *Nomen* definiert als Wörter, die ausschließlich in Artikel-Syntagma auftreten.

Zweitens zeigt Abbildung 19 (S. 115), dass von den 14 im Korpus auftretenden Distributionsklassen, i.e. (Kombinationen aus) Syntagmen im System von Elbert und Pukui (1979), zwar sieben erfasst werden. Allerdings können ebenso sieben Distributionsklassen keiner Wortart im System von Elbert und Pukui (1979) zugeordnet werden.

Elbert und Pukui (1979) können solche Wörter keiner *Wortart* zurechnen, die ausschließlich in attributivem Slot, sowohl in attributivem Slot als auch „Nominalisierungskonstruktion“, sowohl in Artikel-Syntagma als auch in attributivem Slot, sowohl in „Nominalisierungskonstruktion“ als auch Artikel-Syntagma, in attributivem Slot, „Nominalisierungskonstruktion“ und Artikel-Syntagma, in Tempus-Aspekt-Modus-Syntagma, „Nominalisierungskonstruktion“ und Artikel-Syntagma und schließlich in allen vier Syntagmen auftreten. Betrachtet man dazu die Frequenzen der Wörter, die in den jeweiligen Syntagmen auftreten (vgl. Tabelle 11, S. 108), zeigt sich, dass die Mehrheit der im Korpus auftretenden Wörter durch das System von Elbert und Pukui (1979) nicht erfasst werden (vgl. Tabelle 12, S. 117).

Tabelle 12 (S. 117) lässt sich entnehmen, dass von den insgesamt 2250 Wörtern (Types), die im Korpus auftreten, im System von Elbert und Pukui (1979) nur 962 (43%) einer *Wortart* zugeordnet werden können, demgegenüber 1288 (57%) der im Korpus auftretenden Wörter jedoch nicht. Bei den Tokenanzahlen ist der Unterschied noch deutlicher: Während 62.219 (26%) der 243.514 im Korpus auftretenden Tokens vom System von Elbert und Pukui (1979) erfasst werden, bleibt bei 181.295 (74%) Tokens die *Wortart* unbestimmt.

Bis hierhin lässt sich zusammenfassend festhalten, dass aus der im Korpus beobachteten Verteilung von Wörtern auf (Kombinationen von) Syntagmen nicht ohne weiteres Distributionsklassen im Hawaiischen bestimmbar sind, da Wörter in allen (Kombinationen von) Syntagmen auftreten und letztlich ein-

Syntagmen	absolute Frequenz		relative Frequenz	
	Types	Tokens	Types	Tokens
erfasst				
TAM	54	448	0,02	0,00
ART	115	1415	0,05	0,00
TAM \wedge ART	109	979	0,05	0,00
TAM \wedge ATTR	100	914	0,05	0,00
TAM \wedge NMLZ	40	374	0,02	0,00
TAM \wedge ART \wedge ATTR	492	57402	0,22	0,24
TAM \wedge ATTR \wedge NMLZ	52	687	0,02	0,00
Σ	962	62219	0,43	0,26
nicht erfasst				
ATTR	85	958	0,04	0,00
ART \wedge ATTR	453	18317	0,20	0,08
ART \wedge NMLZ	6	104	0,00	0,00
ATTR \wedge NMLZ	4	23	0,00	0,00
TAM \wedge ART \wedge NMLZ	67	1504	0,03	0,01
ART \wedge ATTR \wedge NMLZ	31	3201	0,01	0,01
TAM \wedge ART \wedge ATTR \wedge NMLZ	642	157188	0,29	0,65
Σ	1288	181295	0,57	0,74
Σ	2250	243514	1	1

Tabelle 12: Vergleich erfasste und nicht erfasste Wörter im Korpus

deutige Kriterien für die Bildung von Distributionsklassen fehlen. Auch das von Elbert und Pukui (1979) vorgeschlagene Wortartsystem hilft hier nicht weiter. Denn erstens ist anhand der Wortdistribution nicht ersichtlich, warum Wörter, die in Tempus-Aspekt-Modus-Syntagma und einem beliebigen anderen Syntagma auftreten, eine Klasse *Verben* bilden sollen, und in Gegenüberstellung dazu Wörter, die ausschließlich in Artikel-Syntagma auftreten, eine Klasse *Nomen*, insbesondere nicht, wenn man berücksichtigt, dass Wörter sowohl in Tempus-Aspekt-Modus-Syntagma als auch Artikel-Syntagma auftreten. Zweitens ist die Hälfte der möglichen Kombinationen aus Syntagmen, in denen Wörter im Korpus auftreten und sogar die Mehrheit der Worttypes und Worttokens im Korpus keiner *Wortart* zuweisbar.

Ein Wortartsystem wie das von Elbert und Pukui (1979) muss dann mindestens revidiert, wenn nicht gar völlig aufgegeben werden. Betrachtet man nur die Wortdistribution im Korpus, bleibt zumindest unklar, welche Kombinationen aus Syntagmen und warum genau diese Kombinationen Distributionsklassen darstellen.

Ein wichtiger Aspekt, der die Repräsentativität des Korpus betrifft (vgl. Kapitel 3.3.2, S. 67–69), muss hier allerdings noch beachtet werden: Es ist *a*

priori nicht auszuschließen, das ein anders zusammengesetztes Korpus, auch andere Ergebnisse liefert. Das bedeutet, dass möglicherweise Wörter, die im Korpus z.B. ausschließlich in Artikel-Syntagma oder ausschließlich in Artikel-Syntagma und attributivem Slot auftreten, im Hawaiischen auch in weiteren Syntagmen auftreten können. Es ist zumindest vorstellbar, dass im Korpus mehr potentielle Distributionsklassen beobachtet wurden, als sich durch andere Studien bestätigen ließen. Die Ergebnisse aus der Untersuchung der Wortverteilung auf bestimmte (Kombinationen von) Syntagmen im Korpus sollten also, wenn möglich, durch Ergebnisse aus anderen Arbeiten evaluiert werden. Da bisher keine Arbeiten zur Verteilung von Wörtern auf Syntagmen im Hawaiischen vorliegen, bietet sich als Möglichkeit der Evaluation an, einen Blick in hawaiische Wörterbücher zu werfen und zur genaueren Erfassung potentieller Distributionsklassen im Hawaiischen die Beobachtungen zur Verteilung von Wörtern im Korpus mit den Einträgen dieser Wörter bzw. der Klassifizierung dieser Wörter in den Wörterbüchern abzugleichen.

Die Ergebnisse aus dem Abgleich von Korpusbeobachtungen und Wörterbucheinträgen und die daraus folgenden Überlegungen zu Distributionsklassen und *Wortarten* im Hawaiischen sind Gegenstand des nächsten Kapitels.

4.1.2 Abgleich Korpusbeobachtungen und Wörterbucheinträge

Für eine verbesserte Evaluation der Ergebnisse aus der Korpusuntersuchung und um ein möglichst genaues Profil von potentiellen Distributionsklassen im Hawaiischen zu erhalten, wurde sowohl das Wörterbuch von Pukui und Elbert (1986) als auch das Wörterbuch von Parker (1922) herangezogen.

Die Erstauflage des Wörterbuchs von Pukui und Elbert (1986) stammt von 1957 und nach mehrfacher Revidierung und Erweiterung stellt die sechste Auflage von 1986 mit ungefähr 29.000 Einträgen (im Abschnitt Hawaiisch-Englisch) vermutlich das umfangreichste Wörterbuch einer polynesischen Sprache dar (vgl. Pukui und Elbert 1986: vii).

Das Wörterbuch von Parker (1922) ist eine revidierte und erweiterte Fassung des bereits 1865 erschienenen, ersten größeren Wörterbuchs des Hawaiischen von Lorrin Andrews und enthält mehr als 15.000 Einträge (vgl. Parker 1922: vi).

Während in der aus den Korpusdaten erstellten Liste nur festgehalten ist, wie häufig Inhaltswörter (Types) in welchen Syntagmen auftreten (vgl. Kapitel 3.5, S. 104–106), setzen sowohl Pukui und Elbert (1986) als auch Parker

(1922) im Hawaiischen Wortartkategorien an und versehen ein Wort im jeweiligen Wörterbucheintrag entsprechend mit einer Wortartetikette, an der aber auch ablesbar ist, in welchen Syntagmen das Wort potentiell auftreten soll. In Tabelle 13 sind die Kategorienetiketten von Elbert und Pukui (1986) und Parker (1922) den im Korpus erfassten Kategorien gegenübergestellt, vgl.:

Korpus Wort (Type) in	Elbert und Pukui (1986) word-class	Parker (1922) word-class
Artikel-Syntagma	noun noun-verb transitive noun-verb intransitive noun-verb stative noun-verb	noun
Tempus-Aspekt-Modus-Syntagma	verb noun-verb transitive verb transitive noun-verb intransitive verb intransitive noun-verb stative verb stative noun-verb	verb
attributer Slot	stative verb stative noun-verb	adjective adverb

Tabelle 13: Angleichung der im Korpus und in Wörterbüchern verwendeten Kategorien

Eine Einschränkung des Vergleichs ist hier dadurch gegeben, dass weder im Wörterbuch von Pukui und Elbert (1986) noch im Wörterbuch von Parker (1922) Angaben zum Auftreten eines Wortes in „Nominalisierungskonstruktion“ gemacht werden.

Der Einteilung von Wörtern in *Wortarten* im Wörterbuch von Pukui und Elbert (1986) liegt das von Elbert und Pukui (1979) entwickelte Wortartensystem zugrunde (vgl. Kapitel 4.1.1, S. 112–114). Im Wörterbuch von Pukui und Elbert (1986) wird dann z.B. *pōpoki* „cat“ als „noun“ (Pukui und Elbert 1986: 342), *hu'e* „to remove“ als „transitive verb“ (Pukui und Elbert 1986: 85), *wiki* „to hurry, quick“ als „stative verb“ (Pukui und Elbert 1986: 385) und *hana*

„work, to work“ als „transitive noun-verb“ (Pukui und Elbert 1986: 55) angeben.

Vereinzelt finden sich in den Einträgen in Pukui und Elbert (1986) Angaben ohne weitere Spezifizierung der Subklasse eines Wortes, z.B. *‘aulama* „to light with a torch“ nur als „verb“ (Pukui und Elbert 1986: 32) oder *anapuni* „boundary, to encircle“ nur als „noun-verb“ (Pukui und Elbert 1986: 24). Daher ist in Tabelle 13 (S. 119) in der zweiten Spalte unter den Kategorien von Pukui und Elbert (1986) auch „verb“ und „noun-verb“ separat angeführt.

Im Gegensatz zu Pukui und Elbert (1986) übernimmt Parker (1922: xviii) die auf der Grammatik von Andrews (1854) aufbauende Wortartunterteilung aus dem Wörterbuch von Andrews (1865) in „verbs“, das heißt Wörter, die in Tempus-Aspekt-Modus-Syntagma auftreten, „nouns“, das heißt Wörter, die in Artikel-Syntagma auftreten, „adjectives“, das heißt Wörter, die attributiv in Artikel-Syntagma auftreten und „adverbs“, das heißt Wörter, die attributiv in Tempus-Aspekt-Modus-Syntagma auftreten. Wörter, die in mehr als einem Syntagma auftreten, sind in Parker (1922) mit einem eigenen Eintrag für das Auftreten im jeweiligen Syntagma mit entsprechender Wortartetikette versehen. So werden dann die zuvor betrachteten Beispiele aus Pukui und Elbert (1986) folgendermaßen in Parker (1922) angegeben:³³ *popoki* „cat“ als „noun“ (Parker 1922: 566), *hue* „to remove“ als „verb“ (Parker 1922: 208), *wiki* einmal mit „quick“ als „adjective“ (Parker 1922: 622) und einmal mit „to hasten“ als „verb“ (Parker 1922: 622) und schließlich *hana* einmal mit „work“ als „noun“ (Parker 1922: 103) und einmal mit „to work“ als „verb“ (Parker 1922: 104). Nach der Angleichung der von Pukui und Elbert (1986) und Parker (1922) verwendeten Termini konnte überprüft werden, ob die Verteilung von Wörtern auf bestimmte Syntagmen im Korpus den Angaben im Wörterbuch von Pukui und Elbert (1986) und den Angaben im Wörterbuch von Parker (1922) entspricht.

In einem ersten Schritt sollte nur bei denjenigen Wörtern, die im Korpus ausschließlich in einem Syntagma auftreten, überprüft werden, ob diese auch in den Wörterbüchern über die verwendeten Termini als solche ausgezeichnet sind, um sich zunächst einen besseren Überblick darüber zu verschaffen, wie viele Wörter im Hawaiischen potentiell nur in einem Syntagma auftreten.

³³In Parker (1922) sind langer Vokal und Glottal graphemisch nicht realisiert (vgl. dazu Kapitel 3.3.2, S. 73–74). Zwar gibt Parker (1922) nach jedem Wort mithilfe der Markierung durch Diakritika die Aussprache des Wortes an. Diese Angaben sind aber nicht identisch mit der Schreibweise der neuen Orthographie, wie sie z.B. Pukui und Elbert (1986) verwenden.

Tabelle 14 zeigt einen Ausschnitt aus der Vergleichstabelle zwischen den Ergebnissen aus der Korpusuntersuchung und den Angaben von Pukui und Elbert (1986) und von Parker (1922) mit Wörtern, an denen unter anderem gefundene Unterschiede exemplarisch herausgestellt sind, vgl.:

Wort	Korpus			Pukui und Elbert (1986)			Parker (1922)		
	TAM	ART	ATTR	TAM	ART	ATTR	TAM	ART	ATTR
kākala	0	1	0	1	1	1	1	1	1
kēpā	0	1	0	0	1	0	1	1	0
ko‘eko‘e	0	1	0	1	0	1	1	1	1
limahana	0	1	0	1	1	1	-	-	-
makewai	0	1	0	-	-	-	1	1	1
pehi	1	0	0	1	1	1	1	0	0
pōpoki	0	1	0	0	1	0	0	1	0

Tabelle 14: Ausschnitt Vergleichstabelle zum Wortauftreten

Die Angaben in Tabelle 14 sind binär kodiert, das heißt 1 kodiert „kommt vor“ und 0 kodiert „kommt nicht vor“. Beispielsweise tritt *pōpoki* im Korpus ausschließlich in Artikel-Syntagma auf und ist auch von Pukui und Elbert (1986: 342–343) und von Parker (1922: 566) mit dem entsprechenden Wortartetikett versehen. Neben Übereinstimmungen treten allerdings auch Diskrepanzen zwischen den Ergebnissen der Korpusuntersuchung und den Angaben von Pukui und Elbert (1986) und von Parker (1922) auf. So tritt z.B. *kākala* im Korpus ausschließlich in Artikel-Syntagma auf, sollte aber den Angaben von Pukui und Elbert (1986: 118) und von Parker (1922: 247) zufolge ebenso in Tempus-Aspekt-Modus-Syntagma und attributiven Slot auftreten können.

Es finden sich auch Diskrepanzen zwischen den Angaben von Pukui und Elbert (1986) und von Parker (1922). Beispielsweise sollte Parker (1922) zufolge *kēpā* sowohl in Artikel-Syntagma als auch Tempus-Aspekt-Modus-Syntagma auftreten können, Pukui und Elbert (1989) zufolge hingegen so wie im Korpus ausschließlich in Artikel-Syntagma.

In wenigen Fällen weichen die Angaben von Pukui und Elbert (1986) oder von Parker (1922) sehr stark von den Beobachtungen um Korpus ab, z.B. bei *ko‘eko‘e*, das im Korpus ausschließlich in Artikel-Syntagma auftritt, Pukui und Elbert (1986) zufolge aber gar nicht in Artikel-Syntagma, sondern ausschließlich in Tempus-Aspekt-Modus-Syntagma und attributivem Slot auftreten sollte.

Nicht bei jedem im Korpus nicht-flexibel auftretenden Wort war ein Vergleich mit den Angaben von Pukui und Elbert (1986) und von Parker (1922) möglich,

da es schlicht keinen Eintrag zu dem entsprechenden Wort im Wörterbuch von Pukui und Elbert (1986), z.B. zu *makewai*, oder im Wörterbuch von Parker (1922), z.B. zu *limahana*, gibt. Zu insgesamt 10 Wörtern, die im Korpus ausschließlich in einem Syntagma auftreten, sind weder von Pukui und Elbert (1986) noch von Parker (1922) Angaben zu finden. Diejenigen Wörter, zu denen in keinem der beiden Wörterbücher ein Eintrag existiert, sind dann aus dem Vergleich gestrichen worden.

In Tabelle 15 sind die Ergebnisse aus dem Vergleich zwischen den Beobachtungen im Korpus mit den Einträgen in dem Wörterbuch von Pukui und Elbert (1986) und dem Wörterbuch von Parker (1922) zusammengefasst, vgl.:

Korpus	flexibel	nicht-flexibel	anders	Σ	n/a	Σ
Pukui und Elbert (1989)						
TAM	21	0 (19)	0	40	14	54
ART	32	64	3	99	16	115
ATTR	55	0 (22)	24	79	6	85
Σ	108	83	27	218	36	254
Parker (1922)						
TAM	13	25	1	39	15	54
ART	32	51	3	86	29	115
ATTR	36	10	21	67	18	85
Σ	81	86	25	192	62	254
Ergebnis aus						
	$PE \vee Pa$	$PE \wedge Pa$			$PE \wedge Pa$	
TAM	28	0 (21)	-	49	5	54
ART	42	69	-	111	4	115
ATTR	84	0 (16)	-	84	1	85
Σ	154	90	-	244	10	254

Tabelle 15: Zusammenfassung Vergleich Korpus und Wörterbücher

In Tabelle 15 ist in der ersten Spalte das Syntagma angegeben, in dem Wörter im Korpus ausschließlich auftreten. Im ersten Teil von Tabelle 15 ist dann angegeben, wie viele Wörter im Wörterbuch von Pukui und Elbert (1986) als flexibel (Spalte 2), das heißt in mehr als einem Syntagma auftretend, als nicht-flexibel (Spalte 3), das heißt so wie im Korpus ausschließlich in dem entsprechenden Syntagma auftretend, oder anders (Spalte 4), das heißt in einem anderen Syntagma auftretend, angegeben sind. In Spalte 5 ist die Summe aus der Anzahl der Wörter angegeben, zu denen Einträge in Pukui und Elbert (1986) existieren. In Spalte 6 ist die Anzahl der Wörter festgehalten, zu denen es kei-

ne Angaben in Pukui und Elbert (1986) gibt und in Spalte 7 schließlich die Gesamtanzahl der Wörter, die im Korpus auftreten.

Problematisch ist der Vergleich zum Auftreten in attributivem Slot insofern, dass Pukui und Elbert (1986) und Elbert und Pukui (1979) Wörter, die attributiv auftreten als „stative verbs“ klassifizieren und demzufolge keine Wörter existieren können, die ausschließlich attributiv auftreten. In Klammern angegeben ist dann die Anzahl der Wörter, die Pukui und Elbert (1986) als „stative verb“ klassifiziert haben und damit nur in Tempus-Aspekt-Modus-Syntagma und in attributivem Slot auftreten sollten. Alle anderen wurden von Pukui und Elbert (1986) anders klassifiziert, z.B. als „intransitive noun-verb“. Ein ähnliches Problem besteht beim Vergleich der Wörter, die im Korpus ausschließlich in Tempus-Aspekt-Modus-Syntagma auftreten. Denn Elbert und Pukui (1979) gehen in der Klassifizierung von *Verben* davon aus, dass jedes Wort, das in Tempus-Aspekt-Modus-Syntagma auftritt auch in „Nominalisierungskonstruktion“ auftreten kann und somit kein Wort potentiell ausschließlich in Tempus-Aspekt-Modus auftritt. In Tabelle 15 (S. 122) ist dann in Klammern die Anzahl der Wörter angegeben, die Pukui und Elbert (1986) als „verb“ und nicht z.B. als „noun-verb“ klassifiziert haben.

Im zweiten Teil von Tabelle 15 (S. 122) sind die gleichen Angaben zu Parker (1922) angeführt.

Da Parker (1922) „adjectives“ als eigenständige Wortart ansetzt und die „Nominalisierungskonstruktion“ nicht berücksichtigt, können hier die Wörter angegeben werden, die Parker (1922) zufolge ausschließlich in attributivem Slot oder ausschließlich in Tempus-Aspekt-Modus-Syntagma auftreten sollten.

Aus Tabelle 15 (S. 122) lässt sich nun beispielsweise entnehmen, dass von 40 Wörtern, die im Korpus ausschließlich in Tempus-Aspekt-Modus-Syntagma auftreten, Pukui und Elbert (1986) zufolge 19 auch in anderen Syntagmen auftreten können und von 39 Wörtern, die im Korpus ausschließlich in Tempus-Aspekt-Modus-Syntagma auftreten, Parker (1922) zufolge sogar 26 Wörter in einem anderen Syntagma auftreten können, oder dass von 99 Wörtern, die im Korpus ausschließlich in Artikel-Syntagma auftreten, Pukui und Elbert (1986) zufolge nur 32 Wörter ausschließlich in Artikel-Syntagma auftreten können und von 89 Wörtern, die im Korpus ausschließlich in Artikel-Syntagma auftreten, Parker (1922) zufolge ebenfalls nur 32 Wörter ausschließlich in Artikel-Syntagma auftreten können.

Insgesamt sind von 218 im Korpus ausschließlich in einem Syntagma auftretenden Wörter nur 108 auch in Pukui und Elbert (1986) als nicht-flexibel und

von 192 im Korpus ausschließlich in einem Syntagma auftretenden Wörter nur 81 auch von Parker (1922) als nicht-flexibel kategorisiert.

Im dritten Abschnitt von Tabelle 15 (S. 122) ist schließlich das Ergebnis aus der Verknüpfung der Angaben von Pukui und Elbert (1986) und Parker (1922) zusammengestellt.

Als „flexibel“ wurden diejenigen Wörter zusammengefasst, die entweder in Pukui und Elbert (1986) oder Parker (1922) oder sowohl in Pukui und Elbert (1986) als auch Parker (1922), also in mindestens einem der beiden Wörterbücher als in mehr als einem Syntagma auftretend angegeben sind. In Tabelle 15 (S. 122) sind diese als nicht-ausschließende disjunkte Verknüpfung der Angaben von Pukui und Elbert (1986) und Parker (1922) notiert. Die Kategorie „anders“ fällt hier weg, da kein Wort sowohl in Pukui und Elbert (1986) als auch in Parker (1922) zwar als ausschließlich in einem Syntagma, aber in einem anderen Syntagma als im Korpus auftretend, klassifiziert ist.

Als „nicht-flexibel“ wurden im Gegensatz dazu nur diejenigen Wörter zusammengefasst, die sowohl in Pukui und Elbert (1986) als auch in Parker (1922) als in ausschließlich einem Syntagma auftretend angegeben sind. Diese sind in Tabelle 15 (S. 122) als konjunkte Verknüpfung der Angaben von Pukui und Elbert (1986) und Parker (1922) notiert.

Zusätzlich angeführt sind auch noch diejenigen Wörter, zu denen weder in Pukui und Elbert (1986) noch in Parker (1922) Einträge zu finden sind (n/a).

Im Ergebnis bleiben von 49 Wörtern, die im Korpus ausschließlich in Tempus-Aspekt-Modus-Syntagma auftreten, nur 21 Wörter, von 111 Wörtern, die im Korpus ausschließlich in Artikel-Syntagma auftreten, nur 69 Wörter, von 84 Wörtern, die im Korpus ausschließlich attributiv auftreten, gar keine Wörter, und insgesamt von 244 Wörtern, die im Korpus ausschließlich in einem Syntagma auftreten, nur 90 Wörter übrig, die einem Vergleich mit den Wörterbüchern standhalten.

Aus dem Vergleich der im Korpus nicht-flexibel auftretenden Wörter mit den Angaben zu diesen in den Wörterbüchern von Pukui und Elbert (1986) und Parker (1922) geht dann folgendes hervor: Erstens kann angenommen werden, dass im Hawaiischen weniger Wörter nicht-flexibel auftreten als die Auftretensweise im Korpus vermuten lässt. Zweitens kann aber bestätigt werden, dass wenige Wörter im Hawaiischen ausschließlich in einem Syntagma auftreten, zumindest ausschließlich in Artikel-Syntagma, folgt man der Annahme, dass alle Wörter, die in attributivem Slot auftreten, auch in Tempus-Aspekt-Modus-Syntagma und alle Wörter, die in Tempus-Aspekt-Modus-Syntagma auftreten

auch in „Nominalisierungs-konstruktion“ auftreten können.

Vergleicht man auch die Wörter, die im Korpus in mehr als einem Syntagma auftreten, mit den Angaben von Pukui und Elbert (1986) und von Parker (1922), ergibt sich ein ähnliches Bild, vgl.:

Syntagmen	Korpus	PE \wedge Pa	
		=	+
TAM-ART	109	71	38
TAM-ATTR	100	0	100 (42)
TAM-NMLZ	40	35	5
ART-ATTR	453	236 ³⁴	217
ART-NMLZ	6	2	4
ATTR-NMLZ	4	0	4
Σ	712	344	368
TAM-ART-ATTR	492	492	n/a
TAM-ART-NMLZ	67	51	16
TAM-ATTR-NMLZ	52	36	16
ART-ATTR-NMLZ	31	11	20
Σ	642	579	63
Σ	1354	923	431

Tabelle 16: Korpus-Wörterbuch-Vergleich flexibler Wörter

In der zweiten Spalte von Tabelle 16 ist angegeben, wie viele Wörter im Korpus ausschließlich in der entsprechenden Kombination aus Syntagmen auftreten, in der dritten Spalte, wie viele davon den Einträgen in den Wörterbüchern von Pukui und Elbert (1986) und Parker (1922) entsprechen (=) und in der vierten Spalte, wie viele in einem weiteren Syntagma auftreten sollten (+).

Beispielsweise treten 109 Wörter im Korpus ausschließlich in Tempus-Aspekt-Modus-Syntagma und Artikel-Syntagma auf. Den Angaben von Pukui und Elbert (1986) und Parker (1922) zufolge können allerdings 38 davon auch in attributivem Slot auftreten, sodass nach dem Vergleich 71 Wörter übrig bleiben, die potentiell ausschließlich in in Tempus-Aspekt-Modus-Syntagma und Artikel-Syntagma auftreten.

Wie bei den nicht-flexiblen Wörtern ist hier wieder die Behandlung der „Nominalisierungs-konstruktion“ schwierig. So müssten – zumindest Elbert und

³⁴Zu 18 der 453 Wörter im Korpus finden sich keine Angaben im Wörterbuch von Pukui und Elbert (1986) oder Parker (1922). Diese sind hier trotzdem zu den 236 gezählt worden.

Pukui (1979) zufolge – alle Wörter, die im Korpus ausschließlich in Tempus-Aspekt-Modus-Syntagma und attributivem Slot auftreten, theoretisch auch in „Nominalisierungs-konstruktion“ auftreten können. In Tabelle 16 (S. 125) ist dann in Klammern die Anzahl der Wörter angegeben, die Pukui und Elbert (1986) und Parker (1922) zufolge auch in Artikel-Syntagma auftreten sollten. Da weder in Pukui und Elbert (1986) noch in Parker (1922) Angaben zum Auftreten in „Nominalisierungs-konstruktion“ enthalten sind, kann die Kombination aus Tempus-Aspekt-Modus-Syntagma, Artikel-Syntagma und attributivem Slot nicht verglichen werden. In den anderen Kombinationen wurde überprüft, ob Pukui und Elbert (1986) und Parker (1922) zufolge die Wörter auch in einem anderen Syntagma vorkommen. So treten z.B. 6 Wörter im Korpus ausschließlich in Artikel-Syntagma und „Nominalisierungs-konstruktion“ auf, den Wörterbüchern von Pukui und Elbert (1989) und Parker (1922) zufolge, sollten jedoch 4 davon auch in Tempus-Aspekt-Modus-Syntagma auftreten können und nur 2 ausschließlich in Artikel-Syntagma.

Anzumerken sei hier noch, dass Wörter, die im Korpus in mehr Syntagmen auftreten als den Angaben im Wörterbuch von Pukui und Elbert und von Parker (1922) entsprechen, in Tabelle 16 (S. 125) nicht in besonderer Weise markiert worden sind, weil das Auftreten im Korpus als Beleg für das potentielle Auftreten eines Wortes angesehen wurde. Ein Beispiel für ein solches Wort ist *naue*, das im Korpus in allen vier syntaktischen Slots auftritt, von Pukui und Elbert (1986: 263) allerdings als „intransitive verb“ und von Parker (1922: 456) als „verb“ klassifiziert ist.

Ähnlich wie bei den nicht-flexiblen Wörtern (Tabelle 15, S. 122) geht dann aus Tabelle 16 (S. 125) hervor, dass erstens im Hawaiischen viele Wörter flexibler als im Korpus, also anders als im Korpus in mehr als zwei oder drei Syntagmen, auftreten können. Zweitens bestätigt sich aber auch, dass manche Wörter vermutlich ausschließlich in zwei oder drei Syntagmen auftreten.

Insgesamt kann als Ergebnis aus dem Vergleich der Beobachtungen zur Wortverteilung auf (Kombinationen von) Syntagmen im Korpus mit den Angaben in den Wörterbüchern von Pukui und Elbert (1986) und von Parker (1922) festgehalten werden, dass im Korpus Wörter als weniger flexibel repräsentiert sind, als sie es den Angaben in den Wörterbüchern zufolge sein sollten. Es muss zwar nicht davon ausgegangen werden, dass Wörter in einem Korpus genauso flexibel auftreten, wie sie es potentiell können und auch Unterschiede zwischen Wörterbüchern sind nicht unbedingt überraschend. Beachtet man aber zudem, dass auch Wörter im Korpus flexibler auftreten als aus den Angaben in den

Wörterbüchern von Pukui und Elbert (1986) und von Parker (1922) hervor- geht, lässt dies erste Zweifel an der Annahme aufkommen, Wörter seien im Hawaiischen für das Auftreten in ganz bestimmten Syntagmen prädisponiert. Zur Frage von Distributionsklassen lässt sich weiter festhalten, dass manche, aber nicht alle im Korpus beobachteten Verteilungen von Wörtern auf bestimmte (Kombinationen von) Syntagmen durch den Vergleich mit den Wörterbüchern bestätigt werden.

Die Abweichungen der Angaben in den Wörterbüchern und der Grammatik von Elbert und Pukui (1979) von der beobachteten Wortverteilung im Korpus betreffen das potentielle Auftreten von Wörtern in attributivem Slot und in „Nominalisierungs-konstruktion“. Erstens sollte dann Elbert und Pukui (1979) zufolge kein Wort potentiell ausschließlich in Tempus-Aspekt-Modus-Syntagma auftreten, sondern als *Verb* mindestens noch in „Nominalisierungs-konstruktion“. Zweitens sollte kein Wort potentiell ausschließlich in attributivem Slot auftreten, sondern als *statives Verb* ebenfalls in Tempus-Aspekt-Modus-Syntagma. Das würde bedeuten, dass von den 14 im Korpus beobachteten Wortverteilungen auf bestimmte Syntagmen (vgl. Abbildung 11, S. 108) das Auftreten ausschließlich in attributivem Slot sowie ausschließlich in attributivem Slot und „Nominalisierungs-konstruktion“ nicht mehr als Distributionsklasse in Frage kommt.

Bei vier der verbleibenden Kombinationen aus Syntagmen ist jedoch immer noch ungeklärt, welche Kategorie diese Elbert und Pukui (1979) zufolge definieren sollen. So kann beispielsweise ein Wort, das im Korpus ausschließlich in attributivem Slot und Artikel-Syntagma oder ein Wort, das in Tempus-Aspekt-Modus-Syntagma, Artikel-Syntagma und „Nominalisierungs-konstruktion“ auftritt, keiner *Wortart* zugewiesen werden.

Bezogen auf die Auftretensweise von Wörtern in attributivem Slot ist nun auffällig, dass von den 85 Wörtern, die im Korpus ausschließlich in attributivem Slot auftreten, nur 22 von Pukui und Elbert (1986) als *stative Verben* klassifiziert sind (vgl. Tabelle 15, S. 122) und von den 100 Wörtern, die im Korpus ausschließlich in Tempus-Aspekt-Modus-Syntagma und attributivem Slot auftreten, nur 42 von Pukui und Elbert (1986) als *stative Verben* klassifiziert sind (vgl. Tabelle 16, S. 125). Im Korpus treten also Wörter ausschließlich in Tempus-Aspekt-Modus-Syntagma und attributivem Slot auf, die nicht von Elbert und Pukui (1979) und Pukui und Elbert (1986) als *stative Verben* oder von Parker (1922) als *Adjektive* klassifiziert werden, vgl.:

- (8) *ua ho‘oili ‘ia iho-la [kekahi kaua ho‘omeamea]*
 PFV losgehen PASS DIR-DIST ART Kampf vortäuschen
 ‘Ein vorgetäuschter Kampf ging los.’
 (Desha 1995 [1920,1924]:1)

In Satz (8) tritt das von Pukui und Elbert (1986: 244) als *transitives Verb* klassifizierte Wort *ho‘omeamea* „vortäuschen“, bestehend aus dem Präfix *ho‘o-* und der Reduplikation der Wurzel *mea* in der Bedeutung „sagen“ in attributivem Slot auf.

Es kann dann nicht mehr jedes Wort, das in attributivem Slot auftritt, als *statives Verb* klassifiziert werden und die semantische Subklassifizierung in *stative*, *intransitive* und *transitive Verben* ist nicht gleichzusetzen mit dem Auftreten von Wörtern in bestimmten Syntagmen, da offensichtlich unabhängig von dieser Wörter, die in Tempus-Aspekt-Modus-Syntagma auftreten, potentiell auch in attributivem Slot auftreten können.

Hinzu kommt, dass im Korpus manche Wörter ausschließlich in attributivem Slot und Artikel-Syntagma auftreten (vgl. Abbildung 19, S. 115) und Wörter, die von Pukui und Elbert (1986) und Parker (1922) als *Nomen* klassifiziert sind, auch in attributivem Slot auftreten, vgl.:

- (9) *ua hua‘i ‘ia [nā iwi kupapa‘u]*³⁵
 PFV ausgraben PASS PL Knochen Leichnam
 ‘Leichenknochen wurden ausgegraben.’
 (Silva 2002b:306)

In Satz (9) tritt das von Pukui und Elbert (1986: 185) als *Nomen* klassifizierte Wort *kupapa‘u* „Leichnam“ in attributivem Slot auf.

Im Wörterbuch von Parker (1922) finden sich zwei Einträge: einmal *kupapa‘u* „a dead body, corpse“, klassifiziert als *Nomen* (vgl. Parker 1922: 343) und einmal *kupapa‘u* „of or belonging to a dead body“, klassifiziert als *Adjektiv*, und dazu das Beispiel *hale kupapa‘u* „tomb“ (vgl. Parker 1922: 343), mit *hale*

³⁵Eine Bedingung für die Zitation von Texten in neuer Orthographie aus den im Rahmen von *Papahāna Ho‘oilina ‘Ōlelo Hawai‘i, The Hawaiian Language Legacy Program* entstandenen drei Ausgaben des *Legacy Journal* (vgl. Kapitel 1.3.2, S. 15) ist, zusätzlich die Orthographie des Originaltextes anzugeben (vgl. Silva 2002a: iv - v). Daher wird hier und in nachfolgenden Beispielen der Originaltext in einer Fussnote angeführt.

Originaltext: *ua huai ia na iwi kupapau.*

„Haus“. Dieses Beispiel findet sich auch im Wörterbuch von Pukui und Elbert (1986: 547) als Eintrag unter „tomb“ im Teil Englisch-Hawaiisch. Daran wird deutlich, dass Pukui und Elbert (1986: 185) *kupapa‘u* letztlich semantisch als *Nomen* klassifizieren (vgl. auch Kapitel 4.1.1, S. 112).

Wichtig ist an dieser Stelle, dass im Hawaiischen vermutlich jedes Wort in attributivem Slot auftreten kann, unabhängig von der semantischen Klassifizierung und unabhängig davon, in welchem Syntagma es ebenfalls auftritt. Anders als z.B. im Deutschen kann hier auch nicht zwischen der Partizipbildung eines Verbs wie *vorgetäuscht* und einem nominalen Kompositum wie *Leichenknochen* unterschieden werden.

Das Auftreten in attributivem Slot kann also nicht als Kriterium für die Unterscheidung von Distributionsklassen im Hawaiischen herangezogen werden, seien diese mit *Verb* und *Nomen* etikettiert oder nicht.

Geht man zusammenfassend davon aus, dass jedes Wort, das in Tempus-Aspekt-Modus-Syntagma auftritt, auch in „Nominalisierungs-konstruktion“ und attributivem Slot und jedes Wort, das in Artikel-Syntagma auftritt, auch in attributivem Slot auftreten kann und kein Wort potentiell ausschließlich in attributivem Slot, verschieben sich die im Korpus beobachteten und ebenso die von Elbert und Pukui (1979) vorgeschlagenen Distributionsklassen, vgl.:

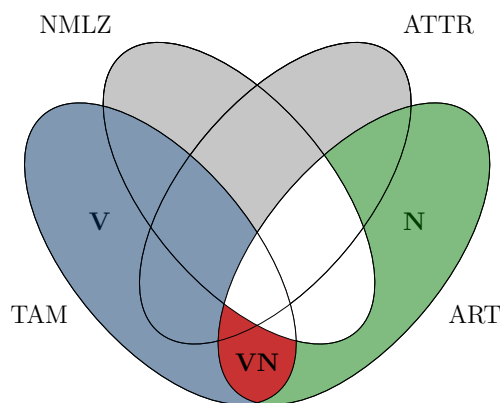


Abbildung 20: Venn-Diagramm korrigierter Distributionsklassen

In Abbildung 20 (S. 129) sind die potentiellen Distributionsklassen des Hawaiischen, die sich nach dem Vergleich der Beobachtungen im Korpus mit den Angaben in den Wörterbüchern von Pukui und Elbert (1986) und Parker (1922) ergeben, in einem korrigierten Venn-Diagramm dargestellt.

Da im Hawaiischen vermutlich kein Wort potentiell ausschließlich in „Nominalisierungs-konstruktion“ und anders als im Korpus beobachtet, kein Wort ausschließlich in attributivem Slot sowie ausschließlich in „Nominalisierungs-konstruktion“ und attributivem Slot auftritt, sind die entsprechenden Teile der Ellipsen im Diagramm grau gefärbt.

Im Wortartsystem von Elbert und Pukui (1979) müssen zuerst *Nomen* dahingehend korrigiert werden, dass nicht Wörter, die ausschließlich in Artikel-Syntagma, sondern Wörter, die potentiell in Artikel-Syntagma und attributivem Slot auftreten, als solche bestimmt sind. Aus distributionaler Perspektive sind dann *Verben* als Wörter definiert, die in allen Syntagmen außer Artikel-Syntagma auftreten, *Nomen* als Wörter, die in allen Syntagmen außer Tempus-Aspekt-Modus-Syntagma und „Nominalisierungs-konstruktion“ und *Nomen-Verben* als Wörter, die in Tempus-Aspekt-Modus-Syntagma, Artikel-Syntagma und attributivem Slot auftreten.

Würde man unter Berücksichtigung der Ergebnisse aus dem Abgleich der Korpusbeobachtungen mit den Angaben in den Wörterbüchern von Pukui und Elbert (1986) und von Parker (1922) die im Korpus auftretenden Wörter nach den in Abbildung 20 (S. 129) gebildeten Distributionsklassen als *Verb*, *Nomen* und *Nomen-Verb* kategorisieren, wären im Gegensatz zur ursprünglichen Beobachtung (vgl. Tabelle 12, S. 117) die Mehrheit der Wörter im Korpus erfasst (vgl. Tabelle 17, S. 131).

Wie sich Tabelle 17 (S. 131) entnehmen lässt, ist es nun zwar möglich, 1485 (0,66%) von den 2250 im Korpus auftretenden Wörter mit einer Kategoriene-tikette zu versehen. Erstens bleibt der kategoriale Status von *Nomen-Verben*, die die Mehrheit der erfassten Wörter bilden, bislang aber weiterhin ungeklärt. Zweitens ist es immer noch bei ungefähr einem Drittel der im Korpus auftretenden Wörter (765 von 2250) nicht möglich, ein Kategorienetikett zu finden. Nimmt man *Nomen-Verben* und die nicht erfassten Wörter zusammen, können 1768 von 2250 (0,79%) der Wörter im Korpus nicht in eindeutiger Weise einer *Wortart* zugewiesen werden.

Neben der Frage des Umgangs mit Wörtern, die sowohl in Tempus-Aspekt-Modus-Syntagma als auch Artikel-Syntagma auftreten, bleibt bei dieser Klassifizierung das größte Problem die „Nominalisierungs-konstruktion“.

Kategorie	absolute Frequenz	relative Frequenz
erfasst		
Verb	173	0,08
Nomen	309	0,14
Nomen-Verb	1003	0,45
Σ	1485	0,66
nicht erfasst		
n/a	765	0,34
Σ	765	0,34
Σ	2250	1

Tabelle 17: erfasste und nicht-erfasste klassifizierte Wörter

Vier die „Nominalisierungsstruktur“ enthaltene Kombinationen von Syntagmen, in denen Wörter im Korpus ausschließlich auftreten, sind keiner Kategorie zuzuordnen (vgl. Abbildung 19, S. 115). Im Kern kann bei der Ansetzung von *Verben*, *Nomen* und *Nomen-Verben* nicht erklärt werden, warum nicht ausschließlich als *Verben* klassifizierte Wörter *nominalisiert* auftreten. Im nächsten Kapitel soll die „Nominalisierungsstruktur“ genauer betrachtet werden.

4.1.3 Nominalisierung im Hawaiischen

Der Grund dafür, in der Korpusanalyse das Auftreten von Wörtern in „Nominalisierungsstruktur“ separat vom Auftreten in Artikel-Syntagma zu erfassen und die „Nominalisierungsstruktur“ nicht unter das Artikel-Syntagma zu subsumieren, war folgender: Wenn die Annahme zutrifft, dass Wörter im Hawaiischen prädisponiert sind für das Auftreten in bestimmten syntaktischen Slots, sollten sich auch Unterschiede in der Verteilung von Wörtern auf die „Nominalisierungsstruktur“ zeigen (vgl. Kapitel 3.4.5, S. 94–95).

Im Wesentlichen muss bei der Bestimmung einer Konstruktion als „Nominalisierungsstruktur“ gezeigt werden, dass nur bestimmte Wörter darin auftreten, und zwar solche Wörter, die nicht als *Nomen* klassifiziert sind. Das bedeutet also, dass von einer potentiellen Beschreibung von *Nominalisierung* im Hawaiischen auch die potentielle Unterscheidung zwischen *Verb* und *No-*

men abhängt. Die Rede von *Nominalisierung* setzt die Definition einer *nominalen* Kategorie und mindestens einer nicht-*nominalen* Kategorie voraus, die *nominalisiert* werden kann. Dazu passt zunächst die Beobachtung, dass im Korpus kein Wort ausschließlich in „Nominalisierungs-konstruktion“ auftritt (vgl. Tabelle 11, S. 108). Ebenso kann festgehalten werden, dass die „Nominalisierungs-konstruktion“ insgesamt im Korpus deutlich seltener als die anderen Syntagmen vorkommt, vgl.:

Syntagma	absolute Frequenz	relative Frequenz
ART	103750	0,43
TAM	63898	0,26
ATTR	65989	0,27
NMLZ	9877	0,04
Σ	243514	1

Tabelle 18: Vorkommenshäufigkeit von Syntagmen im Korpus

Tabelle 18 lässt sich entnehmen, dass die „Nominalisierungs-konstruktion“ nur 4% aller im Korpus vorkommenden Syntagmen ausmacht bzw. dass nur 4% aller Worttokens im Korpus in „Nominalisierungs-konstruktion“ auftreten.

Zuerst sei hier wieder die Beschreibung der „Nominalisierungs-konstruktion“ im von Elbert und Pukui (1979) für das Hawaiische vorgeschlagene Wortartsystem betrachtet. Über Distributionsklassen sind *Verben* bestimmt als Wörter, die in Tempus-Aspekt-Modus-Syntagma und „Nominalisierungs-konstruktion“ auftreten und *Nomen* als Wörter, die in Artikel-Syntagma auftreten, wobei das Auftreten in attributivem Slot keine Rolle bei der Differenzierung spielt (vgl. Kapitel 1.1.2, S. 127–129). Genau genommen ist die Voraussetzung der Bestimmung von *Nominalisierung* eine Definition von *Verben* als Wörter, die für das Auftreten in Tempus-Aspekt-Modus-Syntagma prädisponiert sind und *Nomen* als Wörter, die für das Auftreten in Artikel-Syntagma prädisponiert sind. *Nominalisierung* bedeutet dann, dass ein *Verb* für das Auftreten in Artikel-Syntagma besonders ausgezeichnet werden muss, i.e. in „Nominalisierungs-konstruktion“ auftritt.

Ein Beispiel für einen solchen Fall ist das Wort *pa‘ē* „rufen“, das im Korpus ausschließlich in Tempus-Aspekt-Modus-Syntagma und „Nominalisierungs-konstruktion“ auftritt, vgl.:

- (10a) [ua **pa‘ē** mai-la] kona leo
 PFV rufen DIR-DIST 3POSS Stimme
 ‘Ihre Stimme rief.’
 (Desha 1995 [¹1920,1924]:114)

- (10b) Ua ho‘opū‘iwa nui ‘ia na‘e ‘o Kamehameha
 PFV erschrecken viel PASS aber TOP Kamehameha
- [i ka **pa‘ē** ‘ana mai] o kekahi
 PREP DEF rufen NMLZ DIR POSS ART
- leo ‘ano malihini iā ia
 Stimme Art fremd PREP 3SG

‘Kamehameha wurde aber sehr aufgeschreckt beim Rufen einer fremdartigen Stimme zu ihm.’
 (Desha [¹1920,1924]:252)

Im Wörterbuch von Pukui und Elbert (1986: 298) ist *pa‘ē* als *statives Verb* klassifiziert. Im Korpus tritt *pa‘ē* dann auch nicht in Artikel-Syntagma, sondern nur wie in Satz (10a) in Tempus-Aspekt-Modus-Syntagma oder wie in Satz (10b) in „Nominalisierungs-konstruktion“ auf.

Die „Nominalisierungs-konstruktion“ wird von Elbert und Pukui (1979: 80) allerdings nicht ausschließlich über die Distribution von Wörtern beschrieben, vgl.: „[...] the combination verb + ‘ana usually represents an ongoing process, frequently translated into English by the present participle.“

Zugrunde liegt hier die Hypothese einer *verbalen* Semantik, also der Definition von *Verb* als ein Wort, dessen Bedeutung der ontologischen Klasse EREIGNIS zugeordnet werden kann (vgl. die Ansätze in Kapitel 2.2.4, S. 25–55). Neben der Distribution eines Wortes soll sich dann an einer solchen *verbalen* Semantik von Wörtern im Auftreten in der Konstruktion mit ‘ana zeigen, dass es sich um *Nominalisierung* von *Verben* bzw. überhaupt um *Nominalisierung* handelt und nicht etwa um *Nomen* im Auftreten in Artikel-Syntagma.

Die von Elbert und Pukui (1979: 80) angesprochene Übersetzung ins Englische kann, wenn überhaupt, nur auf semantische Parallelen hinweisen und ist freilich kein Kriterium für die formale Beschreibung der *Nominalisierung* eines *Verbs* im Hawaiischen. Während das Gerundium im Englischen oder die infinite Verbform im Deutschen ein nominalisiertes Verb auch formal als solches aus-

zeichnen, lässt sich im Hawaiischen ein *Verb* morphologisch nicht analysieren. In der Definition von Elbert und Pukui (1979) sind *Verben* und *nominalisierte Verben* also nur distributional und semantisch beschrieben. Dies wirft im Hawaiischen Probleme auf, wie später noch zu zeigen ist.

Zunächst muss aber, um das Bild der *Nominalisierung* im Hawaiischen zu vervollständigen, eine weitere *Nominalisierungs*-Strategie betrachtet werden, die bislang keine Beachtung gefunden hat. Neben der „Nominalisierungs-konstruktion“, in der das ungebundene Morphem ‘*ana* auftritt, beschreiben Elbert und Pukui (1979: 80–83) *Nominalisierung* als Wortbildungsprozess, bei dem von einer *verbalen* Basis durch das Anhängen des *Nominalisierungs*-Suffix, das die Allomorphe *-na*, *-ana*, *-hana*, *-kana* und *-lana* hat, ein *Nomen* abgeleitet wird. Als Beispiel nennen Elbert und Pukui (1979: 82) unter anderem das als *Nomen* klassifizierte Wort *holoholo-na* „animal“, gebildet von der als *Verb* klassifizierten Basis *holoholo* „to go about“, oder das als *Nomen* klassifizierte Wort *kau-lana* „resting place“, das von der als *Verb* klassifizierten Basis *kau* „to place“ abgeleitet ist.

Elbert und Pukui (1979: 80–81) gehen davon aus, dass sich im Vergleich zu Wörtern, die in „Nominalisierungs-konstruktion“ auftreten, die Bedeutung eines derivierten *Nomen* stärker von der Bedeutung der Basis unterscheidet, spezifischer ist und „that *-na* words usually designate a single act or object.“ (Elbert und Pukui 1979: 80).

Im Kern unterscheiden Elbert und Pukui (1979) bei der *Nominalisierung* im Hawaiischen also ähnlich zur Gegenüberstellung von *verbal noun* und *deverbal noun* im Englischen (vgl. z.B. Comrie und Thompson 2007) oder von *substantiviertem Infinitiv* und *deverbalem Nomen* im Deutschen (vgl. z.B. Motsch 1999) zwischen *nominalisierten Verben* im Auftreten in „Nominalisierungs-konstruktion“ und derivierten *Nomen*, die von *Verben* abgeleitet sind und in Artikel-Syntagma auftreten. Die Analogie zum Deutschen und Englischen wird an späterer Stelle noch problematisiert.

Im Gegensatz zum Gebrauch von ‘*ana* sehen Elbert und Pukui (1979: 81) die Wortbildung mit *-na* und Alternanten als nicht mehr produktiv an, so dass zwar potentiell jedes *Verb* in „Nominalisierungs-konstruktion“ auftritt, nicht aber von jedem *Verb* ein *Nomen* durch Suffigierung abgeleitet werden kann. Sowohl ‘*ana* als auch *-na*, *-ana*, *-hana*, *-kana* und *-lana* lassen sich auf proto-polynesisch *-(*C*)*anga* zurückführen (vgl. Chung 1974, Clark 1981).

Auch in anderen polynesischen Sprachen sind Reflexe von PPN *-(*C*)*anga* zu finden. Allerdings differiert die Anzahl der Formen je nach Sprache deutlich,

vgl.:

Sprache	Reflexe von <i>-(C)anga</i>	Quelle
Tahitianisch	<i>-ra'a</i>	Tryon (1970)
Rarotongaisch	<i>=anga</i>	Buse (1963)
Marquesisch	<i>-na, -tina</i>	Krupa (2005)
Rapanuisch	<i>-(V)na, -hana</i>	du Feu (1996), Kieviet (2017)
Tuvaluisch	<i>-saga, -maga, -taga, -aga, -kaga</i>	Besnier (2000)
Tongaisch	<i>-'anga, -fanga, -nanga, -tanga, -nga</i>	Churchward (1953), Broschart (1997)
Hawaiisch	<i>-na, -ana, -hana, -kana, -lana, 'ana</i>	Elbert und Pukui (1979)
Renellese	<i>-ganga, -hanga, -manga, -nanga, -sanga, -tanga, -nga</i>	Elbert (1988)
Samoisch	<i>-ga, -faga, -'aga, -maga, -laga, -aga, -saga, -taga</i>	Mosel und Hovdhau- gen (1992)
Maorisch	<i>-nga, -anga, -inga, -hanga, -ranga, -kanga, -tanga, -manga</i>	Biggs (1969), Bauer (2001), Harlow (2006)

Tabelle 19: Reflexe von **(C)anga* in polynesischen Sprachen

Im Hawaiischen und in anderen polynesischen Sprachen (vgl. z.B. Harlow 2006: 121–123 und passim zum Maorischen) ist die Parallelität in der Konsonantenalternation in *Nominalisierungs*-Morphem, Transitivierungssuffix und Passivmorphem auffällig. Im Hawaiischen trägt das Transitivierungssuffix die Formen *-a'i* bzw. *-Ca'i* mit den Konsonanten *h, k, l, m, n* und *-i* bzw. *-Ci* mit *h, k* oder *'*, das Passivmorphem die ungebundene Form *'ia* oder z.B. die Suffixform *-ia* bzw. *-Cia* mit *h, k, l, m* und *n* (vgl. Elbert und Pukui 1979: 79). Sowohl im Hawaiischen als auch anderen polynesischen Sprachen findet sich bei einem Wort häufig der gleiche Konsonant in *Nominalisierungs*-Suffix und Passivsuffix, z.B. hawaiisch *komo* „to enter“ (vgl. Pukui und Elbert 1986: 164) mit Passivform *komo-hia* (vgl. Pukui und Elbert 1986: 165) und *nominalisierter* Form *komo-hana* „West“ (vgl. Pukui und Elbert 1986: 165) oder maorisch *whakaatu* „to show“ mit Passivform *whakaatu-ria* und *nominalisierter* Form *whakaatu-ranga* „exhibition“ (vgl. Harlow 2006: 121).

Hale (1968) und Sanders (1991) haben die Alternation in den Passivformen der polynesischen Sprachen darauf zurückgeführt, dass der Konsonant historisch zur Endung des Wortstamms gehört, z.B. proto-austronesisch **inum* mit hawaiisch *inu* „to drink“ und Passivform *inu-mia* (vgl. Pukui und Elbert 1986: 101) und analog maorisch *inu* „to drink“ mit Passivform *inu-mia* und *nominalisierter* Form *inu-manga* „drink“ (vgl. Harlow 2006: 121), wobei Uneinigkeit

zwischen Hale (1968) und Sanders (1991) darin herrscht, warum die Stammkonsonanten nach dem Verlust in den Passivformen bestehen geblieben sind, und wie erklärt werden kann, dass sowohl beim Passivsuffix als auch beim *Nominalisierungs*-Suffix synchron vereinzelt mehrere Formen nebeneinander bestehen (vgl. dazu auch Besnier 2000: 446 Anmerkung 25 und Bauer 2001: 391–392), z.B. maorisch *hopu* „to catch“ mit Passivform *hopu-kia*, aber *nominalisiert* *hopu-kanga*, *hopu-tanga* und *hopu-ranga* (vgl. Bauer 2001: 391).

Synchron lassen sich dann beim Passiv auch keine Regeln für das Auftreten einer bestimmten Morphemalternante angeben. Churchward (1953: 239 und *passim*) hat zwar versucht, für das Tongaische semantische Unterschiede im Gebrauch der verschiedenen Formen des *Nominalisierungs*-Suffix aufzuzeigen. Bei Churchward (1953: 239 und *passim*) bleibt allerdings unklar, ob es sich tatsächlich um systematische Bedeutungsunterschiede oder nur Einzelbeispiele handelt, und auch für das Hawaiische und andere polynesischen Sprachen können solche Unterschiede nicht eindeutig festgestellt werden.

Der einzige Unterschied im Gebrauch der Formen des *Nominalisierungs*-Suffix scheint wie beim Hawaiischen auch in anderen polynesischen Sprache mit der Differenzierung zweier Arten von *Nominalisierung* zusammenzuhängen.

In Abbildung 21 (S. 137) ist zum Vergleich der Beschreibung der *Nominalisierung* im Hawaiischen von Elbert und Pukui (1979) mit der Beschreibung anderer polynesischer Sprachen, beispielhaft die Beschreibung der *Nominalisierung* im Marquesischen von Krupa (2005) und im Samoanischen von Mosel und Hovdhaugen (1992) aufgeführt.

Grundlegend findet sich aber in dem meisten Arbeiten zur *Nominalisierung* in den polynesischen Sprachen (vgl. z.B. Chung 1974 oder Clark 1981) und in Einzelgrammatiken (vgl. z.B. Biggs 1969 zum Maorischen, Besnier 2000 zum Tuvaluischen oder Kieviet 2017 zum Rapanuischen) eine ähnliche Beschreibung.

Krupa (2005) für das Marquesische und Mosel und Hovdhaugen (1992: 194 und *passim*) für das Samoanische unterscheiden zwischen *syntaktischer Nominalisierung* und *lexikalischer Nominalisierung*.

Anders als im Hawaiischen kann zwar im Marquesischen und im Samoanischen formal nicht zwischen *Nominalisierungs*-Suffix und ungebundenem Morphem unterschieden werden. Wie *‘ana* in der hawaiischen „Nominalisierungs-konstruktion“ wird bei der *syntaktischen Nominalisierung* aber nur eine bestimmte Form der Reflexe auf PPN **(C)anga* verwendet: im Marquesischen das Suffix *-tina* im Gegensatz zu *-na* (vgl. Krupa 2005: 504), im Samoanischen

Hawaiisch (Elbert und Pukui 1979):						
verb		nmlz verb		verb		noun
EVENT	→	EVENT	vs	EVENT	→	OBJECT
<i>pa‘ē</i>		<i>ka pa‘ē ‘ana</i>		<i>holoholo</i>		<i>holoholo-na</i>
to call		the calling		to go about		animal
Marquesisch (Krupa 2005):						
verb		verbal noun		verb		noun
EVENT	→	EVENT	vs	EVENT	→	OBJECT
<i>ono</i>		<i>te ono-tina</i>		<i>moe</i>		<i>moe-na</i>
to hear		the hearing		to sleep		mat
Samoanisch (Mosel und Hovdhaugen 1992):						
verb		verbal noun		verb		deverbal noun
EVENT	→	EVENT	vs	EVENT	→	OBJECT
<i>su‘e</i>		<i>le su‘e=ga</i>		<i>tau</i>		<i>tau-laga</i>
to search		the searching		to anchor		harbour

Abbildung 21: Vergleich Nominalisierungsstrategien in polynesischen Sprachen

ausschließlich *-ga*, das neben den anderen Allomorphen zwar auch in der *lexikalischen Nominalisierung* auftritt, in *syntaktischer Nominalisierung* aber als Klitikon analysierbar ist, da es statt an das *Verb* auch an eine auf das *Verb* folgende Direktionalpartikel angehängt werden kann (vgl. Mosel und Hovdhaugen 1992: 197).

Die *lexikalische Nominalisierung* durch Anhängen eines Suffixes definieren Krupa (2005) und Mosel und Hovdhaugen wie Elbert und Pukui (1979: 80–83) als Wortbildungsprozess mit der Funktion „to derive new lexemes“ (Krupa 2005: 504), das heißt der Ableitung von *Nomen* bzw. „deverbal nouns“ (Mosel und Hovdhaugen 1992: 194) von *Verben*. Ähnlich wie bei Elbert und Pukui (1979: 80–81) fällt dann auch die semantische Analyse der Derivate aus. Mosel und Hovdhaugen (1992: 194) schreiben dazu: „The deverbal nouns denote the

object or the place of the event denoted by the verb [...].“

Die Abgrenzung der *syntaktischen Nominalisierung* von der *lexikalischen* basiert im Wesentlichen wie bei Elbert und Pukui (1979) auf der Annahme einer *verbalen* Semantik, über die Wörter im Auftreten in Artikel-Syntagma als *nominalisierte Verben* bzw. *Verbalnomen* analysiert werden können. So nehmen Mosel und Hovdhaugen (1991: 197) bei der *syntaktischen Nominalisierung* für das Samoanische an: „verbal nouns denote the action in question as a whole [...]“ und Krupa (2005: 507) vergleichbar für das Marquesische: „verbal nouns [...] usually express an action [...]“.

Ebenso wie Elbert und Pukui (1979: 81) für das Hawaiische gehen Mosel und Hovdhaugen (1992: 194 und passim) für das Samoanische davon aus, dass die *lexikalische Nominalisierung* anders als die *syntaktische Nominalisierung* nicht produktiv ist, während Krupa (2005) zur Produktivität im Marquesischen keine Angaben macht. Chung (1974: 646) behauptet allerdings, dass die *lexikalische Nominalisierung* in keiner polynesischen Sprache mehr produktiv ist. In den meisten polynesischen Sprachen scheint sich zumindest nur eine geringe Anzahl von Beispielen für die *lexikalische Nominalisierung* aufzählen zu lassen (vgl. z.B. Churchward 1953: 239 zum Tongaischen oder Besnier 2000: 592 zum Tuvaluischen).

An dieser Stelle muss angemerkt werden, dass mit der Beschreibung von *lexikalischer Nominalisierung* und *syntaktischer Nominalisierung* auf Basis der Unterscheidung zwischen einer *verbalen* und einer *nominalen* Semantik, in vielen Arbeiten zu polynesischen Sprachen davon ausgegangen wird, dass zudem *Nominalisierung* ohne morphologische Markierung durch ein Suffix analysiert werden kann (vgl. z.B. Chung 1974, Mosel und Hovdhaugen 1992, Besnier 2000, Cablitz 2000, Moyse-Faurie 2016, Kieviet 2017) und damit letztlich *Verb* und *Nomen* semantisch unterscheidbar sind.

Im Folgenden soll im Detail auf die Probleme, die bei der Beschreibung der *Nominalisierung* im Hawaiischen entstehen, eingegangen werden. Da es vorrangig um die Beschreibung von *Nominalisierung* geht, wird dabei die Frage der Unterscheidung von *Verb* und *Nomen* im Hawaiischen über die Einteilung lexikalisierten Wortbedeutungen in ontologische Klassen nur angeschnitten und erst später im Detail behandelt (vgl. Kapitel 4.3, S. 196–215).

Zunächst sei die *lexikalische Nominalisierung* betrachtet. Der Rede von *lexikalischer Nominalisierung* liegt die Annahme zugrunde, dass es sich dabei um einen Wortbildungsprozess, genauer um kategorienverändernde Derivation, handelt: Von einem *Verb*, also einem Wort, dass ausschließlich in Tempus-

Aspekt-Modus-Syntagma, attributivem Slot und „Nominalisierungs-konstruktion“ auftritt und dessen Bedeutung ontologisch als EREIGNIS klassifizierbar ist, wird durch Suffigierung mit einem der Allomorphe *-(a)na*, *-hana*, *-kana* oder *-lana* ein *Nomen* gebildet, also ein Wort, das ausschließlich in Artikel-Syntagma und attributivem Slot auftritt und dessen Bedeutung ontologisch als OBJEKT klassifizierbar ist (vgl. auch Kapitel 2.2.4, S. 25–55).

Wenn derivierte Wörter für das Auftreten in Artikel-Syntagma prädisponiert sind, ließe sich damit zeigen, dass im Hawaiischen – beachtet man die fehlende Produktivität – zumindest einige Wörter formal auch ohne Flexionsmorphologie als *Nomen* definiert werden können.

Da aus rein distributionaler Perspektive aus der Annahme von kategorienverändernder Derivation folgt, dass derivierte *Nomen* ausschließlich in Artikel-Syntagma und attributivem Slot und dessen *verbale* Basen ausschließlich in Tempus-Aspekt-Modus-Syntagma, attributivem Slot und „Nominalisierungs-konstruktion“ auftreten, lohnt es sich noch vor einer semantischen Analyse, einen Blick auf die Distribution derivierter Wörter und dessen Basen im Korpus zu werfen.

Tabelle 20 (S. 140) bietet einen Überblick über die (Kombinationen aus) Syntagmen, in denen *-(C)(a)na*-Derivate und dessen Basen im Korpus auftreten. Insgesamt treten im Korpus 49 Derivate auf. Nur von 23 dieser Derivate treten auch die Basen im Korpus auf.³⁶

In den Spalten von Tabelle 20 (S. 140) sind diejenigen (Kombinationen aus) Syntagmen angegeben, in denen die Derivate auftreten und in den Zeilen diejenigen (Kombinationen aus) Syntagmen, in denen die Basen auftreten, wobei diejenigen (Kombinationen aus) Syntagmen, in denen kein Wort im Korpus auftritt, in Tabelle 20 (S. 140) nicht aufgeführt sind.

In der dritten Zeile von Tabelle 20 (S. 140) ist zudem die Gesamtfrequenz des Auftretens eines Derivats (Types) in dem jeweiligen Syntagma bzw. der Kombination aus Syntagmen angegeben. So treten beispielsweise 11 Derivate im Korpus ausschließlich in Artikel-Syntagma auf, während nur ein Derivat ausschließlich in Tempus-Aspekt-Modus-Syntagma und Artikel-Syntagma auftritt. Von den 11 Derivaten, die ausschließlich in Artikel-Syntagma auftreten, lassen sich wiederum vier Basen im Korpus identifizieren, die alle in Tempus-

³⁶Betrachtet man die Version des Korpus vor der Eliminierung von Homonymen (vgl. Kapitel 3.4.6, S. 102–104), treten von allen Derivaten die Basen auf. Diese Basen kommen ausnahmslos sowohl in Tempus-Aspekt-Modus-Syntagma als auch Artikel-Syntagma, attributivem Slot und „Nominalisierungs-konstruktion“ vor.

Basis	Derivat									Σ
	ART	ATTR	TAM ^ ART	ART ^ ATTR	ART ^ NMLZ	TAM ^ ART ^ ATTR	TAM ^ ART ^ NMLZ	ART ^ ATTR ^ NMLZ	TAM ^ ART ^ ATTR ^ NMLZ	
Σ	11	1	1	12	1	14	2	4	3	49
ART						1				1
TAM ^ ART ^ ATTR						3				3
TAM ^ ART ^ NMLZ			1							1
TAM ^ ATTR ^ NMLZ							1			1
TAM ^ ART ^ ATTR ^ NMLZ	4			5	1	5		2		17

Tabelle 20: Vorkommen Basen und *-na*-Derivate im Korpus

Aspekt-Modus-Syntagma, Artikel-Syntagma, attributivem Slot und „Nominalisierungskonstruktion“ auftreten.

Insgesamt kann für die Verteilung von Basis-Derivat-Paaren auf (Kombinationen von) Syntagmen im Korpus folgendes festgehalten werden: Die aus der Annahme von kategorienverändernder Derivation hervorgehende Distribution von Basis-Derivat-Paaren kann überhaupt nicht belegt werden. Zwar finden sich unter den Derivaten im Korpus 23 Elbert und Pukui (1979) folgend als *Nomen* klassifizierte Wörter – das sind 11 Wörter, die ausschließlich in Artikel-Syntagma und 12 Wörter, die ausschließlich in Artikel-Syntagma und attributivem Slot, auftreten. Jedoch kann keinem dieser Derivate eine Basis zugeordnet

werden, die Elbert und Pukui (1979) folgend als *Verb* klassifiziert ist. Vielmehr tritt die Mehrheit der Wörter (17 von 23), die die Basis für eine Derivation bilden, im Korpus in allen vier Syntagmen auf. Diese Wörter wären damit Elbert und Pukui (1979) folgend als *Nomen-Verben* klassifiziert. Unter den im Korpus auftretenden Basen lässt sich nur ein Wort, das ausschließlich in Tempus-Aspekt-Modus-Syntagma, attributivem Slot und „Nominalisierungs-konstruktion“ auftritt, Elbert und Pukui (1979) folgend als *Verb* klassifizieren. Das von diesem Wort gebildete Derivat allerdings tritt in Tempus-Aspekt-Modus-Syntagma, Artikel-Syntagma und „Nominalisierungs-konstruktion“ auf und kann damit Elbert und Pukui (1979) folgend nicht als *Nomen* klassifiziert werden.

Auffällig bei den 49 Derivaten im Korpus ist dann nicht nur, dass 20 von diesen in Tempus-Aspekt-Modus-Syntagma auftreten, sondern vor allem, dass insgesamt 10 Derivate auch in „Nominalisierungs-konstruktion“ auftreten, das heißt, dass es sich bei 10 Wörtern im Korpus theoretisch um eine *Nominalisierung* von derivierten *Nomen* handeln müsste.

Wie bei der Verteilung aller 2250 im Korpus auftretenden Wörter auf (Kombinationen von) Syntagmen ist es nun möglich, dass die Auftretensweisen der Basen und der Derivate im Korpus nicht repräsentativ für das Hawaiische sind (vgl. Kapitel 4.1.1, S. 117–118). Das bedeutet – insbesondere unter Berücksichtigung der geringen Anzahl identifizierbarer Basen im Korpus – dass die potentielle Auftretensweise von Basen und Derivaten im Korpus nicht abgebildet ist. Wie bei der Evaluation der Verteilung aller Wörter besteht die Möglichkeit, die Ergebnisse aus der Korpusuntersuchung mit den Angaben in den Wörterbüchern von Pukui und Elbert (1986) und Parker (1922) zu vergleichen, um einen besseren Überblick über die potentielle Auftretensweise von Basen und Derivaten zu erhalten. Um das Bild der Distribution von Basis-Derivat-Paaren noch zu erweitern, können zusätzlich die Wörter mit in den Vergleich einbezogen werden, die von Elbert und Pukui (1979: 80–83) als Beispiele für die *lexikalische Nominalisierung* genannt werden und nicht im Korpus auftreten.

In Tabelle 21 (S. 142) sind die Ergebnisse aus dem Vergleich zusammengefasst. Im ersten Teil von Tabelle 21 (S. 142) sind die theoretischen *Wortartkategorien* festgehalten, die nach dem Vergleich zwischen den Angaben im Wörterbuch von Pukui und Elbert (1986) und Parker (1922) mit der Distribution im Korpus den im Korpus auftretenden Basis-Derivat-Paaren zugeschrieben werden können.

Korpus				
Basis	Derivat			Σ
	Nomen	Verb	Nomen-Verb	
Verb	3	0	5	8
Nomen	1	0	3	4
Nomen-Verb	20	0	17	37
Σ	24	0	25	49

Elbert und Pukui (1979)				
Basis	Derivat			Σ
	Nomen	Verb	Nomen-Verb	
Verb	7	1	6	14
Nomen	2	0	2	4
Nomen-Verb	5	0	14	19
Σ	14	1	22	37

Korpus \vee Elbert und Pukui (1979)				
Basis	Derivat			Σ
	Nomen	Verb	Nomen-Verb	
Verb	9	1	9	19
Nomen	3	0	3	6
Nomen-Verb	25	0	25	50
Σ	37	1	37	75

Tabelle 21: Zusammenfassung Basis-Derivat-Kategorien

Im zweiten Teil von Tabelle 21 sind die theoretischen *Wortartkategorien* für die von Elbert und Pukui (1979: 80–83) genannten Beispiele angegeben.

Im dritten Teil von Tabelle 21 sind schließlich die Ergebnisse aus den ersten beiden Teilen zusammengelegt, das heißt, es sind diejenigen Basis-Derivat-Paare angegeben, die entweder nur im Korpus auftreten oder nur von Elbert und Pukui (1979: 80–83) als Beispiele genannt werden, oder sowohl im Korpus auftreten als auch von Elbert und Pukui (1979: 80–83) als Beispiele angeführt sind.

Zunächst kann festgehalten werden, dass nach dem Abgleich mit den Angaben in den Wörterbüchern von Elbert und Pukui (1986) und von Parker (1922) zwar drei Basis-Derivat-Paare im Korpus vorkommen mit einem als *Verb* klassifizierten Wort als Basis und einem als *Nomen* klassifizierten Wort als Derivat. Allerdings ändert dies nicht, dass die Mehrheit der Wörter, die die Basis für ein Derivat bilden, als *Nomen-Verb* klassifiziert sind. Ebenso sind fast gleich viele Derivate als *Nomen-Verb* klassifiziert wie als *Nomen*.

Dem zweiten Teil von Tabelle 21 lässt sich entnehmen, dass Elbert und Pukui

(1979: 81) sogar ein Beispiel für ein Basis-Derivat-Paar nennen, bei dem nicht nur die Basis sondern auch die derivierte Form als *Verb* klassifiziert ist.

Zum Auftreten von Derivaten in „Nominalisierungs-konstruktion“ kann aufgrund fehlender Angaben in den Wörterbüchern von Elbert und Pukui (1986) und von Parker (1922) kein Vergleich stattfinden.

Aus der Zusammenlegung der im Korpus beobachteten Wörter und der Beispiele von Elbert und Pukui (1979: 80–83) lassen sich insgesamt 75 Derivate bzw. Basis-Derivat-Paare im Hawaiischen identifizieren, wobei sich die Anzahl der Fälle, die nicht über kategorienverändernde Derivation erklärt werden können, im Vergleich mit der im Korpus beobachteten Auftretensweise noch erhöht (vgl. erster und dritter Teil von Tabelle 21, S. 142). So finden sich z.B. statt einem Wort, drei Wörter mit einem *Nomen* als Basis und insgesamt wären von 75 Basis-Derivat-Paaren nur 9 mit kategorienverändernder Derivation erfasst.

Zusammenfassend lassen sich aus dem unteren Teil von Tabelle 21 (S. 142) und den Ergebnissen zum Auftreten von Derivaten in „Nominalisierungs-konstruktion“ im Korpus aus Tabelle 20 (S. 140) insgesamt sieben verschiedene Basis-Derivat-Paare extrahieren (vgl. Abbildung 22, S. 144).

Unter der distributionalen Annahme, dass *Verben* Wörter sind, die in Tempus-Aspekt-Modus-Syntagma, attributivem Slot und „Nominalisierungs-konstruktion“, nicht aber in Artikel-Syntagma auftreten, und von *Verben* derivierte *Nomen* als *Nomen* in Artikel-Syntagma und attributivem Slot, nicht aber in Tempus-Aspekt-Modus-Syntagma und „Nominalisierungs-konstruktion“ auftreten, kann nur eines der sieben identifizierten Basis-Derivat-Paaren im Hawaiischen (vgl. Abbildung 22, S. 144) beschrieben werden.

Möchte man an der Definition von *Wortarten* als Distributionsklassen festhalten, kann im Hawaiischen nicht von *Nominalisierung* gesprochen werden, weil sich dann bei den Derivaten das Auftreten in Tempus-Aspekt-Modus-Syntagma und in „Nominalisierungs-konstruktion“ und bei den Basen das Auftreten in Artikel-Syntagma nicht mit kategorienverändernder Derivation beschreiben lässt.

Wenn unter Beachtung der Verteilung von Basen und Derivaten auf Syntagmen nicht überzeugend ist, kategorienverändernde Derivation anzunehmen, stellt sich sogleich die Frage, welche Folgen dies für die semantische Beschreibung von *Nominalisierung* hat.

Der semantischen Beschreibung der *Nominalisierung* von Elbert und Pukui (1979: 80–83) liegt im Kern die Annahme zugrunde, *Wortarten* ließen sich im

	Basis		Derivat
1.	V	→	N-na → N-na in NMLZ
2.	V	→	V-na
3.	V	→	NV-na → NV-na in NMLZ
4.	N	→	N-na
5.	N	→	NV-na
6.	NV	→	N-na → N-na in NMLZ
7.	NV	→	NV-na → NV-na in NMLZ

Abbildung 22: Kategorien identifizierbarer Basis-Derivat-Paare im Hawaiischen

Hawaiischen semantisch unterscheiden: *Verben* sind Wörter, dessen Bedeutung der ontologischen Klasse EREIGNIS und *Nomen* sind Wörter, dessen Bedeutung der ontologischen Klasse OBJEKT angehört, wobei die ontologischen Klassen auf die Distributionsklassen abdeckt werden. Aus dieser Perspektive wird bei kategorienverändernder Derivation die ontologische Klasse der Bedeutung einer *verbalen* Basis, die prädisponiert ist für das Auftreten in Tempus-Aspekt-Modus-Syntagma, durch die Ableitung in eine andere ontologische Klasse überführt, die einem *nominalen* Derivat, das für das Auftreten in Artikel-Syntagma prädisponiert ist, entspricht. Bei einem Basis-Derivat-Paar muss demzufolge die Bedeutung der Basis der ontologischen Klasse EREIGNIS und die Bedeutung des Derivats der ontologischen Klasse OBJEKT zugeordnet werden können.

Im Folgenden sollen nacheinander die sieben in Abbildung 22 (S. 144) festgehaltenen Fälle von Basis-Derivat-Paaren anhand von Beispielen aus dem

Korpus und von Elbert und Pukui (1979) genauer analysiert werden.

Zu 1.: Ein Beispiel für einen unter 1. in Abbildung 22 (S. 144) fesgehaltenen Fall eines Basis-Derivat-Paares ist *loi* und *loi-na*. Die Basis *loi* klassifizieren Pukui und Elbert (1986: 209) als *Verb* mit der Bedeutung „to look over critically“ und das Derivat *loi-na* klassifizieren Pukui und Elbert (1986: 210) als *Nomen* unter anderem mit der Bedeutung „observation“. In einem Fall wie *loi* und *loi-na* läuft die Beschreibung von kategorienverändernder Derivation parallel zum Englischen bei *observe* und *observ-ation* oder Deutschen bei *beobachten* und *Beobacht-ung* mit der Klassifizierung von *loi* als *Verb* und *loi-na* als *Nomen Actionis*.

Bei Basis-Derivat-Paaren wie *loi* und *loi-na* sollte der Definition von *Verben* von Elbert und Pukui (1979) zufolge die Basis potentiell auch in „Nominalisierungs-konstruktion“ auftreten (vgl. Kapitel 4.1.1, S. 112–115). So tritt z.B. die von Pukui und Elbert (1986: 301) als *Verb* mit der Bedeutung „to push“ klassifizierte Basis *pahu* des Derivats *pahu-na*, das Pukui und Elbert (1986: 301) als *Nomen* mit der Bedeutung „a pushing“ klassifizieren, im Korpus *nominalisiert* in der Konstruktion *ka pahu ‘ana* mit der Bedeutung „the pushing“ auf. Auch wenn es nicht möglich scheint, einen semantischen Unterschied zwischen *pahu-na* und *ka pahu ‘ana* zu analysieren, könnte distributional und formal zwischen einem derivierten *Nomen Actionis* und einem *nominalisierten Verb* unterschieden werden. An der Gegenüberstellung von *pahu-na* und *ka pahu ‘ana* zeigt sich, dass die Semantik bei der Unterscheidung von *Nomen Actionis* und *nominalisierten Verb* nur eine untergeordnete Rolle spielt und kategorienverändernde Derivation primär nicht über die Semantik, sondern die Abdeckung von Formklassen auf Distributionsklassen definiert ist.

Solange gezeigt werden kann, dass *pahu* für das Auftreten in Tempus-Aspekt-Modus-Syntagma prädisponiert ist und im Gegensatz dazu *pahu-na* für das Auftreten in Artikel-Syntagma, ist für die Unterscheidung zwischen *pahu* als *nominalisiertem Verb* in *ka pahu ‘ana* und *pahu-na* als deriviertem *Nomen* die ontologische Klassifizierung der Bedeutung von *pahu* und *pahu-na* irrelevant. Problematisch wird es aber dann, wenn bei Basis-Derivat-Paaren mit einer als *Verb* klassifizierten Basis und einem als *Nomen* klassifizierten Derivat nicht nur die Basis sondern auch das Derivat in „Nominalisierungs-konstruktion“ auftreten kann. Im Korpus tritt z.B. sowohl die Basis *hiki*, die Pukui und Elbert (1986: 69) als *Verb* klassifizieren mit der Bedeutung „to come“, als auch das Derivat *hiki-na*, das Pukui und Elbert (1986: 69) als *Nomen* klassifizieren mit der Bedeutung „East“, in „Nominalisierungs-konstruktion“ auf, vgl.:

(11a) *ma* [*ka hiki* 'ana] *i Kealakekua*
 PREP DEF ankommen NMLZ PREP Kealakekua
 'Bei der Ankunft in Kealakekua'
 (Desha 1996 [1920,1924]:134)

(11b) *ha'o wale ho'i* [*ka hiki-na* 'ana mai]
 ersehnen INTSF INTSF DEF ankommen-NMLZ NMLZ DIR

e o'u piko
 PREP 1SG.POSS Verwandter

'Die Ankunft von meinem Verwandten war lange ersehnt.'
 (Desha 1996 [1920,1924]:114)

Ich sehe hier zwei Alternativen der Analyse, die beide Probleme für die Definition von *Nominalisierung* im Hawaiischen mit sich bringen: Wenn die Beschreibung der Konstruktion mit 'ana als *Nominalisierung* von *Verben* beibehalten werden soll, kann es sich bei *hiki-na* nicht um ein deriviertes *Nomen* handeln. Wenn dagegen an der Klassifizierung von *hiki-na* als *Nomen* festgehalten wird, kann die Konstruktion mit 'ana nicht als *Nominalisierung* von *Verben* beschrieben werden. Dieses Dilemma lässt sich auch nicht mithilfe einer ontologischen Klassifizierung der Bedeutungen von *hiki* und *hiki-na* lösen. Denn erstens ist kein klarer Bedeutungsunterschied zwischen *hiki* in *ka hiki* 'ana in Satz (11a) und *hiki-na* in *ka hiki-na* 'ana in Satz (11b) erkennbar. Zweitens ist zwar ein Unterschied zwischen der Bedeutung des Auftretens von *hiki-na* in „Nominalisierungs-konstruktion“ und in Artikel-Syntagma erkennbar. Während *hiki-na* in „Nominalisierungs-konstruktion“ in Satz (11b) die Ankunft an einem Ort bezeichnet, ist die Bedeutung von *hiki-na* mit „Osten“, also dem Ort, wo die Sonne ankommt (vgl. Parker 1922: 123), im Auftreten in Artikel-Syntagma spezifischer. Der semantische Unterschied zwischen den Auftretensweisen von *hiki-na* lässt sich jedoch nicht über *Nominalisierung* erklären.

Die Frage, ob Wörter, die in „Nominalisierungs-konstruktion“ auftreten, als *nominalisierte Verben* klassifiziert werden können, wird an späterer Stelle bei der genaueren Analyse der „Nominalisierungs-konstruktion“ wieder aufgegriffen. Zur Frage der kategorienverändernden Derivation seien zunächst weiterhin die in Abbildung 22 (S. 144) festgehaltenen Fälle betrachtet.

Zu 2.: Weder aus distributionaler noch aus semantischer Perspektive kann kategorienverändernde Derivation angesetzt werden, wenn sowohl Basis als auch Derivat ausschließlich in Tempus-Aspekt-Modus-Syntagma auftreten. Ein auf diese Weise verteiltes Basis-Derivat-Paar kommt zwar nicht im Korpus vor, Elbert und Pukui (1979: 81) aber haben mit *lūlā* und *lūlā-na* ein Beispiel für einen solchen Fall identifiziert: Elbert und Pukui (1979: 81) sowie Pukui und Elbert (1986: 214–215) klassifizieren sowohl *lūlā* als auch *lūlā-na* als *Verb* mit der Bedeutung „calm“. Um die Derivation mit *-na* und Alternanten im Hawaiischen weiterhin als kategorienverändernde Derivation zu beschreiben, muss entweder davon ausgegangen werden, dass es sich bei *lūlā-na* um einen Sonderfall handelt oder die Analyse als Derivat nicht zutrifft. Es ist allerdings nicht auszuschließen, dass sich im Hawaiischen noch mehr Basis-Derivat-Paare ausfindig machen lassen, bei denen sowohl Basis als auch Derivat als *Verb* klassifiziert sind. Zumindest kommen in anderen polynesischen Sprachen mehrere solcher Fälle vor (vgl. z.B. Mosel und Hovdhaugen 1992: 195 zum Samoanischen).

Zu 3.: In Fällen, in denen die Basis zwar als *Verb* kategorisiert ist, das Derivat allerdings Elbert und Pukui (1979) zufolge als *Nomen-Verb*, das heißt sowohl in Tempus-Aspekt-Modus-Syntagma als auch Artikel-Syntagma auftritt, ist dann genauso unklar, wie kategorienverändernde Derivation analysiert werden soll.

So findet sich z.B. die Basis *pai*, die Pukui und Elbert (1986: 301) als *Verb* klassifizieren mit der Bedeutung „to lift“ und davon abgeleitet das Wort *pai-na*, das Pukui und Elbert (1986: 302) als *Nomen-Verb* klassifizieren mit der Bedeutung „to lift, lifting“. Kategorienverändernde Derivation lässt sich bei *pai* und *pai-na* nur dann beschreiben, wenn man ignoriert, dass *pai-na* ohne Bedeutungsunterschied in dem gleichen Syntagma auftritt wie *pai*.

Fälle, in denen die Basis als *Verb* und das Derivat als *Nomen-Verb* klassifiziert ist, werden für die distributionale und die semantische Begründung von kategorienverändernder Derivation auch dann problematisch, wenn ein semantischer Unterschied zwischen Basis und Derivat analysierbar ist. So betrachten Elbert und Pukui (1979: 83) z.B. *kule-ana* als mögliches Derivat von *kūlē*. Pukui und Elbert (1986: 179) klassifizieren *kūlē* als *Verb* mit der Bedeutung „to seize, meddle in the concerns of others“ und *kule-ana* als *Nomen-Verb* unter anderem mit der Bedeutung „right, concern, cause, property“ im Auftreten in Artikel-Syntagma und unter anderem mit der Bedeutung „to cause, entitle, give right to process“ im Auftreten in Tempus-Aspekt-Modus-Syntagma.

Weder aus distributionaler noch aus semantischer Perspektive kann bei *kule-ana* Derivation als kategorienverändernde Derivation analysiert werden und die Bedeutungsveränderung des Derivats *kule-ana* gegenüber der Basis *kūlē* muss offensichtlich nicht gleichgesetzt werden mit der Überführung einer der ontologischen Klasse EREIGNIS zugehörigen Bedeutung des *Verbs* in eine der ontologischen Klasse OBJEKT zugehörigen Bedeutung des *Nomens*.

Zu 4.: Dass der Bedeutungsunterschied zwischen Basis und Derivat nicht mit der Unterscheidung der ontologischen Klassen EREIGNIS und OBJEKT beschrieben werden muss, zeigt sich dann noch deutlicher in Fällen, in denen sowohl die Basis als auch das Derivat als *Nomen* klassifiziert wird, aber trotzdem ein Bedeutungsunterschied zwischen Basis und Derivat feststellbar ist. Elbert und Pukui (1979: 82) nennen für einen solchen Fall als Beispiel die Basis *kumu*, die Pukui und Elbert (1986: 182) als *Nomen* klassifizieren mit der Bedeutung „source, bottom, teacher“ und das davon abgeleitete Wort *kumu-hana*, das Pukui und Elbert (1986: 182) ebenso als *Nomen* klassifizieren mit der Bedeutung „topic, subject“. Entscheidend ist hier, dass trotz Bedeutungsunterschied zwischen *kumu* und *kumu-hana* weder über Distributionsklassen noch über semantisch-ontologische Klassen ein kategorialer Unterschied zwischen *kumu* und *kumu-hana* bestimmt werden kann.

Es muss angemerkt werden, dass sich allerdings nicht bei jedem Basis-Derivat-Paar, bei dem sowohl Basis als auch Derivat als *Nomen* klassifiziert werden, ein semantischer Unterschied in klarer Weise analysieren lässt, so z.B. nicht bei *hope* „result“ und *hope-na* „result“ (Pukui und Elbert 1986: 82), wobei das Problem der Beschreibung von kategorienverändernder Derivation freilich in gleicher Weise besteht.

Zu 5.: Es ist dann auch nicht möglich, bei Basis-Derivat-Paaren *Nominalisierung* anzusetzen, bei denen die Basis als *Nomen* und das Derivat als *Nomen-Verb* kategorisiert wird.

Als Beispiel für einen solchen Fall findet sich die Basis *maka*, die Pukui und Elbert (1986: 224) als *Nomen* klassifizieren mit der Bedeutung „beloved one, favorite“ und das Derivat *maka-na*, das Pukui und Elbert (1986: 227) als *Nomen-Verb* klassifizieren mit der Bedeutung „gift“ im Auftreten in Artikel-Syntagma und der Bedeutung „to give a gift“ im Auftreten in Tempus-Aspekt-Modus-Syntagma. Ein anderes Beispiel ist das Wort *‘ohā*, das Pukui und Elbert (1986: 276) als *Nomen* klassifizieren mit der Bedeutung „taro corm“ als Basis für das Derivat *‘oha-na*, das Pukui und Elbert (1986: 276) als *Nomen-Verb* klassifizieren mit der Bedeutung „family“ im Auftreten in Artikel-Syntagma und der Be-

deutung „to be related“ in Tempus-Aspekt-Modus-Syntagma. Weder bei *maka* und *maka-na* noch bei *'ohā* und *'oha-na* kann eine Ableitung aus *verbaler* Basis mit *verbaler* Semantik beschrieben werden und der Bedeutungsunterschied zwischen *maka* „beloved person“ und *maka-na* „gift“ sowie zwischen *'ohā* „tarom“ und *'oha-na* „family“ ist nicht als Unterschied ontologischer Klassen zu analysieren.

Dasselbe gilt auch, wenn nicht eindeutig zwischen der Bedeutung der Basis und der Bedeutung des Derivats im Auftreten in Artikel-Syntagma unterschieden werden kann, wie bei *'ali* „scar“, das Pukui und Elbert (1986: 20) als *Nomen* klassifizieren und *'āli-na* „to scar, scarred“, das Pukui und Elbert (1986: 21) als *Verb* klassifizieren.

Zu 6.: Ein ähnliches Bild wie bei den bisher betrachteten Basis-Derivat-Paaren zeigt sich dann auch bei denjenigen, bei denen die Basis als *Nomen-Verb* und das Derivat als *Nomen* kategorisiert wird.

Während sich in manchen Fällen wieder kein Unterschied zwischen der Bedeutung der Basis im Auftreten in Artikel-Syntagma und der Bedeutung des Derivats analysieren lässt, z.B. bei *hai* „to sacrifice, sacrifice“ (Pukui und Elbert 1986: 47) und *hai-na* „sacrifice“ (Pukui und Elbert 1986: 48), kann in manchen Fällen ein Unterschied in der Bedeutung der Basis im Auftreten in Artikel-Syntagma zur Bedeutung des Derivats im Auftreten in Artikel-Syntagma analysiert werden. So findet sich z.B. zu dem Wort *moku*, das Pukui und Elbert (1986: 252) als *Nomen-Verb* klassifizieren mit der Bedeutung „to be cut“ im Auftreten in Tempus-Aspekt-Modus-Syntagma und der Bedeutung „island“ im Auftreten in Artikel-Syntagma, das Wort *moku-na* als Derivat, das Pukui und Elbert (1986: 252) als *Nomen* klassifizieren mit der Bedeutung „division“. Um bei *moku* und *moku-na* auf Basis semantisch-ontologischer Klassen kategorienverändernde Derivation anzusetzen, müsste man in der Konsequenz zusätzlich Null-Derivation bei *moku* im Auftreten in Artikel-Syntagma beschreiben, so dass von der *verbalen* Basis *moku₁* „to be cut“ ein *nominales* Derivat *moku₂* „island“ und ein *nominales* Derivat *moku-na* „division“ existiert. Unabhängig vom Problem der Bestimmung der Derivationsrichtung zwischen *moku₁* und *moku₂* (vgl. auch Kapitel 2.2.1, S. 13–14), lässt sich unter Berücksichtigung von Fällen wie *kumu* und *kumu-hana*, bei denen ein semantischer Unterschied zwischen einer *nominalen* Basis und einem *nominalen* Derivat besteht, nicht entscheiden, ob sich die Bedeutung von *moku-na* von der *nominalen* Bedeutung „island“ von *moku* im Auftreten in Artikel-Syntagma oder der *verbalen* Bedeutung „to be cut“ im Auftreten in Tempus-Aspekt-Modus-Syntagma ab-

leitet.

Zu 7: Ähnliches gilt dann auch für Fälle, bei denen sowohl Basis als auch Derivat als *Nomen-Verben* klassifiziert werden.

So findet sich z.B. die Basis *make*, die Pukui und Elbert (1986: 228) als *Nomen-Verb* klassifizieren mit der Bedeutung „die“ im Auftreten in Tempus-Aspekt-Modus-Syntagma und der Bedeutung „death“ im Auftreten in Artikel-Syntagma und das davon abgeleitete Wort *make-na*, das Pukui und Elbert (1986: 229) ebenso als *Nomen-Verb* klassifizieren mit der Bedeutung „lament“ im Auftreten in Tempus-Aspekt-Modus-Syntagma und der Bedeutung „lamentation“ im Auftreten in Artikel-Syntagma. Ein weiteres Beispiel, das Elbert und Pukui (1979: 82) nennen ist die Basis *mana‘o* „think, thought“ mit dem Derivat *mana‘o-lana* „to hope, hope“.

Bei den Derivaten *make-na* und *mana‘o-lana* kann zwar ein Bedeutungsunterschied zur jeweiligen Basis festgestellt werden, ohne dass dieser allerdings in Zusammenhang mit einem kategorialen Unterschied steht, sei es über Distributionklassen oder über semantisch-ontologische Klassen definiert.

Dasselbe gilt wie schon bei zuvor betrachteten Basis-Derivat-Paaren auch, wenn sich kein Bedeutungsunterschied zwischen Basis und Derivat analysieren lässt, z.B. bei *ka‘i* „to lead, order“ (vgl. Pukui und Elbert 1986: 115) und *kai‘ina* „to lead, order“ (vgl. Pukui und Elbert 1986: 117).

Betrachtet man zusammenfassend die gerade analysierten, sieben im Hawaiischen identifizierten Fälle von Basis-Derivat-Paaren lässt sich festhalten, dass kategorienverändernde Derivation weder aus rein distributionaler noch aus rein semantischer Perspektive beschrieben werden kann. Ebenso wenig lässt sich die Abdeckung von semantisch-ontologischen auf distributionale Klassen zeigen.

Während in manchen Fällen der Bedeutungsunterschied zwischen Basis und Derivat verloren gegangen zu sein scheint oder es einfach nicht möglich ist, die feinen semantischen Unterschiede des Hawaiischen herauszuarbeiten, kann zwar in vielen Fällen ein Bedeutungsunterschied zwischen Basis und Derivat analysiert werden. Dieser ist aber weder mit dem Unterschied zwischen semantisch-ontologischen Klassen EREIGNIS und OBJEKT gleichzusetzen noch auf bestimmte Distributionsklassen beschränkt. Im Kern lässt sich im Hawaiischen anders als z.B. im Deutschen oder Englischen nicht systematisch zwischen kategorienverändernder und nicht-kategorienverändernder Derivation unterscheiden und damit auch nicht von *Nominalisierung* sprechen.

Unklar bleibt dann, ob es sich bei der Derivation mit *-na* und Alternanten diachron um *Nominalisierung* handelt, die nur synchron nicht mehr analysier-

bar ist, oder ob die Rekonstruktion von **(C)anga* als *Nominalisierungs*-Suffix unzutreffend ist. Für die synchrone Beschreibung von *Nominalisierung* im Hawaiischen ist dies aber ohnehin nicht von Belang.

Der Versuch der Beschreibung von kategorienverändernder Derivation hat ebenfalls zu einem Dilemma in der Beschreibung von *Nominalisierung* im Hawaiischen insgesamt geführt, weil *lexikalische* und *syntaktische Nominalisierung* nicht zugleich angesetzt werden können, wenn derivierte *Nomen* syntaktisch *nominalisiert* auftreten.

Das führt wieder zurück zur eingangs betrachteten Annahme, dass bestimmte Wörter für das Auftreten in „Nominalisierungs-konstruktion“ prädisponiert sind und sich dadurch nicht nur die „Nominalisierungs-konstruktion“ als solche definieren lässt, sondern auch *Verben* in Abgrenzung zu *Nomen* und *Nomen-Verben*.

Wie nun aus der Betrachtung der *nominalisierten* Derivate und schon aus der bisherigen Beschreibung der Wortverteilung im Korpus und dem Vergleich mit den Wörterbüchern von Pukui und Elbert (1986) und Parker (1922) (vgl. Kapitel 4.1.1–4.1.2) hervorgeht, ist das Auftreten in „Nominalisierungs-konstruktion“ im Hawaiischen nicht auf bestimmte *Wortarten* beschränkt, vgl.:

	Kategorie		Kategorie in NMLZ
1.	V	→	V in NMLZ
2.	N	→	N in NMLZ
3.	NV	→	NV in NMLZ

Abbildung 23: Kategorien in „Nominalisierungs-konstruktion“

In Abbildung 23 sind die Kategorienetiketten angegeben, die Elbert und Pukui (1979) für die entsprechend verteilten Wörter in Korpus verwenden. Während mit der Definition der „Nominalisierungs-konstruktion“ als *syntaktische Nominalisierung* von *Verben* aus distributionaler Perspektive Wörter erfasst werden, die im Korpus ausschließlich in Tempus-Aspekt-Modus-Syntagma und „Nominalisierungs-konstruktion“ auftreten, wie *pa‘ē* in den Sätzen (10a) – (10b) (S.

133), fallen diejenigen Wörter, die in Artikel-Syntagma und „Nominalisierungs-konstruktion“ sowie in Tempus-Aspekt-Modus-Syntagma, Artikel-Syntagma und „Nominalisierungs-konstruktion“ auftreten, aus dieser Definition heraus. Ein großes Problem bei der Beschreibung der „Nominalisierungs-konstruktion“ als *Nominalisierung* von *Verben* stellt also zunächst die Beobachtung dar, dass im Korpus Wörter ausschließlich in Artikel-Syntagma und „Nominalisierungs-konstruktion“ auftreten. Demzufolge müssen theoretisch Wörter, die Elbert und Pukui (1979), Pukui und Elbert (1986) und Parker (1922) als *Nomen* klassifizieren, im Hawaiischen *nominalisiert* werden können, vgl.:

(12a) *ma ke kakahiaka nui*
 PREP DEF Morgen früh
 ‘am frühen Morgen’
 (Hale‘ole [1863]:83)

(12b) *ma ke kakahiaka ‘ana a‘e*
 PREP DEF Morgen NMLZ DIR
 ‘am Morgen’
 (Hale‘ole 2001 [1863]:259)

Pukui und Elbert (1986: 118) klassifizieren das Wort *kakahiaka* als *Nomen*, das heißt als Wort, das für das Auftreten in Artikel-Syntagma prädisponiert ist. Im Korpus tritt *kakahiaka* zwar nicht in Tempus-Aspekt-Modus-Syntagma, sondern wie in Satz (12a) in Artikel-Syntagma auf, allerdings wie in Satz (12b) ebenso in „Nominalisierungs-konstruktion“.

Ähnlich wie bei *hiki-na* in Satz (11b) (S. 146) lässt ist bei einem Wort wie *kakahiaka*, das über das Auftreten in Artikel-Syntagma als *Nomen* klassifiziert wird, aus distributionaler Sicht nicht von *Nominalisierung* im Auftreten in der Konstruktion mit *‘ana* sprechen.

Elbert und Pukui (1979: 80) selbst weisen nun darauf hin, dass auch Wörter, die sie als *Nomen* klassifizieren, in „Nominalisierungs-konstruktion“ auftreten können und nennen als Beispiele *‘āina* „land“ mit *ka ‘āina ‘ana* „the giving / forming / distribution of land“ und *mahi ‘ai* „farmer“ mit *mahi ‘ai (‘ia) ‘ana o ka āina* „the farming of the land“. ³⁷ Elbert und Pukui (1979: 80) schreiben

³⁷Es ist nicht ganz nachvollziehbar, warum Elbert und Pukui (1979: 80) *mahi ‘ai* hier als Beispiel nennen, obwohl z.B. im Wörterbuch von Pukui und Elbert (1986: 219) *mahi ‘ai* als „transitive noun-verb“ mit der Bedeutung „farmer, to farm“ angegeben ist.

dazu weiter: „After words commonly used as nouns, ‘*ana* seems to give a verblike meaning to the noun head, which of course remains a noun.“

Im Gegensatz zur Bedeutung von *kakahiaka* im Auftreten in Artikel-Syntagma in Satz (12a) (S. 152) müsste dieser Beschreibung von Elbert und Pukui (1979: 80) entsprechend die Bedeutung von *kakahiaka* im Auftreten in „Nominalisierungs-konstruktion“ in Satz (12b) (S. 152) *verbal*, z.B. als das Werden des Morgens, interpretiert werden. Abgesehen von der Frage der Interpretation der Bedeutung von *kakahiaka* in Satz (12b) (S. 152) und der unabhängigen Bestimmung einer *verbalen* Semantik, wirft der Versuch von Elbert und Pukui (1979: 80), über die Semantik zu erklären, dass nicht nur Wörter, die in Tempus-Aspekt-Modus-Syntagma auftreten, *nominalisiert* werden können, einige Probleme auf.

So definieren Elbert und Pukui (1979: 80) *Nominalisierung* auf zwei unterschiedliche Weisen, und zwar jeweils angepasst an die angenommene *Wortart*, die *nominalisiert* wird. Dafür ist Voraussetzung, dass *Wortarten* nicht nur über die Distribution bestimmt sind, sondern auch als semantische Klassen: ein *Verb*, z.B. *pa‘ē*, syntaktisch-distributional über das Auftreten in Tempus-Aspekt-Modus-Syntagma und semantisch über die Zuweisung der Bedeutung zur ontologischen Klasse EREIGNIS, ein *Nomen*, z.B. *kakahiaka*, syntaktisch-distributional über das Auftreten in Artikel-Syntagma und semantisch über die Zuweisung der Bedeutung zur ontologischen Klasse OBJEKT. Die *Nominalisierung* eines *Verbs* ist dann aber syntaktisch-distributional beschrieben als das Auftreten eines Wortes in Artikel-Syntagma, das für das Auftreten in Tempus-Aspekt-Modus-Syntagma prädisponiert ist, wie bei *pa‘ē* im Vergleich von Satz (10a) zu (10b) (S. 133). Die *Nominalisierung* eines *Nomens* hingegen ist semantisch beschrieben als die Veränderung der ontologischen Klasse der Bedeutung eines Wortes, das nur in Artikel-Syntagma auftritt, wie bei *kakahiaka* im Vergleich von Satz (12b) zu (12a) (S. 152). Genau genommen müssten Elbert und Pukui (1979) bei der *Nominalisierung* von *Nomen* von *Verbalisierung* sprechen, wobei unklar ist, wie ein rein semantisch definierter Begriff der *Verbalisierung* verstanden werden soll.

Dazu kommt noch, dass nicht in jedem Fall ein klarer Bedeutungsunterschied zwischen dem Auftreten eines von Pukui und Elbert (1986) und Parker (1922) als *Nomen* klassifizierten Wortes in Artikel-Syntagma und „Nominalisierungs-konstruktion“ analysierbar ist, vgl.:

(13a) *ho‘ōho a‘e-la ia me [nā nānaina huhū]*
 ausrufen DIR-DIST 3SG PREP PL Aussehen wütend

iā Ivanahō
 PREP Ivanhoe

‘Er rief mit wütendem Aussehen/Gesicht nach Ivanhoe aus.’
 (Antonio 1995 [1871]:524)

(13b) *He nani a mōlale maika‘i [ka nānaina ‘ana]*
 IDF schön und klar gut DEF Aussehen NMLZ

o ua wahi lā
 POSS DEM Ort DIST

‘Das Aussehen dieses Ortes ist schön und klar.’
 (Antonio 1995 [1871]:491)

Pukui und Elbert (1986: 261) und Parker (1922: 453) klassifizieren das Wort *nānaina* als *Nomen*. Wie *kakahiaka* tritt *nānaina* im Korpus ausschließlich in Artikel-Syntagma, wie in Satz (13a), und in „Nominalisierungskonstruktion“, wie in Satz (13b), auf, so dass sich die Auftretensweise von *nānaina* in der Konstruktion mit ‘*ana* syntaktisch-distributional nicht als *Nominalisierung* definieren lässt. Um das Auftreten von *nānaina* in Satz (13b) semantisch als *Nominalisierung* oder *Verbalisierung* zu beschreiben, müsste ein Bedeutungsunterschied bzw. Unterschied in der ontologischen Klasse der Bedeutung zwischen dem Auftreten von *nānaina* in Artikel-Syntagma in Satz (13a) und in „Nominalisierungskonstruktion“ in Satz (13b) interpretiert werden.

Auch wenn es nicht in eindeutiger Weise aus dem Kontext hervorgeht, ist es bei *nānaina* in Satz (13b) ähnlich zu *kakahiaka* in Satz (12b) (S. 152) vielleicht noch möglich, die Bedeutung als das Werden oder die Veränderung des Aussehens des Ortes zu interpretieren. Es lassen sich aber auch Beispiele finden, bei denen ich keinen Spielraum mehr sehe für eine semantische Interpretation, aus der heraus die Bedeutung eines Wortes im Auftreten in Artikel-Syntagma im Vergleich zum Auftreten in „Nominalisierungskonstruktion“ als die Veränderung einer *nominalen* zu einer *verbalen* Bedeutung bestimmt werden könnte, vgl.:

- (14a) *puni me [nā pōhaku nunui]*³⁸
 umgeben PREP PL Stein groß
 ‘umgeben von großen Steinen.’
 (Silva 2003:8-15)

- (14b) *e ‘ike aku ai i ka ho‘oma‘ema‘e ‘ia*
 INF sehen DIR ANAPH PREP DEF säubern PASS

mai, me [ka pōhaku ‘ana] o kahakai, ‘o
 DIR PREP DEF Stein NMLZ POSS Strand TOP

*ia ka sanapepa o ia wā*³⁹
 3SG DEF Sandpapier POSS 3SG Zeit

‘Man sah das Gesäubert-Werden mit dem Stein des Strandes; er war das Sandpapier dieser Zeit.’
 (Silva 2002b: 175)

Im Auftreten von *pōhaku* in „Nominalisierungskonstruktion“ in Satz (14b) lässt sich kein Bedeutungsunterschied zum Auftreten von *pōhaku* in Artikel-Syntagma in Satz (14a) erkennen. Insbesondere unter Beachtung des anaphorischen Bezuges des Pronomens *ia* auf *pōhaku* und der Aussage, dass der Stein das Sandpapier war, sehe ich keine Möglichkeit, *pōhaku* in *ka pōhaku ‘ana* als semantisch *verbal* zu interpretieren.

An Wörtern wie *kakahiaka*, *nānaina* und *pōhaku*, die im Korpus ausschließlich in Artikel-Syntagma und „Nominalisierungskonstruktion“ auftreten, zeigt sich dann erstens, dass aus der Beobachtung der Verteilung von Wörtern die Beschreibung der Konstruktion mit *‘ana* als *Nominalisierung* nicht gerechtfertigt ist. Zweitens kann auch unter Berücksichtigung der Bedeutung dieser Wörter entweder nicht von *Nominalisierung* gesprochen werden, weil sich wie bei *pōhaku* kein Bedeutungsunterschied zwischen dem Auftreten in Artikel-Syntagma und dem Auftreten in der Konstruktion mit *‘ana* zeigen lässt, oder wie bei *kakahiaka* der Bedeutungsunterschied von Elbert und Pukui (1979: 80) als Veränderung einer *nominalen* zu einer *verbalen* Bedeutung analysiert wird.

³⁸Originaltext: *puni me na pohaku nunui*.

³⁹Originaltext: *e ike aku ai i ka hoomaemae ia mai, me ka pohaku ana o kahakai, o ia ka sanapepa o ia wa*.

Als nächstes soll die Beschreibung der *syntaktischen Nominalisierung* von Wörtern betrachtet werden, die im Korpus sowohl in Tempus-Aspekt-Modus-Syntagma, „Nominalisierungskonstruktion“ und Artikel-Syntagma auftreten. Ein Beispiel für ein solches Wort ist ‘*ai*, vgl.:

- (15a) [ua ‘*ai* a‘e-la] ‘o ia i kēla maka manō
 PFV essen DIR-DIST TOP 3SG PREP DIST Auge Hai
 ‘Er aß dieses Haiauge.’
 (Hale‘ole 2001 [1863]:27)

- (15b) a mua o [ko Kekūhaupi‘o ‘*ai* ‘ana] i
 und vor POSS POSS Kekūhaupi‘o essen NMLZ PREP

 kēla manō niuhi
 DIST Hai Niuhi

‘und vor Kekūhaupi‘os Essen von diesem Niuhi-Hai’
 (Hale‘ole 2001 [1863]:26-27)

- (15c) eia [ka ‘*ai*] a me ka i‘a o
 es gibt DEF Essen und PREP DEF Fleisch/Fisch POSS

 loko nei⁴⁰
 drinnen PROX

‘Es gibt hier drinnen Essen und Fleisch/Fisch⁴¹.’
 (Silva 2002a:114)

In Satz (15a) tritt ‘*ai* in Tempus-Aspekt-Modus-Syntagma auf, in Satz (15b) in „Nominalisierungskonstruktion“ und in Satz (15c) in Artikel-Syntagma ohne *Nominalisierung* durch ‘*ana*.

Zwar ist zwischen den Auftretensweisen von ‘*ai* in den Sätzen (15b) und (15c) ein semantischer Unterschied erkennbar: Während ‘*ai* in (15b) die Handlung

⁴⁰Originaltext: eia ka ai a me ka ia o loko nei.

⁴¹Im Hawaiischen wird mit ‘*ai* mehrheitlich pflanzliches Essen, mit *i‘a* dagegen tierisches Essen bezeichnet (vgl. Pukui und Elbert 1986: 9). Hier ist aus dem Kontext nicht eindeutig erkennbar, ob es sich spezifisch um Fleisch oder Fisch handelt.

des Essens bezeichnet, bezieht sich *'ai* in (15c) (S. 156) auf das Objekt des Essens.

Aus formaler und distributionalen Perspektive lässt sich jedoch nicht unterscheiden zwischen *'ai* als *Verb*, das für das Auftreten in Tempus-Aspekt-Modus-Syntagma und „Nominalisierungs-konstruktion“ prädisponiert ist und *'ai* als *Nomen*, das für das Auftreten in Artikel-Syntagma prädisponiert ist.

An dieser Stelle wird deutlich, dass das grundlegende Problem der Klassifizierung von Wörtern, die sowohl in Tempus-Aspekt-Modus-Syntagma als auch Artikel-Syntagma auftreten, auch die Frage der *Nominalisierung* betrifft.

Wenn Elbert und Pukui (1979: 43) *Nomen-Verben* als *Verben* definieren, die ohne *Nominalisierung* mit *'ana* in Artikel-Syntagma auftreten im Gegensatz zu *Verben*, die nur *nominalisiert* in Artikel-Syntagma auftreten (vgl. auch Kapitel 4.1.1, S. 112–115), kann die Auftretensweise eines Wortes wie *'ai* nicht erklärt werden.

Wenn Pukui und Elbert (1986: ix) *Nomen-Verben* rein distributional als Wörter definieren, die sowohl in Tempus-Aspekt-Modus-Syntagma als auch in Artikel-Syntagma auftreten und am Beispiel von *'ai* festgehalten werden muss, dass *Nomen-Verben* auch in „Nominalisierungs-konstruktion“ auftreten, kann nicht von *Nominalisierung* gesprochen werden. Weder ein Wort, das sowohl *Verb* als auch *Nomen* ist, noch ein Wort, das keiner *Wortart* angehört, muss *nominalisiert* werden. Die Rede von *Nominalisierung* setzt eben die Unterscheidung von *Verb* und *Nomen* voraus. Wenn aber nicht zwischen *'ai* als *Verb* und *'ai* als *Nomen* differenziert werden kann, ist es auch nicht möglich, *Nominalisierung* anzusetzen, auch wenn ein semantischer Unterschied im Auftreten in „Nominalisierungs-konstruktion“ und Artikel-Syntagma erkennbar ist.

Um also weiterhin von *Nominalisierung* sprechen zu können, müssen Elbert und Pukui (1979) bei einem Wort wie *'ai* Nullderivation ansetzen und semantisch zwischen einem *Verb* *'ai₁* „essen“, das für das Auftreten in Tempus-Aspekt-Modus-Syntagma prädisponiert ist und *nominalisiert* werden kann und einem *Nomen* *'ai₂* „Essen“, das hingegen für das Auftreten in Artikel-Syntagma prädisponiert ist, unterscheiden. Dann wäre jedoch unter der Annahme von Elbert und Pukui (1979: 80), dass auch als *Nomen* klassifizierte Wörter *nominalisiert* werden können, nicht mehr zu entscheiden, ob es sich bei *'ai* im Auftreten in „Nominalisierungs-konstruktion“ um die syntaktisch definierte *Nominalisierung* von *'ai₁* als *Verb*, wie bei *pa'e* in Satz (10b) (S. 133), oder um die semantisch definierte *Nominalisierung* von *'ai₂* als *Nomen*, wie bei *kaka-hiaka* in Satz (12b) (S. 152), handelt.

Ein ähnliches Problem bei der Entscheidung darüber, ob es sich in Zusammenhang mit der Annahme von Nullderivation um die *Nominalisierung* eines *Verbs* oder *Nomens* handelt, zeigt sich, wenn die Bedeutung eines Wortes innerhalb des Auftretens in einem bestimmten Syntagma variieren kann, vgl.:

- (16a) *a me [ka ‘aihue ‘ana] i kēlā wa‘apā*
 und PREP DEF stehlen NMLZ PREP DIST Skiff
 ‘und mit dem Stehlen von dem Skiff.’
 (Desha 1996 [¹1920,1924]:336)

- (16b) *inā ua ho‘āhewa ‘ai kekahi ma muli o*
 wenn PFV verurteilen PASS ART nach POSS

ke kānāwai no [ka ‘aihue]⁴²
 DEF Gesetz POSS DEF Diebstahl

‘wenn derjenige verurteilt ist nach dem Gesetz über Diebstahl.’
 (Silva 2002b:220)

- (16c) *ka palekaua ikaika nou e hiki ‘ole*
 DEF Schutzschild stark 2SG.POSS INTV können NEG

ai i [nā ‘aihue] ke ki‘i mai
 ANAPH PREP PL Dieb INF angreifen DIR

‘das starke Schutzschild für dich, damit Diebe nicht angreifen können.’
 (Antonio 1995 [¹1871]:168)

Pukui und Elbert (1986: 10) klassifizieren *‘aihue* als *transitives Nomen-Verb* mit der Bedeutung „to steal, thief, theft“. Im Wörterbuch von Parker (1922: 35) finden sich drei Einträge: *‘aihue* „thievish“ klassifiziert als *Adjektiv*, *‘aihue* „thief“ klassifiziert als *Nomen* und *‘aihue* „to thieve, steal“ klassifiziert als *Verb*.

Um über die semantische Differenzierung von *Verb* und *Nomen* an der *Nominalisierung* von *‘aihue* im Auftreten in der Konstruktion in Satz (16a) festzuhalten, müssen erstens zwei *Nomen* angesetzt werden: *‘aihue* als *Nomen* mit

⁴²Originaltext: *ina ua hoahewa ai kekahi ma muli o ke kanawai no ka aihue.*

der Bedeutung „Diebstahl“ im Auftreten in Artikel-Syntagma in Satz (16b) (S. 158) und ‘*aihue* als *Nomen* mit der Bedeutung „Dieb“ im Auftreten in Artikel-Syntagma in Satz (16c) (S. 158). Zweitens muss von diesen beiden ‘*aihue* als *Verb* mit der Bedeutung „stehlen“ und prädisponiert für das Auftreten in Tempus-Aspekt-Modus-Syntagma unterschieden werden. Drittens muss gezeigt werden, dass ‘*aihue* „Stehlen“ im Auftreten in „Nominalisierungs-konstruktion“ die *Nominalisierung* von ‘*aihue* als *Verb* darstellt und sich semantisch als *nominalisiertes Verb* von ‘*aihue* „Diebstahl“ als *Nomen Actionis* unterscheidet. Es muss also nicht nur die Derivationsrichtung zwischen ‘*aihue*₁ „stehlen“ als *Verb*, ‘*aihue*₂ „Diebstahl“ als *Nomen Actionis* und ‘*aihue*₃ „Dieb“ als *Nomen Agentis* und eventuell noch ‘*aihue*₄ „diebisch“ als *Adjektiv* angegeben werden können, sondern auch, dass ‘*aihue*₁ „stehlen“ als *Verb* in „Nominalisierungskonstruktion“ auftritt.

Zudem gilt folgendes: Wenn die Beschreibung von *Nominalisierung*, wie von Elbert und Pukui (1979: 80) bei *kakahiaka* in Satz (12b) (S. 152) angenommen und auf ‘*ai* in Satz (16b) (S. 158) übertragbar, nicht nur syntaktisch als das Auftreten eines *Verbs* in Artikel-Syntagma, sondern auch semantisch als Bedeutungsveränderung eines *Nomens* im Auftreten in Artikel-Syntagma möglich ist und sich an ‘*aihue* zeigt, dass die Bedeutung eines Wortes im Auftreten in Artikel-Syntagma auch ohne *Nominalisierung* nicht identisch sein muss, fehlt zumindest eine Erklärung, warum manche *Nomen* nur mittels *Nominalisierung* die Bedeutung im Auftreten in Artikel-Syntagma ändern.

Im Grunde zeigt sich bei dem Versuch der Beschreibung von *Nominalisierung* von Wörtern wie ‘*aihue* oder ‘*ai* das Problem der Unterscheidung zwischen *Verb* und *Nomen* im Hawaiischen: Es ist weder über die Derivationsrichtung zwischen den postulierten *Wortarten* entscheidbar, noch kann bestimmt werden, welcher *Wortart* ein Wort im Auftreten in der Konstruktion mit ‘*ana* angehört. Ebenso wenig kann andersherum anhand des Auftretens eines Wortes in „Nominalisierungskonstruktion“ gezeigt werden, dass es sich um ein *Verb* handelt bzw. dass z.B. ‘*ai*₁ als *Verb* von ‘*ai*₂ als *Nomen* zu trennen ist. Dabei hilft auch der Versuch, die Unterscheidung von *Verb* und *Nomen* semantisch zu begründen, nicht weiter – insbesondere dann nicht, wenn man beachtet, dass sich wie bei Wörtern, die im Korpus ausschließlich in Artikel-Syntagma und „Nominalisierungskonstruktion“ auftreten, auch bei von Elbert und Pukui (1979) als *Nomen-Verb* klassifizierten Wörtern nicht in jedem Fall in klarer Weise zwischen der Bedeutung im Auftreten in „Nominalisierungskonstruktion“, Artikel-Syntagma und Tempus-Aspekt-Modus-Syntagma unterscheiden

lässt. Ohne die semantische Differenzierung der Auftretensweise eines Wortes wird aber die Entscheidung darüber, ob ein semantisch bestimmtes *Verb* oder von ein semantisch bestimmtes *Nomen* in „Nominalisierungskonstruktion“ auftritt, zusätzlich erschwert, vgl.:

- (17a) *a* [**ho‘āo** *a‘e-la*] *me ka wahine nei*
 und heiraten DIR-DIST PREP DEF Frau PROX
 ‘und [er] heiratete die Frau.’
 (Kaiholo 1998 [1961]:74-75)

- (17b) *mahope o* [*ko Laieikawai ho‘āo ‘ana*] *me*
 nach POSS POSS Laieikawai heiraten NMLZ PREP

Kaonohiokala
 Kaonohiokala

‘Nach dem Heiraten/der Heirat von Laieikawai mit Kaonohiokala.’
 (Hale‘ole 2001 [1863]:167)

- (17c) *e ho‘omākaukau ai no* [*ka ho‘āo*] *o*
 INTV vorbereiten ANAPH PREP DEF Heirat POSS

‘olua
 2DU

‘um auf die Heirat von euch vorzubereiten.’
 (Hale‘ole 2001 [1863]:88)

Pukui und Elbert (1986: 26) klassifizieren *ho‘āo* als *Nomen-Verb*, Parker (1922: 133) klassifiziert *ho‘āo* als *Nomen*. Genauer gesagt sind Pukui und Elbert (1986) etwas unklar in der Klassifizierung: Unter dem Eintrag *ho‘āo* ist die Angabe als *Nomen* mit „marriage“ zu finden (vgl. Pukui und Elbert 1986: 74), allerdings mit dem Verweis auf *‘ao*, unter dessen Eintrag dann *ho‘āo* mit „to marry, marriage“ angegeben ist (vgl. Pukui und Elbert (1986: 26). Bei Parker (1922) ist unklar, warum in der Überarbeitung des Wörterbuchs von Andrews (1865) die Klassifizierung von *ho‘āo* als *Verb* von Andrews (1865: 167) gestrichen wurde. Die Unsicherheit über den Gebrauch von *ho‘āo* in den Wörterbüchern von Pukui und Elbert (1986) und Parker (1922) mag an dem

Umstand liegen, dass es sich bei *ho‘āo* um einen wenig gebräuchlichen Begriff für einen veralteten hawaiischen Hochzeitsbrauch handelt (vgl. Andrews 1865: 167, Parker 1922: 133, Pukui und Elbert 1986: 26), macht aber auch deutlich, dass ohne formale Differenzierung zwischen *marry* und *marriage* die Auftretensweise eines Wortes wie *ho‘āo* rein aus der Semantik schwierig zu bestimmen ist.

Um das Auftreten von *ho‘āo* in Satz (17b) (S. 160) in der Konstruktion mit *‘ana* als *Nominalisierung* zu bezeichnen, muss eine semantische Differenzierung zwischen *ho‘āo*₁ als *Verb* in Satz (17a) (S. 160) und *ho‘āo*₂ als *Nomen* in Satz (17c) (S. 160) vorausgehen, um dann zu zeigen, dass die Bedeutung von *ho‘āo* in „Nominalisierungs-konstruktion“ der Bedeutung von *ho‘āo* im Auftreten in Tempus-Aspekt-Modus-Syntagma entspricht.

Anders als in Satz (17c) (S. 160) scheint *ho‘āo* in Satz (17b) (S. 160) zwar eher den Prozess des Heiratens oder die Handlung des Heiratens von Kaonohiokala durch den Agens Laieikawai auszudrücken, ähnlich der Handlung des Essens von dem Niuhi-Hai durch Kekūhaupi‘o in Satz (15b) (S. 156). Eindeutig ist der Bedeutungsunterschied zwischen dem Auftreten von *ho‘āo* in „Nominalisierungs-konstruktion“ und Artikel-Syntagma jedoch nicht. Die Frage ist hier letztlich, ob über die Gegenüberstellung der semantisch-ontologischen Klassen EREIGNIS und OBJEKT klar zu zeigen ist, dass auch zwei Wörter *ho‘āo*₁ und *ho‘āo*₂ unterscheidbar sind und damit die Richtung der *Nominalisierung* zwar nicht syntaktisch-distributional aber semantisch bestimmt von *ho‘āo*₁ als *Verb* und nicht von *ho‘āo*₂ als *Nomen* ausgeht oder ob sich nur festhalten lässt, dass *ho‘āo* in Tempus-Aspekt-Modus-Syntagma, in Artikel-Syntagma und „Nominalisierungs-konstruktion“ auftritt und die Bedeutung eines Wortes *ho‘āo* lexikalisch nicht einfach über eine semantisch-ontologische Klasse definiert ist, unabhängig von einer möglichen semantischen Profilierung im Auftreten in einem Syntagma.

Das Problem bei dem Versuch, *Nominalisierung* über die semantische Differenzierung bzw. die semantisch-ontologische Klassifizierung von *Verb* und *Nomen* zu beschreiben, wird noch deutlicher an folgendem Beispiel, das Elbert und Pukui (1979: 80) selbst für *Nominalisierung* von *Verben* anführen, vgl.:

- (18) *ka hele ‘āwīwī ‘ana mai*
 DEF come quick NMLZ here
 ‘the quick coming’

Pukui und Elbert (1986: 64) klassifizieren *hele* nun aber als *Nomen-Verb* mit der Bedeutung „come, go, going“. Das bedeutet, dass *hele* nicht ausschließlich in Tempus-Aspekt-Modus-Syntagma, z.B. *ua hele aku* „[He] went away“ (Elbert und Pukui 1979: 48) und „Nominalisierungs-konstruktion“, wie in Satz (18) (S. 161), sondern auch in Artikel-Syntagma auftreten kann, vgl.:

(19)	<i>a</i>	<i>ho‘omaākaukau</i>	<i>nō</i>	<i>ho‘i</i>	<i>no</i>	[<i>ka</i>	<i>hele</i>	<i>aku</i>]
	und	vorbereiten	INTSF	INTSF	PREP	DEF	gehen	DIR
	<i>e</i>	<i>mōhai</i>	<i>i</i>	<i>kēlā</i>	<i>maka</i>	<i>‘ākau</i>		
	INTV	opfern	PREP	DIST	Auge	rechts		
	<i>i</i>	<i>ka</i>	<i>heiau</i>					
	PREP	DEF	Heiau					

‘und sich vorzubereiten auf das Gehen zum Heiau, um das rechte Auge zu opfern.’

(Desha 1996 [¹1920,1924]:26)

Die fehlende Möglichkeit, einer syntakisch-distributionalen Beschreibung der *Nominalisierung* von *hele* kann nicht über eine semantische Unterscheidung von *hele*₁ als *Verb* und *hele*₂ als *Nomen* kompensiert werden. Es lässt sich weder eine lexikalisierte *verbale* Bedeutung „gehen“ von *hele* in Abgrenzung zu einer lexikalisierten *nominalen* Bedeutung „gehen“ von *hele* analysieren, noch ist die Bedeutung von *hele* „gehen“ im Auftreten in der Konstruktion mit *‘ana* dann *hele* als semantisch definiertem *Verb* zuzuweisen. Die Auftretensweisen von *hele* in Tempus-Aspekt-Modus-Syntagma, Artikel-Syntagma und „Nominalisierungs-konstruktion“ können schlicht nicht über eine jeweilige semantisch-ontologische Klassifizierung von *hele* differenziert werden und damit besteht auch keine Möglichkeit, semantisch *Nominalisierung* zu beschreiben. Zudem findet sich auch noch das von *hele* gebildete Derivat *hele-na*, das Pukui und Elbert (1986: 65) als *Nomen* klassifizieren mit der Bedeutung „going“.

Nachdem zuerst die Beschreibung von *lexikalischer Nominalisierung* und danach die Beschreibung von *syntaktischer Nominalisierung* problematisiert wurde, kann nun zusammenfassend der Beschreibungsversuch von *Nominalisierung* im Hawaiischen als Ganzes betrachtet werden (vgl. Abbildung 24, S. 163).

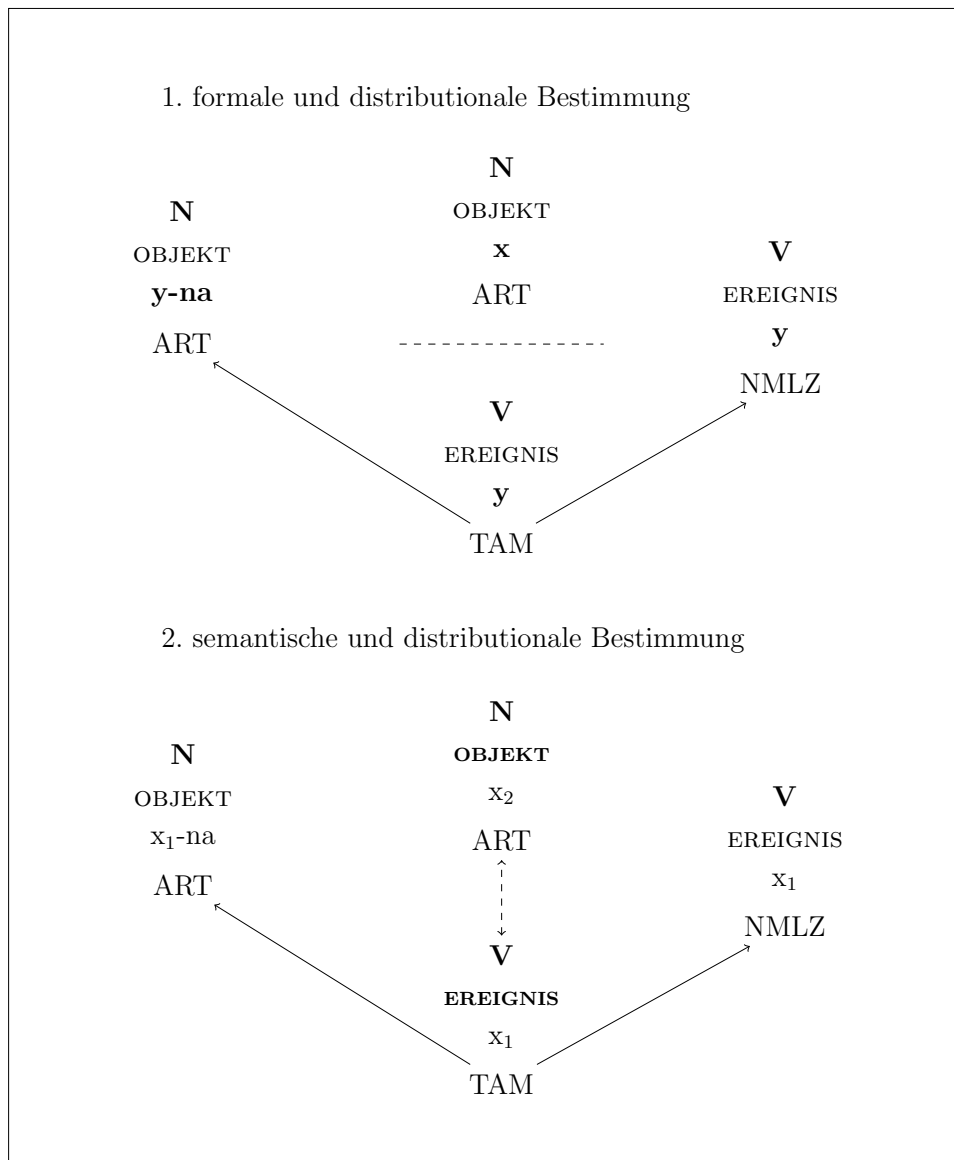


Abbildung 24: Zusammenfassung Beschreibungsversuch *Nominalisierung*

Im ersten Teil von Abbildung 24 ist der grundlegende Versuch der Beschreibung von *lexikalischer* und *syntaktischer Nominalisierung* festgehalten.

Die *lexikalische Nominalisierung* lässt sich formal und distributional als kategorienverändernde Derivation bestimmen: Die Basis ist prädisponiert für das Auftreten in Tempus-Aspekt-Modus-Syntagma, das Derivat ist mit *-na* oder Alternanten suffigiert und prädisponiert für das Auftreten in Artikel-Syntagma. Unter der Voraussetzung, dass die Bedeutung eines Wortes im Auftreten in einem bestimmten Syntagma immer in klarer Weise der ontologischen Klasse

EREIGNIS oder OBJEKT zuzuordnen ist, kann zudem gezeigt werden, dass die Bedeutung einer Basis der Klasse EREIGNIS und die Bedeutung eines Derivats der Klasse OBJEKT angehört.

Die *syntaktische Nominalisierung* kann als *Nominalisierung* eines *Verbs* über die Distribution eines Wortes bestimmt werden: Ein *Verb* ist prädisponiert für das Auftreten in Tempus-Aspekt-Modus-Syntagma und muss im Auftreten in Artikel Syntagma mit *‘ana nominalisiert* werden. Zudem ist die Bedeutung der ontologischen Klasse EREIGNIS zuzuordnen.

Im zweiten Teil von Abbildung 24 (S. 163) ist dann die Übertragung dieser Beschreibung festgehalten auf solche Fälle, in denen ein Wort ohne formale Differenzierung sowohl in Tempus-Aspekt-Modus-Syntagma als auch Artikel-Syntagma auftritt.

Wenn das Wort, das die Basis eines Derivats bildet, nicht ausschließlich in Tempus-Aspekt-Modus-Syntagma auftritt, muss bei diesem, um weiterhin an *lexikalischer Nominalisierung* festzuhalten, über die Unterscheidung von semantischen Klassen und über Nullderivation *Verb* und *Nomen* geschieden werden.

Wenn ein Wort, das in der Konstruktion mit *‘ana* auftritt, auch in Artikel-Syntagma auftritt, muss bei diesem ebenso über die Unterscheidung von semantischen Klassen und über Nullderivation *Verb* und *Nomen* geschieden werden, um weiterhin an *syntaktischer Nominalisierung* festhalten zu können.

Neben der Frage der Bestimmung der lexikalisierten Bedeutung von Wörtern und der Derivationsrichtung zwischen postulierten *Wortarten* kann dieser Beschreibung von *Nominalisierung* vor allem entgegengehalten werden, dass dadurch weder aus formaler und distributionaler noch aus semantischer Perspektive die Beobachtungen im Korpus bzw. die identifizierten Auftretensweisen *nominalisierter* Wörter im Hawaiischen widergespiegelt werden (vgl. Abbildung 25, S. 165).

In Abbildung 25 (S. 165) sind acht der Auftretensweisen von *nominalisierten* Wörtern festgehalten, die im Korpus und durch die Angaben in der Grammatik von Elbert und Pukui (1979) bzw. dem Vergleich mit den Wörterbüchern von Pukui und Elbert (1986) und Parker (1922) für das Hawaiische identifiziert werden konnten. Es ist weder möglich, von *lexikalischer Nominalisierung* zu sprechen, wenn *lexikalisch nominalisierte* Wörter nicht ausschließlich in Artikel-Syntagma, sondern ebenso in Tempus-Aspekt-Modus-Syntagma und „Nominalisierungskonstruktion“ und dessen Basen nicht ausschließlich in Tempus-Aspekt-Modus-Syntagma, sondern ebenso (ausschließlich) in Artikel-

E x TAM	E/O x-na ART	E x TAM	E x-na TAM	E x TAM	O x-na ART	O x ART	O x-na ART
E x NMLZ	E/O x-na NMLZ	E/O x ART	E/O x-na ART	E x NMLZ	E x-na TAM	O x ART	O x-na ART
O x ART	O x-na ART	E x TAM	E x-na ART	E x NMLZ	E x-na NMLZ	E x TAM	E x-na TAM
		E/O x NMLZ	E x-na NMLZ	O x ART	O x NMLZ	E x NMLZ	E x-na NMLZ

Abbildung 25: beobachtete Auftretensweisen *nominalisierter* Wörter

Syntagma auftreten, noch kann *syntaktische Nominalisierung* beschrieben werden, wenn Wörter, die in „Nominalisierungs-konstruktion“ auftreten, nicht als *Verben* bestimmbar sind, weil sie nicht nur in Tempus-Aspekt-Modus-Syntagma, sondern auch in Artikel-Syntagma auftreten oder sogar *lexikalisch nominalisierte* Wörter darstellen können. Bezüglich der Bedeutung von Wörtern scheint es schließlich nicht immer möglich, eindeutig über die ontologische Klasse zu entscheiden, die ontologischen Klassen entsprechend der Auftretensweisen eines Wortes voneinander zu unterscheiden oder systematisch die Bedeutung von *lexikalisch nominalisierten* Wörtern in Zusammenhang mit der ontologischen Klasse OBJEKT und die Bedeutung von *syntaktisch nominalisierten* Wörtern in Zusammenhang mit der ontologischen Klasse EREIGNIS zu stellen. Letztlich ergeben sich aus dem Versuch der semantischen Beschreibung von *Nominalisierung* auch Probleme für die Annahme von über ontologische Klassen definierter *verbaler* und *nominaler* Semantik (siehe auch Kapitel 4.3, S. 196–215).

Für die eingangs gestellte Frage, ob bestimmte Wörter für das Auftreten in „Nominalisierungs-konstruktion“ prädisponiert sind und sich damit anhand des Auftretens von Wörtern in „Nominalisierungs-konstruktion“ Distributionsklassen im Hawaiischen bestimmen lassen, muss weiter festgehalten werden, dass

jedes Wort, unabhängig von seiner Auftretensweise in anderen Syntagmen potentiell in „Nominalisierungs-konstruktion“ auftreten kann. Das führt dazu, dass die von Elbert und Pukui (1979: 43 und passim) vorgeschlagenen Distributionsklassen, über die *Wortarten* bestimmt sein sollen, erneut korrigiert werden müssen, vgl.:

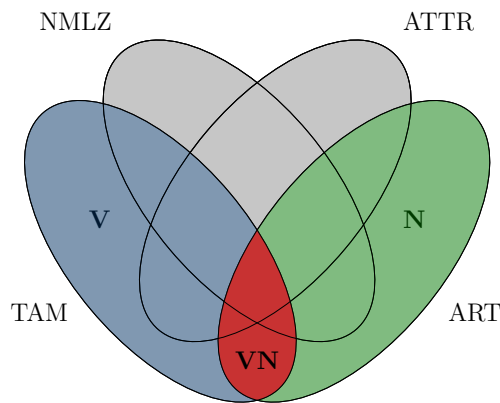


Abbildung 26: Endgültig aus Distributionsklassen gebildete Kategorien

In Abbildung 26 ist wieder in einem Venn-Diagramm dargestellt, welche Distributionsklassen Elbert und Pukui (1979) bei der Kategorisierung von *Wortarten* theoretisch aus der Verteilung von Wörtern im Korpus bilden. Nachdem dargestellt wurde, dass Elbert und Pukui (1979) davon ausgehen, dass nicht nur *Verben*, sondern auch *Nomen* in attributivem Slot (vgl. Kapitel 4.1.2, S. 128–131) und in „Nominalisierungs-konstruktion“ auftreten können, müssen entsprechend aus distributionaler Perspektive *Wortarten* folgendermaßen definiert werden: *Verben* als Wörter, die in allen Syntagmen auftreten außer Artikel-Syntaxma, *Nomen* als Wörter, die in allen Syntagmen auftreten außer in Tempus-Artikel-Syntaxma und *Nomen-Verben* als Wörter, die in allen Syntagmen auftreten.

Im Kern ist diese Definition von *Wortarten* als Distributionsklassen nicht überzeugend, weil die Trennung von *Verb* und *Nomen* nicht aufrechterhal-

ten werden kann, wenn mit dem jeweiligen Kategorienetikett die potentielle Auftretensweise eines Wortes bezeichnet wird. Die Klassenbildung beruht auf der Verallgemeinerung beobachteter Auftretensweisen von Wörtern als voneinander unterscheidbare potentielle Auftretensweisen ohne Differenzierung des tatsächlich beobachteten Auftretens eines jeden einzelnen Wortes.

Wenn man nun aber den Schluss zieht, dass erstens, weil es Wörter im Hawaiischen gibt, die in Tempus-Aspekt-Modus-Syntagma und attributivem Slot oder in Tempus-Aspekt-Modus-Syntagma und „Nominalisierungskonstruktion“ auftreten, potentiell jedes Wort, das in Tempus-Aspekt-Modus-Syntagma auftritt, ebenso in attributivem Slot und „Nominalisierungskonstruktion“ auftreten kann und so *Verben* definieren möchte, wie es Elbert und Pukui (1979) tun, und dass zweitens, weil es Wörter gibt, die in Artikel-Syntagma und attributivem Slot oder in Artikel-Syntagma und „Nominalisierungskonstruktion“ auftreten, potentiell jedes Wort, das in Artikel-Syntagma auftritt, ebenso in attributivem Slot und „Nominalisierungskonstruktion“ auftreten kann, und so *Nomen* definieren möchte, dann sehe ich keinen Grund, nicht genauso zu sagen, dass drittens, weil es Wörter gibt, die in Tempus-Aspekt-Modus-Syntagma und Artikel-Syntagma auftreten, potentiell jedes Wort, das in Tempus-Aspekt-Modus-Syntagma auftritt ebenso in Artikel-Syntagma auftreten kann und umgekehrt.

An Distributionsklassen im Hawaiischen lässt sich also nur dann festhalten, wenn diese auch formal differenziert werden können. Zumindest in dem Versuch der Beschreibung von *lexikalischer Nominalisierung* als kategorienverändernde Derivation war es nicht möglich, über die derivationale Morphologie Distributionsklassen als Formklassen zu bestimmen.

Neben kategorienverändernder Derivation versuchen Elbert und Pukui (1979: 68–89 und *passim*) nun aber auch derivationsmorphologische Prozesse zu beschreiben, die ausschließlich innerhalb einer angenommenen Kategorie *Verb* auftreten, das heißt, über die theoretisch *Verben* formal als Distributionsklasse bestimmbar sind.

Im nächsten Abschnitt soll überprüft werden, ob diese Annahme von Elbert und Pukui (1979) einer genaueren Analyse standhält.

4.1.4 *Verbale Derivationsmorphologie*

Ähnlich wie bei der *lexikalischen Nominalisierung* ist bezüglich der Unterscheidung von *Wortarten* im Hawaiischen die grundlegende Frage bei der Annahme

einer *verbalen* Morphologie, ob sich derivationsmorphologisch Distributionsklassen bestimmen lassen, das heißt, ob Wörter, die bestimmte Derivationsmorpheme tragen auch nur in bestimmten Syntagmen auftreten. Anders als bei der *lexikalischen Nominalisierung* geht es jedoch nicht darum, zu untersuchen, ob die Ableitung eines *Nomens* von einem *Verb* gezeigt werden kann, sondern ob ausschließlich innerhalb einer Kategorie *Verb* bestimmte derivationsmorphologische Prozesse auftreten. Die zu überprüfende Annahme besagt also, dass sich im Hawaiischen Basis-Derivat-Paare identifizieren lassen, bei denen sowohl die Basis als auch das Derivat für das Auftreten in Tempus-Aspekt-Modus-Syntagma prädisponiert ist.

In Tabelle 22 sind diejenigen Affixe zusammengetragen, die Elbert und Pukui (1979: 68–89) zur *verbalen* Morphologie zählen, die also ausschließlich in einer angenommenen Kategorie *Verb* auftreten sollen, vgl.:

	Funktion	Form	Kategorie	
			Basis	Derivat
Präfix	Kausativierung	<i>ha-</i> , <i>ha'a</i> , <i>hā-</i> <i>'ā-</i> , <i>'a-</i> , <i>kā-</i> <i>'ō-</i> , <i>hai-</i>	Verb	kausatives Verb
Suffix	Passivisierung	<i>-a</i> , <i>-hia</i> , <i>-kia</i> , <i>-lia</i> , <i>-mia</i> , <i>-nia</i> , <i>-ia</i>	aktives Verb	passives Verb
	Transitivierung	<i>-a'i</i> , <i>-na'i</i> , <i>-ka'i</i> , <i>-la'i</i> , <i>-hi</i> , <i>-ki</i> , <i>-'i</i> , <i>-i</i>	intransitives Verb	transitives Verb

Tabelle 22: Überblick *verbale* Derivationsmorphologie

Beim Überblick in Tabelle 22 orientiere ich mich an der Beschreibung der Form und Funktion der verschiedenen Präfixe und Suffixe von Elbert und Pukui (1979: 68 - 89). In der dritten und vierten Spalte von Tabelle 22 sind die Kategorienetiketten angegeben, die theoretisch aus der Terminologie von Elbert und Pukui (1979: 68–89) für die jeweilige Basis und das jeweilige Derivat folgen. Im Kern versuchen Elbert und Pukui (1979: 68–89) über die Derivationsmorphologie semantische Subklassen von *Verben* zu bestimmen.

Da hier die Frage des Zusammenhangs von Derivations- und Distributionsklassen im Vordergrund steht, soll im Folgenden nur kurz auf die Funktion der Affixe aus Tabelle 22 eingegangen werden, ohne eine tiefergehende Analyse zu bieten.

Die Formen *ha-*, *ha'a*, *hā-*, *'ā-*, *'a-*, *kā-*, *'ō-* und *hai-* fassen Elbert und Pukui (1979: 68–71) als Allomorphe unter der Funktion „causative / simulative meaning“ (Elbert und Pukui 1979: 68) zusammen. Regeln, in welchem lautlichen Umfeld eine bestimmte Form auftritt, lassen sich nicht angeben (vgl. Elbert und Pukui 1979: 68). Elbert und Pukui (1979: 68–71) nennen einige, in der Bedeutung teilweise stark voneinander abweichende Beispiele für Derivate mit *hā-* oder anderen Morphemalternanten und der jeweiligen Basis, ohne jedoch die Funktion der *Kausativierung* genau zu beschreiben.

Im Deutschen können kausative Verben semantisch als Wörter beschrieben werden, die „[...] ein Bewirken der im Grundwort genannten Tätigkeit ausdrücken.“ (Henzen 1957: 212). Ein Beispiel für ein kausatives Verb im Deutschen ist das von dem Grundwort *fallen* abgeleitete Wort *fällen* (vgl. Fleischer und Barz 2012: 375).

Unter den von Elbert und Pukui (1979: 68–71) genannten Beispielen finden sich dann auch Basis-Derivat-Paare, bei denen eine ähnliche Bedeutungsbeziehung wie bei *fallen* und *fällen* besteht, z.B. *liu* „to turn“ und *hāliu* „to cause to turn“.

In anderen Fällen scheint eher eine Ähnlichkeitsbeziehung zwischen der Bedeutung von Derivat und Basis oder eine Graduierung der Bedeutung der Basis ausgedrückt zu werden, z.B. *'ele* „black“ und *hā-'ele* „blackish“ (Elbert und Pukui 1979: 68).

Bei vielen der von Elbert und Pukui (1979: 68–71) aufgeführten Beispiele ist die Veränderung der Bedeutung von Basis zu Derivat allerdings nicht eindeutig, oder es kann gar kein semantischer Unterschied zwischen Basis und Derivat festgestellt werden,⁴³ z.B. *ehu* „vigorous“ und *hā-ehu* „vigorous“ (Elbert und Pukui 1979: 68), *wili* „to twist“ und *'ā-wili* „to mix“ (Elbert und Pukui 1979: 69) oder *hea* „to call“ und *kā-hea* „to call“ (Elbert und Pukui 1979: 69).

Die Formen *-ia*, *-hia*, *-kia*, *-lia*, *-mia*, *-nia*, fassen Elbert und Pukui (1979: 83–85) unter der Funktion der Bildung *passiver Verben* zusammen, wobei in Einzelfällen die Bedeutung von Derivaten eher als „imperative“ (Elbert und Pukui 1979: 85) analysierbar sein soll. Ähnlich wie bei der *Nominalisierung* findet

⁴³Zieht man hier wieder den Vergleich zum Deutschen, in dem auch nicht immer eine klare semantische Beziehung zwischen Grundwort und Kausativum erkennbar ist (vgl. Fleischer und Barz 2012: 376) und in dem einem Präfix nicht ausschließlich eine Funktion zukommt, mag dies wenig überraschen. Elbert und Pukui (1979: 68–71) bleiben jedoch undeutlich darin, ob auch zwischen den allomorphen Formen Bedeutungsunterschiede bestehen. Damit stellt sich die Frage, ob diese überhaupt funktional unter ein Morphem gefasst werden können. Das kann hier nicht abschließend geklärt werden.

sich mit *'ia* neben den Suffixen zusätzlich eine ungebundene Form, die sich wie die anderen Allomorphe auf eine protopolynesische Form **(C)ia* zurückführen lässt (vgl. Hale 1968, Chung 1974, Sanders 1991) und von der auch in anderen polynesischen Sprachen Reflexe aufzufinden sind (vgl. z.B. Buse 1963: 633 zum Rarotongaïschen oder Harlow 2006: 115–120 zum Maorischen). Während die ungebundene Form im Hawaiïschen wie in anderen polynesischen Sprachen nicht ausschließlich in Tempus-Aspekt-Modus-Syntagma auftritt, sondern ebenso in Artikel-Syntagma und „Nominalisierungs-konstruktion“ (vgl. Elbert und Pukui 1979: 41 und passim, Buse 1963: 633 zum Rarotongaïschen) und damit keiner *Wortart* zugeordnet werden kann, sehen Elbert und Pukui (1979: 83–86) suffigierte Wörter als *Verben* an. Ähnlich wie bei den Kausativen muss auch bei diesen Formen festgehalten werden, dass die Bestimmung der Funktion des *Passivs* eher oberflächlich scheint und eine eindeutige Bedeutungsunterscheidung von Basis und Derivat in vielen Fällen nicht möglich ist (vgl. Elbert und Pukui 1979: 85–86, vgl. auch Churchward 1928 und Cook 1996 zur Diskussion des Passivs im Samoanischen und anderen polynesischen Sprachen). Die Formen *-ka'i*, *-la'i*, *-hi*, *-ki*, *-i*, und *-i* beschreiben Elbert und Pukui (1979: 86–88) zwar als Allomorphe des Transitiv-Morphems. Zugleich bemerken Elbert und Pukui (1979: 86) aber einschränkend, dass nicht nur *intransitive Verben* und *stative Verben*, sondern auch bereits *transitive Verben* mit einer der Formen suffigiert werden können. Ebenso ziehen Elbert und Pukui (1979: 86) einen Vergleich zur Analyse von Churchward (1953) zu den entsprechenden Suffixen im Tongaïschen. Churchward (1953) schreibt zu diesen, dass die eigentliche Funktion nicht *Transitivierung* ist, sondern „to form new verbs with new meanings – meanings which in some cases happen to be transitive.“ (Churchward 1953: 242).

Insgesamt scheint also die Funktion der verschiedenen Affixe nicht gut beschrieben. Entscheidend ist hier aber, dass Elbert und Pukui (1979: 86–88) unabhängig von der genauen Funktion der Suffixe davon ausgehen, Basis und Derivat seien jeweils als *Verb* klassifiziert. Der Bestimmung von *Wortarten* über die Derivationsmorphologie liegt die Annahme zugrunde, dass beispielsweise ein Wort mit Kausativpräfix *hā-* wie *hā-liu* als *Verb* identifiziert werden kann, weil es genauso wie die Basis *liu* für das Auftreten in Tempus-Aspekt-Modus-Syntagma prädisponiert ist und die Bedeutung der ontologischen Klasse EREIGNIS zuzuordnen ist, wobei genau genommen die Bildung von *hā-liu* aus *liu* nur dann als nicht-kategorienverändernde Derivation analysierbar ist, wenn auch die Prädisposition von *liu* für das Auftreten in Tempus-Aspekt-

Modus-Syntagma vorliegt. Im Gegensatz zur Basis kann bei einem Derivat wie *hā-liu* aber in jedem Fall die Prädisposition für das Auftreten in Tempus-Aspekt-Modus-Syntagma, ähnlich wie bei Flexionsklassen, auch formal über die Morphologie gezeigt werden.

Betrachtet man zunächst nur die Verteilung von *verbalen* Derivaten auf die verschiedenen Kombinationen von Syntagmen im Korpus, lässt sich feststellen, dass diese nicht ausschließlich als *Verben* klassifiziert werden können, vgl.:

Syntagma	Präfix	Suffix	
	caus-	-pass	-trans
TAM	4	4	1
ART	4	0	0
ATTR	9	1	1
NMLZ	0	0	0
Σ	17	5	2
TAM \wedge ART	11	7	0
TAM \wedge ATTR	12	2	0
TAM \wedge NMLZ	2	3	1
ART \wedge ATTR	27	0	0
ART \wedge NMLZ	1	0	0
ATTR \wedge NMLZ	0	0	0
Σ	53	1	1
TAM \wedge ART \wedge ATTR	23	19	2
TAM \wedge ART \wedge NMLZ	8	2	0
TAM \wedge ATTR \wedge NMLZ	6	0	0
ART \wedge ATTR \wedge NMLZ	1	0	1
Σ	38	8	3
TAM \wedge ART \wedge ATTR \wedge NMLZ	26	6	1
Σ	134	31	7

Tabelle 23: Verteilung *verbaler* Derivate im Korpus

Tabelle 23 lässt sich entnehmen, dass alle *verbalen* Derivate im Korpus in Artikel-Syntagma auftreten und somit nicht als *Verben* definiert sein können. Die meisten Wörter (Types) mit Kausativ-Präfix (27 von 134) treten sogar ausschließlich in Artikel-Syntagma und attributivem Slot auf. Auch von den im Korpus nur sieben vorkommenden Wörtern (Types) mit Transitiv-Suffix treten vier in Artikel-Syntagma auf.

Bemüht man nun wie zuvor schon bei den *nominalisierten* Derivaten (vgl. Kapitel 4.1.3, S. 140–143) wieder den Vergleich zwischen den Korpusbeobachtun-

gen und der Kategorisierung der entsprechenden Wörter in den Wörterbüchern von Pukui und Elbert (1986) und Parker (1922) und stellt die Kategorien der Basen und Derivate zusammen, kann im Ergebnis nicht an einer *verbalen* Derivationsmorphologie festgehalten werden, vgl.:

Kausativ-Suffix				
Basis	Derivat			Σ
	Nomen	Verb	Nomen-Verb	
Verb	3	8	26	37
Nomen	4	0	4	8
Nomen-Verb	8	2	73	83
Σ	15	10	103	128
Passiv-Suffix				
Basis	Derivat			Σ
	Nomen	Verb	Nomen-Verb	
Verb	0	4	12	16
Nomen	0	0	0	0
Nomen-Verb	0	4	9	13
Σ	0	8	21	29
Transitivierungs-Suffix				
Basis	Derivat			Σ
	Nomen	Verb	Nomen-Verb	
Verb	0	2	3	5
Nomen	0	0	0	0
Nomen-Verb	0	1	1	2
Σ	0	3	4	7

Tabelle 24: Korrigierte Kategorien *verbaler* Basen und *verbaler* Derivate

In oberen Teil von Tabelle 24 sind die Kategorien der Basen und der mit Kausativ-Präfix gebildeten Derivate festgehalten, im mittleren Teil von Tabelle 24 die Kategorien der Basen und mit Passiv-Suffix gebildeten Derivate und im unteren Teil von Tabelle 24 die Kategorien der Basen und mit Transitiv-Suffix gebildeten Derivate. Die Summen der Derivate entsprechen nicht genau den Summen der Derivate in Tabelle 23 (S. 171), da in einzelnen Fällen die Basis eines Derivats im Hawaiischen nicht mehr vorhanden ist. So lässt sich beispielsweise bei *hā'awi* keine Basis mehr auffinden.

Tabelle 24 kann nun entnommen werden, dass bei *Kausativierung*, *Passivierung* und *Transitivierung* weder alle Basen noch alle Derivate als *Verb* klassifiziert werden können und diese somit nicht als derivationsmorphologische Prozesse einer Kategorie *Verb* analysierbar sind.

Im Wesentlichen liegt das Problem des Versuchs von Elbert und Pukui (1979), die Affixe in Tabelle 22 (S. 168) systematisch über die als *verbal* bestimmten Funktionen *Kausativierung*, *Passivierung* und *Transitivierung* zu beschreiben, ähnlich wie schon beim *Nominalisierungs*-Suffix, an der terminologischen Orientierung an Sprachen, in denen eine Unterscheidung von Verb und Nomen möglich ist. Abgesehen von der Frage, wie gut die teilweise stark divergierenden Bedeutungen der Basen und Derivate überhaupt mithilfe dieser Funktionsetiketten erfasst werden können, ist es im Hawaiischen einfach nicht möglich, die Prädisposition von derivierten Wörtern für ein bestimmtes Syntagma zu zeigen. Auch unter der Annahme, dass die Funktion von Suffixen z.B. *Transitivierung* ist, kann diese Funktion eben keiner *verbalen* Derivationsmorphologie zugeschrieben werden, weil bei einem *transitivierten* Wort im Hawaiischen nicht zwischen *Verb* und *Nomen* unterscheidbar ist.

Das wird noch deutlicher, vergleicht man wieder das Hawaiische mit dem Deutschen: Im Deutschen kann beispielsweise bei einer verbalen Basis die Hauptfunktion des Präfix *be-* als Transitivierung bestimmt werden (vgl. Fleischer und Barz 2012: 383). Die Basis ist klassifiziert als intransitives Verb, z.B. *steigen*, das Derivat als transitives Verb, z.B. *be-steigen* (vgl. Fleischer und Barz 2012: 384). Im Deutschen lässt sich *be-* der Wortbildung von Verben zuordnen, weil das Derivat der Wortart Verb angehört – Die Basis kann auch substantivisch oder adjektivisch sein, wobei die Funktion dann nicht mehr Transitivierung ist (vgl. Fleischer und Barz 2012: 385). Das Präfix *be-* tritt zwar auch in dem Wort *Be-steigung* auf, allerdings kann *Be-stei-gung* in klarer Weise als Nomen von dem Verb *be-steigen* unterschieden und zudem als Derivation mit dem der nominalen Wortbildung zuweisbaren Suffix *-ung* von *be-steigen* analysiert werden.

Im Hawaiischen ist es nicht möglich, in Analogie zum Deutschen innerhalb von Wortartkategorien bestimmte Derivationsmorpheme zu identifizieren bzw. *Wortarten* über Wortbildungsprozesse zu definieren.

Um das Bild der Derivation im Hawaiischen abschließend zu vervollständigen, sollen folgend noch weitere, bislang nicht berücksichtigte Wortbildungsprozesse betrachtet werden, die Elbert und Pukui (1979: 64–89) beschreiben, ohne sie ausschließlich innerhalb einer Kategorie *Verb* oder *Nomen* zu verorten. Dazu gehören unter anderem zwei Wortbildungsprozesse, die Elbert und Pukui (1979: 71–78) ähnlich wie die in Tabelle 22 (S. 168) beschriebenen unter einer bestimmten Funktion zusammenzufassen versuchen (vgl. Tabelle 25, S.174).

	Funktion	Form	Kategorie	
			Basis	Derivat
Präfix	Kausativierung	<i>ho'o-</i> , <i>ho'-</i> , <i>hō'-</i> , <i>hō-</i> , <i>ho-</i>	Verb, Nomen, Nomen-Verb	Verb, Nomen Nomen-Verb
	Stativierung	<i>mā-</i> , <i>ma-</i> , <i>mo-</i> , <i>mō-</i> <i>nā-</i> , <i>na-</i> , <i>nō-</i> , <i>no-</i> , <i>pā-</i> , <i>pa-</i> , <i>pū-</i> , <i>pu-</i> , <i>kū-</i> , <i>ku-</i> , <i>u-</i> , <i>'u-</i> , <i>'ū-</i>	Verb, Nomen, Nomen-Verb	Verb, Nomen, Nomen-Verb

Tabelle 25: Nicht-kategoriale derivationsmorphologische Prozesse

Die Unterscheidung der Allomorphe *ho'o-*, *ho'-*, *hō'-*, *hō-* und *ho-* in Tabelle 25 von den in Tabelle 22 (S. 168) unter der Funktion *Kausativierung* genannten Formen basiert im Wesentlichen darauf, dass Elbert und Pukui (1979: 77) davon ausgehen, dass mit diesen nicht ausschließlich *verbale* sondern auch als *Nomen-Verb* klassifizierte Derivate gebildet werden. Da wie gezeigt auch Derivate, die mit den in Tabelle 22 (S. 168) aufgeführten Formen gebildet sind, nicht als *Verben* kategorisiert werden können, ist die Unterscheidung von Elbert und Pukui (1979) letztlich überflüssig.

Auch bei den Formen, die Elbert und Pukui (1979: 71) mit der Funktion „qualitative / stative meaning“ beschreiben, lässt sich Elbert und Pukui (1979: 71) zufolge weder jede Basis noch jedes Derivat ausschließlich als *Verb* oder *Nomen* klassifizieren. Die Funktion der *Stativierung* kann also nicht als Bildung von *Zustandsverben* verstanden werden, wobei ähnlich wie bei der *Kausativierung* unklar ist, ob sich alle von Elbert und Pukui (1979: 71) genannten Morphemalternanten eindeutig unter einer Funktion *Stativierung* subsumieren lassen. Während zwar Regeln angegeben werden können, in welchem lautlichen Umfeld eine bestimmte Form auftritt (vgl. Elbert und Pukui 1979: 71–73), scheint es weder möglich, die Derivate aller Formen noch die Derivate mit bestimmten Formen semantisch einheitlich zu beschreiben oder überhaupt einen Bedeutungsunterschied zwischen Basis und Derivat zu analysieren, vgl.: „[...] in most instances there is little difference between meaning of base and prefix-plus-base.“ (Elbert und Pukui 1979: 71).

Vereinzelt analysieren Elbert und Pukui (1979: 74–79) auch Affixe mit Intensivierungsfunktion, z.B. *kī-* in *kī-au* „to gallop, fast movement“ mit *au* „to move, movement“ (vgl. Elbert und Pukui 1979: 76), oder pejorativer Funktion, z.B. *-ea* in *puku-ea* „body odor“ mit *puku* „moist“ (vgl. Elbert und Pukui 1979: 78). Zudem nennen Elbert und Pukui (1979: 74 - 79) eine ganze Reihe weiterer, semantisch durchsichtiger Affixe, dessen einzelne Bedeutungen aber unter keiner bestimmten Funktion zusammengefasst werden können. Dazu gehören

aka- „carefully, slowly“ in z.B. *aka-‘ai* „to eat slowly“ von *‘ai* „to eat, food“ (vgl. Elbert und Pukui 1979: 74) und *-kīkē* „back and forth“ in z.B. *pā-kīkē* „to toss back and forth“ von *pā* „to tilt backwards“ (vgl. Elbert und Pukui 1979: 78). Die meisten der von Elbert und Pukui (1979: 74–79) angeführten Affixe sind nicht Teil der produktiven Wortbildung im Hawaiischen und in manchen Fällen kann auch nicht zwischen Affix und Wortstamm unterschieden werden. So analysieren Elbert und Pukui (1979: 76–78) z.B. in dem Wort *‘ele-hine* „old woman“ sowohl *‘ele-* als Präfix mit der Bedeutung „old“ als auch *-hine* als Suffix mit der Bedeutung „female“ (vgl. Elbert und Pukui 1979: 78).

Die abschließende Betrachtung von Affixen, die Elbert und Pukui (1979) zufolge nicht innerhalb einer *Wortart* auftreten, sollte hier vor allem dazu dienen, aufzuzeigen, dass es in der Gesamtperspektive auf die Derivation im Hawaiischen problematisch erscheint, Affixe unter bestimmten Funktionen zusammenzufassen, weil die Funktionen einfach nicht ausreichend verstanden sind. Es ist dann kaum überraschend, dass sich *Kausativierung*, *Passivierung* und *Transitivierung* nicht als Funktionen innerhalb der Kategorie *Verb* beschreiben lassen und die Annahme einer *verbalen* Derivationsmorphologie im Hawaiischen genauso wenig überzeugend ist wie Annahme von kategorienverändernder Derivation. Die Folgen, die daraus und aus den bisherigen Untersuchungen zur Wortverteilung im Korpus für die Frage von über Distributionsklassen bestimmten *Wortarten* im Hawaiischen entstehen, sollen im nächsten Kapitel noch einmal zusammenfassend dargestellt werden.

4.1.5 Zusammenfassung Distribution

Die Annahme, Wörter seien im Hawaiischen für das Auftreten in bestimmten Syntagmen prädisponiert, muss zunächst aus der Beobachtung der Verteilung von Wörtern auf Tempus-Aspekt-Modus-Syntagma, Artikel-Syntagma, „Nominalisierungskonstruktion“ und attributiven Slot im Korpus zurückgewiesen werden. Denn wie gezeigt wurde, treten Wörter im Korpus nicht ausschließlich in einem dieser Syntagmen auf, sondern in 14 aus insgesamt 15 möglichen Kombinationen von Syntagmen (vgl. Tabelle 11, S. 108). Zudem tritt eine deutliche Mehrheit der Wörter im Korpus in mehr als einem Syntagma auf und es gibt Hinweise darauf, dass mit der Gesamtauftrittshäufigkeit eines Wortes auch der Flexibilitätsgrad steigt (vgl. Kapitel 4.1.1, S. 110–112). Damit ist auch die unter anderem von Dixon (2010: 45) hervorgebrachte Vermutung, dass flexibel auftretende Wörter in den polynesischen Sprachen einen Sonderfall darstellen, zumindest für das Hawaiische widerlegt. Der Abgleich der Beobachtungen im

Korpus mit den Angaben in den Wörterbüchern von Pukui und Elbert (1986) und von Parker (1922) hat dann auch die potentielle Flexibilität von Wörtern im Hawaiischen in großen Teilen bestätigt (vgl. Kapitel 4.1.2, S. 118–131).

Der Versuch von Elbert und Pukui (1979), Ordnung in die Vielfalt der verschiedenen Auftretensweisen von Wörtern zu bringen und über die Bildung von Distributionsklassen *Wortarten* zu bestimmen, kann schließlich nicht überzeugen. Denn erstens können keine formalen Parameter herangezogen werden, um Distributionsklassen voneinander zu unterscheiden: Weder ist die Beschreibung von *lexikalischer Nominalisierung* als kategorienverändernde Derivation möglich, wenn sowohl Basen als auch derivierte Wörter in allen Syntagmen auftreten (vgl. Kapitel 4.1.3, S. 139–151). Noch kann eine *verbale* Derivationsmorphologie angesetzt werden, wenn vermeintlich derivierte *Verben* in allen Syntagmen auftreten (vgl. Kapitel 4.1.4, S. 167–173).

Zweitens kann nicht begründet werden, dass Wörter, die potentiell in allen Syntagmen außer in Artikel-Syntagma auftreten, eine Distributionsklasse *Verben* bilden und demgegenüber Wörter, die potentiell in allen Syntagmen auftreten außer in Tempus-Aspekt-Modus-Syntagma, eine Distributionsklasse *Nomen* bilden, wenn Wörter potentiell sowohl in Artikel-Syntagma als auch Tempus-Aspekt-Modus-Syntagma auftreten (vgl. Kapitel 4.1.3, S. 165–167).

Zusammenfassend kann also festgehalten werden, dass die Annahme der Prädisposition von Wörtern für das Auftreten in bestimmten Syntagmen im Hawaiischen aus formaler und distributionaler Perspektive abgelehnt werden muss. Im nächsten Kapitel soll eine alternative Annahme überprüft werden, in der davon ausgegangen wird, dass sich die Prädisposition von Wörtern für bestimmte syntaktische Slots im Hawaiischen durch Frequenzklassen anstelle von Form- oder Distributionsklassen nachweisen lässt.

4.2 Häufigkeitsverteilung und Frequenzklassen

Bei der Frage nach der Bildung von Distributionsklassen ging es bisher um die Untersuchung der potentiellen absoluten Verteilung einzelner Wörter, das heißt um das Vorkommen versus Nicht-Vorkommen von Wörtern in bestimmten (Kombinationen von) Syntagmen. Dass Wörter im Korpus unterschiedlich häufig in verschiedenen Syntagmen auftreten, hat dafür im Wesentlichen keine Rolle gespielt.

Bevor im Folgenden genauer analysiert wird, ob sich in der Häufigkeitsverteilung von Wörtern auf bestimmte Syntagmen im Korpus Klassen unterscheiden las-

sen, soll ein Überblick über die Annahmen gegeben werden, die dem Versuch zugrunde liegen, über Frequenzklassen *Wortarten* in Sprachen wie dem Hawaiischen zu identifizieren.

4.2.1 Frequenz und Prototypikalität

Die Frage, ob im Hawaiischen Frequenzklassen statt Distributionsklassen analysiert werden können, ergibt sich aus den Ansätzen, die von einer prototypischen Abdeckung semantisch-ontologischer Klassen auf syntaktische Slots ausgehen (vgl. Kapitel 2.2.4, S. 25–55).

Diese Ansätze eint die Grundannahme, dass es möglich ist, die Bedeutung eines Wortes einer der semantisch-ontologischen Klassen EREIGNIS, OBJEKT oder EIGENSCHAFT zuzuordnen und entsprechend dieser Klassifizierung die prototypische Auftretensweise des Wortes in Tempus-Aspekt-Modus-Syntagma (*Verbalphrase*), Artikel-Syntagma (*Nominalphrase*) oder attributivem Slot gegeben ist. Ähnliche Annahmen über die semantische Bestimmbarkeit von *Wortarten* finden sich bereits – wenn auch meist nicht explizit ausgearbeitet – in vielen Einzelbeschreibungen polynesischer Sprachen, beispielsweise der hawaiischen Grammatik von Elbert und Pukui (1979) (vgl. Kapitel 2.2.1, S. 11–14). In der Konzeption der Auftretensweise von Wörtern als *prototypische* Auftretensweise spielt jedoch die Bildung von Distributionsklassen eine untergeordnete Rolle, da es im Kern nicht darum geht zu zeigen, dass bestimmte Wörter ausschließlich in ganz bestimmten (Kombinationen von) Syntagmen auftreten, sondern darum, dass auch wenn Wörter potentiell in allen (Kombinationen von) Syntagmen auftreten, eine der Auftretensweisen als die *prototypische* identifizierbar ist. Was genau unter *Prototypikalität* zu verstehen ist, wird in solchen Ansätzen entweder stillschweigend vorausgesetzt (vgl. z.B. Völkel 2017) oder mit Verweis auf die Arbeiten von Rosch (1978) dargelegt (vgl. z.B. Croft 2001: 73). Entscheidend ist dann auch weniger eine genaue Bestimmung des Begriffs der *Prototypikalität* als vielmehr die Annahme, dass sich Parameter angeben lassen, anhand derer der prototypische Fall vom nicht-prototypischen Fall unterschieden werden kann.

Da der von Croft (2001: 89 und passim) konstruierte Parameter der *typologischen Markiertheit*, bei dem einem formal unmarkierten Fall ein formal markierter Fall gegenübergestellt wird, in den polynesischen Sprachen nicht anwendbar ist (vgl. Kapitel 2.2.4, S. 36–37), hat unter anderem Völkel (2017) vorgeschlagen, die Häufigkeit, mit der ein Wort in einem bestimmten Syn-

tagma auftritt, als Parameter für die prototypische Auftretensweise heranzuziehen. *Prototypikalität* lässt sich dann insofern als abgeschwächte Form der Prädisposition betrachten, dass damit die Tendenz eines Wortes, in einem bestimmten Syntagma häufiger als in anderen Syntagmen aufzutreten, bezeichnet wird.

Bei dem Versuch, *Wortarten* in den polynesischen Sprachen über die prototypische Auftretensweise von Wörtern zu definieren, werden also Distributionsklassen durch Frequenzklassen ersetzt. Zunächst unabhängig von der semantischen Klassifizierung sind dann im Hawaiischen nicht, wie von Elbert und Pukui (1979) vorgeschlagen, *Verb*, *Nomen* und *Verb-Nomen* anhand des Auftretens in verschiedenen (Kombinationen von) Syntagmen zu unterscheiden, sondern *Verb*, *Nomen* und *Adjektiv* anhand des häufigeren Auftretens in einem bestimmten Syntagma, vgl.:

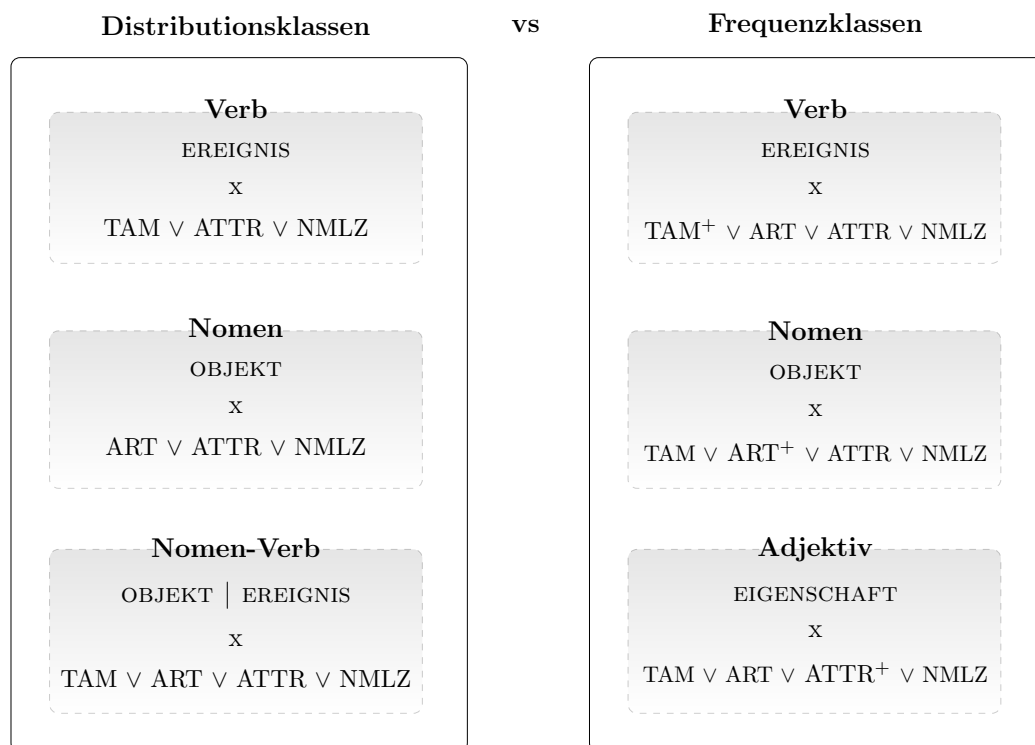


Abbildung 27: Vergleich Distributionsklassen und Frequenzklassen

In Abbildung 27 ist zu erkennen, dass je nach zugrundegelegter Klassenbildung, andere *Wortarten* im Hawaiischen angesetzt werden: Während *Adjektive*

keine eigene Distributionsklasse darstellen, sind *Nomen-Verben* als Frequenzklasse nicht definiert. Daran wird deutlich, dass der Versuch, *Wortarten* über die prototypische Auftretensweise zu bestimmen, Bemühungen geschuldet ist, alle Wörter eindeutig einer *Wortart* zuzuordnen und damit Probleme bei der Konzeption einer flexiblen *Wortart* wie *Nomen-Verb* zu überwinden. Ungeklärt bleibt bei der Ansetzung von Frequenzklassen die Frage des Umgangs mit der „Nominalisierungs-konstruktion“.

In gleicher Weise wie nun die Annahme, Wörter seien gemäß ihrer Zugehörigkeit zu einer semantisch-ontologischen Klasse für das Auftreten in spezifischen Kombinationen aus Syntagmen prädisponiert, die eindeutige Bildung potentieller Distributionsklassen voraussetzt, gilt für die Annahme, Wörter seien gemäß ihrer Zugehörigkeit zu einer semantisch-ontologischen Klasse für das mehrheitliche Auftreten in einem ganz bestimmten Syntagma prädisponiert, dass zunächst unabhängig von der Frage der semantischen Klassifizierung eindeutig Frequenzklassen identifiziert werden müssen. Es geht also nicht darum, dass einzelne Wörter häufiger oder weniger häufig in einem bestimmten Syntagma auftreten, sondern dass im Hawaiischen Klassen von Wörtern differenziert werden können.

Im folgenden Abschnitt soll überprüft werden, ob sich in den Daten im Korpus über die Auftretenshäufigkeit von Wörtern in bestimmten Syntagmen Klassen analysieren lassen.

4.2.2 Frequenz und Klassenbildung

Eine naheliegende Möglichkeit zu untersuchen, ob sich Wörter im Korpus über die Häufigkeit, mit der sie jeweils in einem bestimmten Syntagma auftreten, in Klassen einteilen lassen, besteht darin, die Auftretenshäufigkeiten aller Wörter in den einzelnen Syntagmen gegenüberzustellen.

In Abbildung 28 (S. 180)⁴⁴ sind des besseren Überblicks wegen die Anteile der Häufigkeiten, mit denen die 2251 im Korpus vorkommenden Wörter in Tempus-Aspekt-Modus-Syntagma, Artikel-Syntagma, attributivem Slot und „Nominalisierungs-konstruktion“ auftreten, in einem jeweils eigenen Diagramm zusammengefasst.

⁴⁴Zur Erstellung der Diagramme wurde die Programmiersprache *R* verwendet (vgl. <https://www.r-project.org/>).

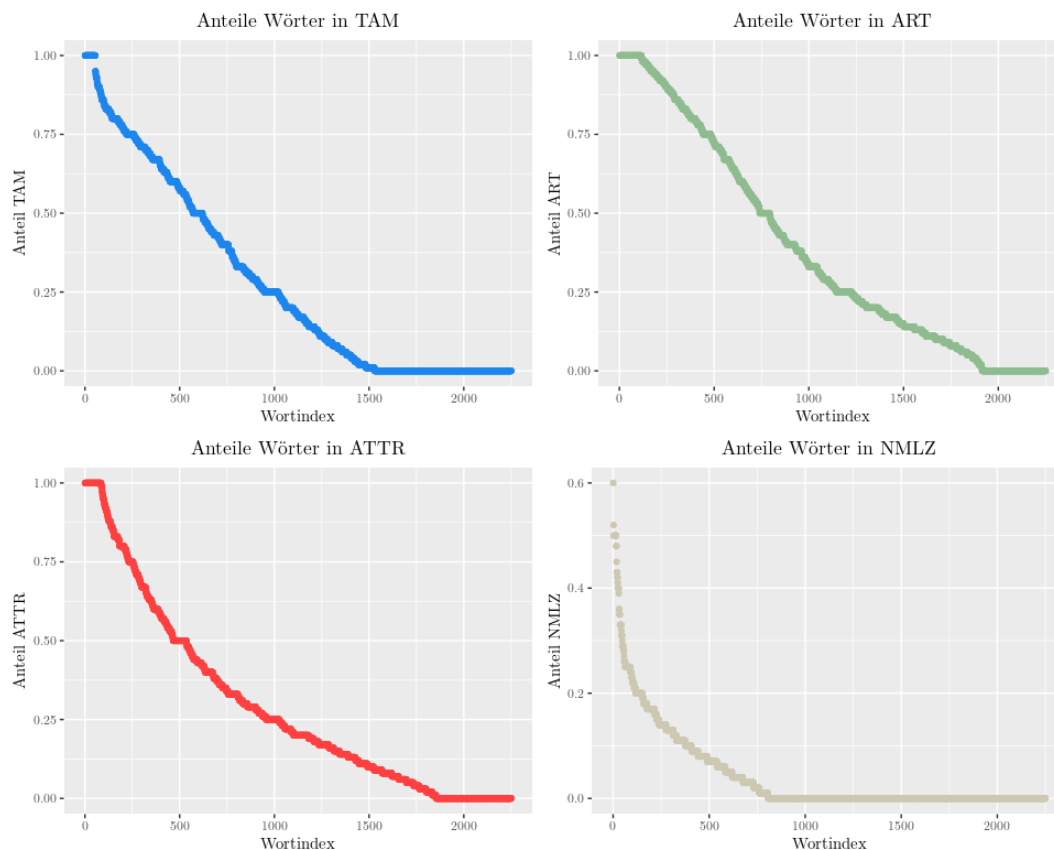


Abbildung 28: Vergleich Anteile in Syntagmen

Auf der x-Achse eines Diagramms in Abbildung 28 ist der Index eines Wortes angegeben, auf der y-Achse der Anteil an der Gesamtfrequenz, mit dem das Wort in dem entsprechenden Syntagma auftritt. Die Berechnung des Anteils ergibt sich beispielsweise für das Wort *kaupaona* folgendermaßen: Im Korpus tritt *kaupaona* insgesamt 30 mal auf, davon 16 mal in Tempus-Aspekt-Modus-Syntagma, 4 mal in Artikel-Syntagma, 6 mal in attributivem Slot und 4 mal in „Nominalisierungs-konstruktion“. Der Anteil der Auftretenshäufigkeit von *kaupaona* in Tempus-Aspekt-Modus-Syntagma beträgt dann aufgerundet auf zwei Nachkommastellen $\frac{16}{30} = 0,53$ oder 53%, in Artikel-Syntagma $\frac{4}{30} = 0,13$ oder 13%, in attributivem Slot $\frac{6}{30} = 0,20$ oder 20% und in „Nominalisierungs-konstruktion“ $\frac{4}{30} = 0,13$ oder 13%.

Die Wörter sind in den Diagrammen entlang der x-Achse in abnehmender Reihenfolge vom höchsten Anteilswert zum niedrigsten Anteilswert angeordnet. Ein Anteil von 100 % spiegelt dann wider, dass ein Wort ausschließlich in dem Syntagma auftritt, ein Anteil von 0 %, dass ein Wort gar nicht in dem

Syntagma auftritt. Die „Nominalisierungsstruktur“ ist das einzige Syntagma, in dem keine Wörter im Korpus ausschließlich auftreten und in dem mehr Wörter als in den anderen Syntagmen nicht auftreten.

Auch wenn die „Nominalisierungsstruktur“ in Ansätzen zu Frequenzklassen keine Berücksichtigung findet und bislang die Beschreibung von *Nominalisierung* im Hawaiischen kritisiert wurde (vgl. Kapitel 4.1.3, S. 131–167), sind die Anteile von Wörtern in „Nominalisierungsstruktur“ in Abbildung 28 (S. 180) separat betrachtet worden. Es ist aber ebenso möglich, die „Nominalisierungsstruktur“ zum Artikel-Syntagma zu zählen. Die Anteile von Wörtern in Tempus-Aspekt-Modus-Syntagma und attributivem Slot ändern sich dadurch nicht wesentlich und da Wörter in „Nominalisierungsstruktur“ insgesamt eher niedrig frequent auftreten, steigen die Anteile von Wörtern in Artikel-Syntagma dadurch nicht erheblich. In Abbildung 29 sind die Anteile von Wörtern in Artikel-Syntagma mit und ohne Einbezug der „Nominalisierungsstruktur“ gegenübergestellt, vgl.:

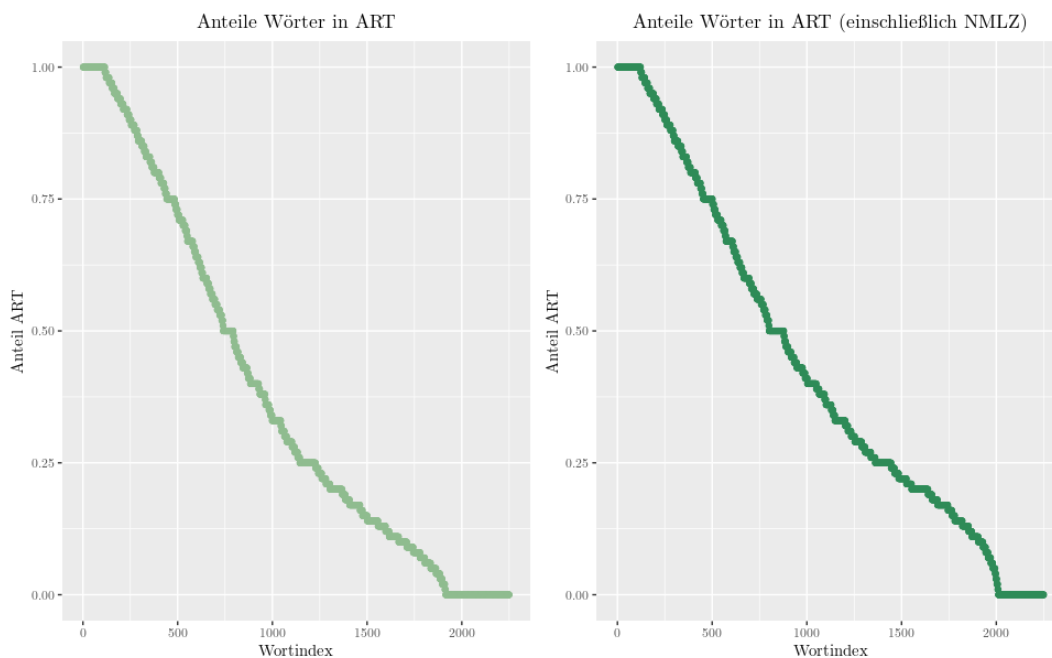


Abbildung 29: Vergleich Anteile in Artikel-Syntagma mit und ohne NMLZ

Abbildung 129 (S. 181) lässt sich dann entnehmen, dass keine wesentlichen Unterschiede bei den Anteilen der Wörter in Artikel-Syntagma bestehen. Entscheidend ist nun, dass aus der Betrachtung der Anteile, mit denen Wörter in einzelnen Syntagmen auftreten, Unklarheiten bei der Bildung von Frequenzklassen hervorgehen. Denn wie in Abbildung 28 (S. 180) zu erkennen ist, treten Wörter in einem Spektrum von hohen bis niedrigen Werten zu allen möglichen Anteilen in Tempus-Aspekt-Modus-Syntagma, Artikel-Syntagma, attributivem Slot und „Nominalisierungs konstruktion“ auf.⁴⁵ Das bedeutet, dass nicht ohne weiteres ersichtlich ist, durch welche Wertebereiche genau Frequenzklassen eingegrenzt sind. Soll beispielsweise ein Wort wie *kaupaona*, das im Korpus zu 53% in Tempus-Aspekt-Modus-Syntagma auftritt, als *Verb* klassifiziert werden, ein Wort wie *inu*, das nur zu 49% in Tempus-Aspekt-Modus-Syntagma auftritt, hingegen nicht? Bei Wörtern wie *welina*, die im Korpus genau zu 50% in Tempus-Aspekt-Modus-Syntagma und zu 50% in Artikel-Syntagma auftreten, ist die prototypische Auftretensweise schließlich gar nicht zu bestimmen.

An diesen Beispielen wird deutlich, dass die Annahme von Frequenzklassen im Kern ähnliche Schwachstellen aufweist wie die Annahme von Distributionsklassen. So wie Wörter in allen möglichen Kombinationen aus Syntagmen auftreten und es nicht eindeutig möglich ist, dabei bestimmte Kombinationen als potentielle Distributionsklassen zu differenzieren, treten Wörter in allen möglichen Häufigkeiten in einem bestimmten Syntagma auf und es ist nicht eindeutig möglich, diskrete Frequenzklassen zu identifizieren. Es ist nicht überzeugend, einfach Wörter, die zu mehr als 50% in einem Syntagma auftreten, einer *Wortart* zuzuweisen. Würde man aber willkürlich andere Wertegrenzen wählen, innerhalb derer eine Frequenzklasse definiert ist, z.B. 75% – 100% Auftreten eines Wortes in Tempus-Aspekt-Modus-Syntagma als Frequenzklasse *Verb* und 25% – 0% Auftreten in Tempus-Aspekt-Modus-Syntagma als andere Frequenzklasse, wäre immer noch der Bereich zwischen 25% und 50% des Auftretens keiner Frequenzklasse und damit keiner *Wortart* zuzurechnen. Mit der Annahme von Frequenzklassen kann das Problem der Definition einer Klasse *Verb-Nomen* letztlich nicht überwunden werden. Denn so wenig wie Wörter, die sowohl in Tempus-Aspekt-Modus-Syntagma als auch in Artikel-Syntagma auftreten, einer der beiden Distributionsklassen

⁴⁵Da sehr viele Wörter betrachtet wurden, kommt es zu einem gewissen „Overplotting“ in den Diagrammen in Abbildung 28 (S. 180). Dies wurde so belassen, weil daran deutlich wird, wie nahe die Frequenzwerte einzelner Wörter beieinander liegen.

Verb oder *Nomen* angehören können, lassen sich Wörter, die ähnlich häufig in Tempus-Aspekt-Modus-Syntagma und Artikel-Syntagma auftreten, einer der beiden Frequenzklassen *Verb* oder *Nomen* zuordnen.

Überdies finden sich zu fast allen Kombinationen von Syntagmen Wörter im Korpus, die in gleichen Anteilen in den Syntagmen auftreten, vgl.:

Anteile in Syntagma				in Korpus	Beispiel
TAM	ART	ATTR	NMLZ		
NMLZ separat					
0,50	0,50			✓	<i>welina</i>
0,50		0,50		✓	<i>maha'oi</i>
0,50			0,50	✓	<i>nalohia</i>
	0,50	0,50		✓	<i>lomaloma</i>
	0,50		0,50	✓	<i>ha'anui</i>
		0,50	0,50	✓	<i>kūloa</i>
0,33	0,33	0,33		✓	<i>uihā</i>
0,33	0,33		0,33	✗	n/a
	0,33	0,33	0,33	✗	n/a
0,25	0,25	0,25	0,25	✓	<i>nunulu</i>
ART inklusive NMLZ					
0,50	0,50			✓	<i>liki</i>
0,50		0,50		✓	<i>hainā</i>
	0,50	0,50		✓	<i>nu'u</i>
0,33	0,33	0,33		✓	<i>popopo</i>

Tabelle 26: Kombinationen gleiche Anteile in Syntagmen

Im oberen Teil von Tabelle 26 ist die „Nominalisierungskonstruktion“ separat betrachtet, im unteren Teil zum Artikel-Syntagma gezählt. Tabelle 26 kann entnommen werden, dass im Korpus in allen Kombinationen von Syntagmen außer den Kombinationen aus drei Syntagmen mit „Nominalisierungskonstruktion“ Wörter vorkommen, die zu gleichen Anteilen auf die in der Kombination enthaltenen Syntagmen verteilt sind. Das bedeutet, dass auf Basis der Auftretenshäufigkeiten nicht nur, wie schon bei den Distributionsklassen, eine Frequenzklasse der *Verb-Nomen* angesetzt werden müsste, sondern zusätzlich eine Klasse der *Verb-Adjektive*, *Nomen-Adjektive* und *Verb-Nomen-Adjektive*.

Aus diesen Beobachtungen ergibt sich auch die Frage, ob Wörter tendenziell

auf alle Syntagmen, in denen sie auftreten, gleich verteilt sind bzw. ein Zusammenhang zwischen dem Häufigkeitsanteil, mit dem ein Wort in einzelnen Syntagmen auftritt und der Anzahl der Syntagmen, in denen ein Wort auftritt, besteht.

In Abbildung 30 sind jeweils für das Tempus-Aspekt-Modus-Syntagma, das Artikel-Syntagma einschließlich der „Nominalisierungskonstruktion“ und den attributiven Slot die Anteile in Kombinationen aus Syntagmen in einem Streifendiagramm gegenübergestellt, vgl.:

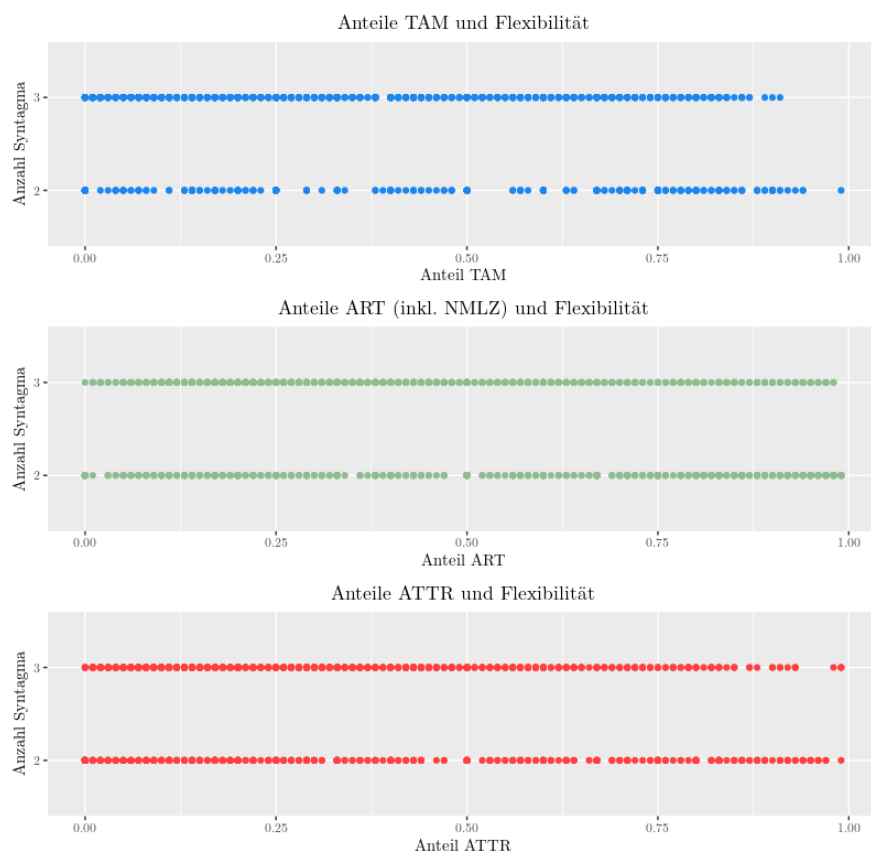


Abbildung 30: Vergleich Anteile in Syntagmen und Flexibilität

In Abbildung 30 sind diejenigen Wörter nicht berücksichtigt, die ausschließlich in einem Syntagma auftreten, weil es für die Frage des Zusammenhangs von Auftretensanteil und Flexibilität keinen informativen Beitrag liefert, Wörter mit einem Auftretensanteil von 100% zu betrachten. Da die „Nominalisierungskonstruktion“ hier zum Artikel-Syntagma gezählt wurde, treten Wörter

höchstens in einer Kombination aus drei Syntagmen auf.

Aus den Diagrammen in Abbildung 30 (S. 184) gehen keine Hinweise hervor, dass sich die Anteile, mit denen ein Wort in einem Syntagma auftritt, wesentlich nach der Anzahl der Syntagmen, in denen ein Wort auftritt, unterscheiden. Beispielsweise erreichen sowohl Wörter, die in zwei als auch in drei Syntagmen auftreten, Häufigkeitsanteile zwischen 1% und 99% in Artikel-Syntagma.

An dieser Stelle kann vorerst zusammengefasst werden, dass sich unter Berücksichtigung der Häufigkeitsanteile, mit denen Wörter in bestimmten Syntagmen auftreten, nicht in eindeutiger Weise Frequenzklassen im Korpus bilden lassen. Unabhängig von der in einer Kombination enthaltenen Anzahl an Syntagmen treten Wörter in allen möglichen Anteilen in Tempus-Aspekt-Modus-Syntagma, Artikel-Syntagma, attributivem Slot und „Nominalisierungs-konstruktion“ im Korpus auf und es ist nicht möglich, Wertebereiche für klar voneinander abgegrenzte Frequenzklassen anzugeben.

Damit sind – zumindest über die Häufigkeitsanteile von flexiblen Wörtern in einzelnen Syntagmen – auch keine prototypischen Auftretensweisen von Wörtern für die Festlegung von *Wortarten* nachweisbar.

Ohne den Fokus auf die Bildung von Frequenzklassen zu legen, sind in der korpuslinguistischen Literatur nun eine ganze Reihe weiterer statistischer Kenngrößen zur Messung der Stärke des Zusammenhangs zwischen einem Wort und einer syntaktischen Konstruktion vorgeschlagen worden (vgl. Schmid und Küchenhoff 2013: 532–533 für einen Überblick), besonders in den Arbeiten von Stefanowitsch und Gries (2003) und Gries und Stefanowitsch (2004) im Rahmen der unter dem Begriff *collostructional analysis* zusammengefassten Methoden. Allen voran Kilgariff (2005) hat jedoch kritisiert, dass die Anwendung von Hypothesen-testenden Verfahren in Korpusuntersuchungen, wie in der *collostructional analysis*, substanzwissenschaftlich nicht begründet ist, weil die dafür notwendige Voraussetzung unabhängiger Beobachtungen in einer Zufallsstichprobe in Sprachdaten niemals gegeben sein kann (vgl. auch Schmid und Küchenhoff 2013: 537–538, Bestgen 2014, Lijffijt et al. 2016, Koplenig 2017). Da ich die Skepsis gegenüber der Anwendung Hypothesen-testender Verfahren teile, sollen die entsprechenden Zusammenhangsmaße aus der induktiven Statistik hier keine Berücksichtigung finden. Stattdessen werden die von Schmid und Küchenhoff (2013) dazu vorgeschlagenen, weitgehend voraussetzungslosen Alternativen in Betracht gezogen, um neben den bislang untersuchten Häufigkeitsanteilen von Wörtern im Auftreten in bestimmten Syntagmen auch andere Möglichkeiten bei der Bildung von Frequenzklassen im Korpus

mit einzubeziehen.

Schmid und Küchenhoff (2013: 548) diskutieren unter anderem die Zusammenhangsmaße *Attraction* und *Reliance* (vgl. auch Schmid 2000: 54–55).

Das von Schmid und Küchenhoff (2013: 548) als *Reliance* bezeichnete Maß wurde hier bereits als Auftretensanteil eines Wortes in einem bestimmten Syntagma behandelt.

Das Maß der *Attraction* entspricht hingegen der relativen Frequenz eines Wortes in einer Konstruktion, wird also berechnet, indem die Auftretenshäufigkeit eines Wortes in einem Syntagma durch die Auftretenshäufigkeit dieses Syntagmas im gesamten Korpus geteilt wird. Wenn Konstruktionen deutlich häufiger im Korpus auftreten als einzelne Wörter, das heißt ein niedriger Zählerwert durch einen hohen Nennerwert geteilt wird, fallen die Ergebniswerte dieser Berechnung entsprechend niedrig aus. Um damit umzugehen, können die Ergebniswerte durch Multiplikation mit 100 in Prozentangaben übertragen werden (vgl. Schmid und Küchenhoff 2013: 548). Beispielsweise ergibt sich mit einer Gesamtfrequenz des Tempus-Aspekt-Modus-Syntagma im Korpus von 63.898 bei dem 561 mal im Korpus in Tempus-Aspekt-Modus-Syntagma auftretenden Wort *aloha*:

$$\text{Attraction}_{aloha;TAM} = \frac{561}{63898} \cdot 100 = 0,88\%.$$

Den Namen *Attraction* haben Schmid und Küchenhoff (2013: 548) gewählt, weil sie dieses Maß metaphorsich als die Anziehungskraft, die von einem Syntagma auf ein Wort ausgeht, interpretieren. Je höher also der *Attraction*-Wert, desto höher die Anziehungskraft eines Syntagmas auf ein Wort.

Für die Annahme von über die Auftretenshäufigkeit messbare prototypische Auftretensweise von Wörtern, stellt sich dann die Frage, ob im Korpus mithilfe des *Attraction*-Maßes Klassen analysiert werden können.

In Abbildung 31 (S. 187) sind in dem jeweiligen Diagramm die *Attraction*-Werte der im Korpus auftretenden Wörter für das Tempus-Aspekt-Modus-Syntagma, das Artikel-Syntagma, den attributiven Slot und die „Nominalisierungskonstruktion“ zusammenfassend dargestellt.

Auf der x-Achse eines Diagramms ist der Wortindex eingetragen, auf der y-Achse der berechnete und auf zwei Nachkommastellen aufgerundete *Attraction*-Wert als Prozentangabe. Die Werte sind in allen Diagrammen entlang der x-Achse in abnehmender Reihenfolge eingetragen.

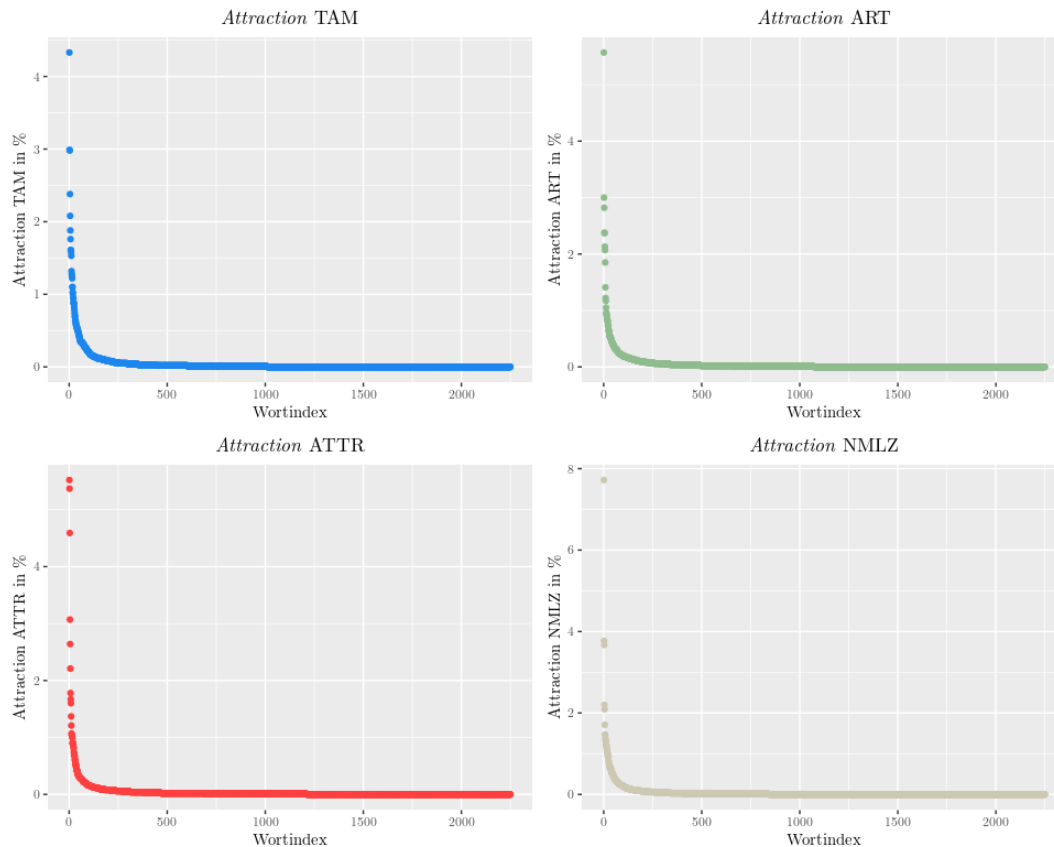


Abbildung 31: Vergleich *Attraction* in Syntagmen

Zunächst kann nach Betrachtung der Diagramme in Abbildung 31 festgehalten werden, dass wie bei den *Reliance*-Werten auch keine wesentlichen Unterschiede der *Attraction*-Werte zwischen Tempus-Aspekt-Modus-Syntagma, Artikel-Syntagma, attributivem Slot und „Nominalisierungskonstruktion“ zu beobachten sind. Das ist wenig überraschend, weil in beiden Maßen letztlich zum Ausdruck kommt, dass Wörter im Korpus in jedem Syntagma in allen möglichen Frequenzen auftreten.

Die Formen der Kurven in den Diagrammen in Abbildung 31 weisen allerdings auf eine Zipfsche Verteilung der Wörter in einzelnen Syntagmen bzw. auf den *Zipfsches Gesetz* genannten Zusammenhang zwischen dem Rang eines Wortes und seiner Häufigkeit hin (vgl. Baayen 2001: 13–24, Pustet 2004, Popescu 2009: 127–130, Cantos Gómez 2013: 180–182, Desagulier 2017: 116–118, Brezina 2018: 44–46). Dieses Gesetz, von Zipf (1949: 24) mit $R \times f = C$ formalisiert, besagt, dass das Produkt aus Rang und der jeweiligen Häufigkeit eines Wortes konstant ist (vgl. auch Zipf: 1935). Der Rang ist dabei der „Platz“ des Wor-

tes in einer nach Häufigkeit geordneten Liste. Rang 1 wäre das am häufigsten auftretende Wort, Rang 2 das am zweithäufigsten auftretende usw.

Um zu überprüfen, wie stark sich die Häufigkeitsverteilungen von Wörtern in einzelnen Syntagmen an eine Zipf-Verteilung annähern, kann eine doppelt logarithmische Darstellung herangezogen werden, in der sowohl die Rangwerte als auch die Frequenzwerte logarithmisch skaliert sind. Bei einer perfekten Zipf-Verteilung sollte sich in dieser Darstellung eine gerade Linie mit einem Anstieg von -1 zeigen (vgl. Baayen 2001: 15, Baayen 2008: 42–43, Cantos Gómez 2013: 181–182). Bei den Wortverteilungen auf einzelne Syntagmen ist dies nicht klar der Fall, vgl.:

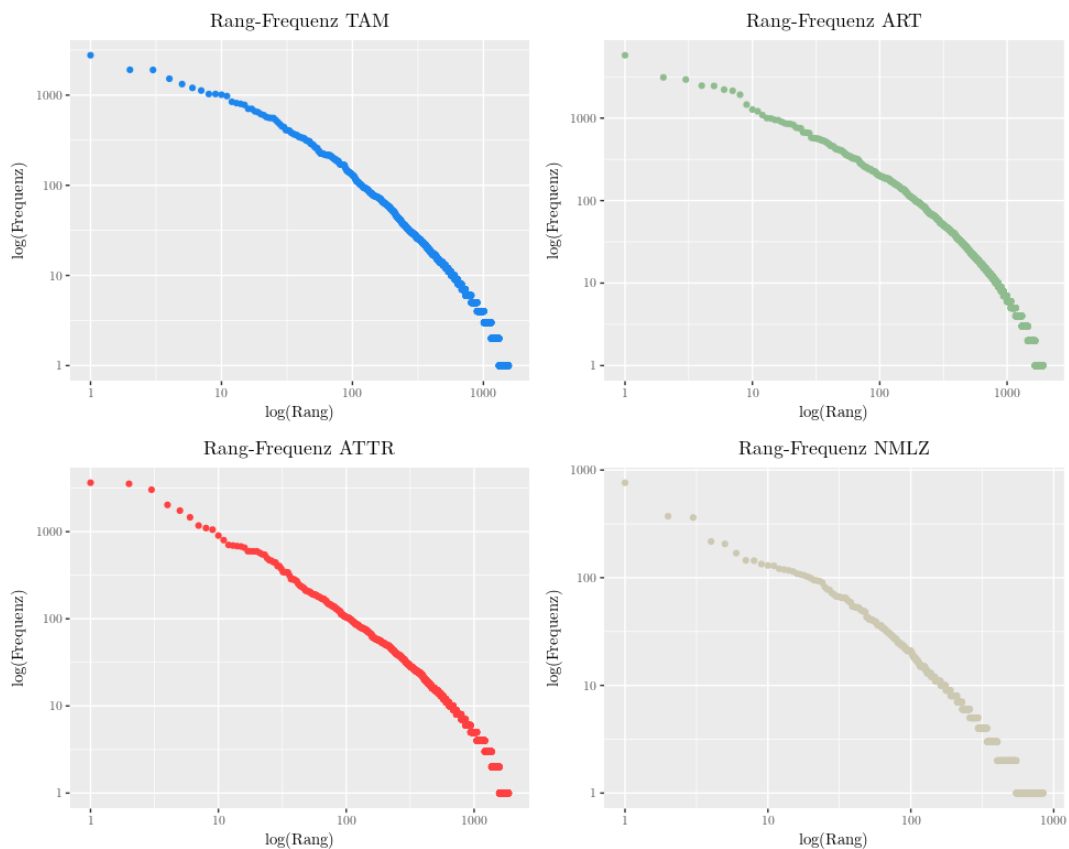


Abbildung 32: Rang-Frequenz-Plot auf einer logarithmischen Skala

Auch wenn in Abbildung 32 (S. 188) Abweichungen der Häufigkeitsverteilungen von einer Zipf-Verteilung erkennbar sind, kann die Kernaussage des *Zipfschen Gesetzes* für die jeweilige Verteilung von Wörtern in Tempus-Aspekt-Modus-Syntagma, Artikel-Syntagma, attributivem Slot und „Nominalisierungs-konstruktion“ übernommen werden: Es treten wenige Wörter mit hoher Frequenz und viele Wörter mit niedriger Frequenz auf.

Die Beobachtung, dass im Korpus Wörter in allen Syntagmen mehrheitlich niedrig frequent auftreten, bedeutet im Kern, dass auch die „Anziehungskraft“ der einzelnen Syntagmen auf die Mehrheit der Wörter schwach ausfällt. Für die Bildung von Frequenzklassen müssten dann unter der Voraussetzung, dass hohe *Attraction*-Werte in einem Syntagma mit niedrigen *Attraction*-Werten in anderen Syntagmen zusammenfallen, die hohen *Attraction*-Werte seltener Wörter herangezogen werden, wobei unklar ist, in welche Klasse Wörter mit niedrigen *Attraction*-Werten und Wörter, die ähnliche *Attraction*-Werte in einzelnen Syntagmen aufweisen, gehören.

Interessant ist diesem Zusammenhang, die *Attraction*-Werte mit dem Flexibilitätsgrad von Wörtern zu vergleichen. In Abbildung 33 (S. 190) sind dazu in dem jeweiligen Diagramm die *Attraction*-Werte von Wörtern in Tempus-Aspekt-Modus-Syntagma, Artikel-Syntagma, attributiven Slot und „Nominalisierungs-konstruktion“ der Anzahl der Syntagmen, in denen die Wörter auftreten, gegenübergestellt.

Betrachtet man Abbildung 33 (S. 190) ist auffällig, dass mit höherer Anzahl an Syntagmen in einer Kombination, in der ein Wort auftritt, auch die *Attraction*-Werte des Wortes steigen. Dies gilt insbesondere für die *Attraction*-Werte von Tempus-Aspekt-Modus-Syntagma und Artikel-Syntagma und im Vergleich von Wörtern, die ausschließlich in einem Syntagma auftreten mit Wörtern, die in allen Syntagmen auftreten. Verwendet man also in der Annahme von Frequenzklassen statt der *Reliance*-Werte die *Attraction*-Werte als Maß für die prototypische Auftretensweise eines Wortes und die Definition von *Wortarten*, bedeutet das folgendes: Ein Wort, das in allen Syntagmen auftritt, gehört eher der Klasse der *Verben* an, als ein Wort, das ausschließlich in Tempus-Aspekt-Modus-Syntagma auftritt. Und gleichermaßen gehört ein Wort, das in allen Syntagmen auftritt, eher der Klasse der *Nomen* an, als ein Wort, das ausschließlich in Artikel-Syntagma auftritt.

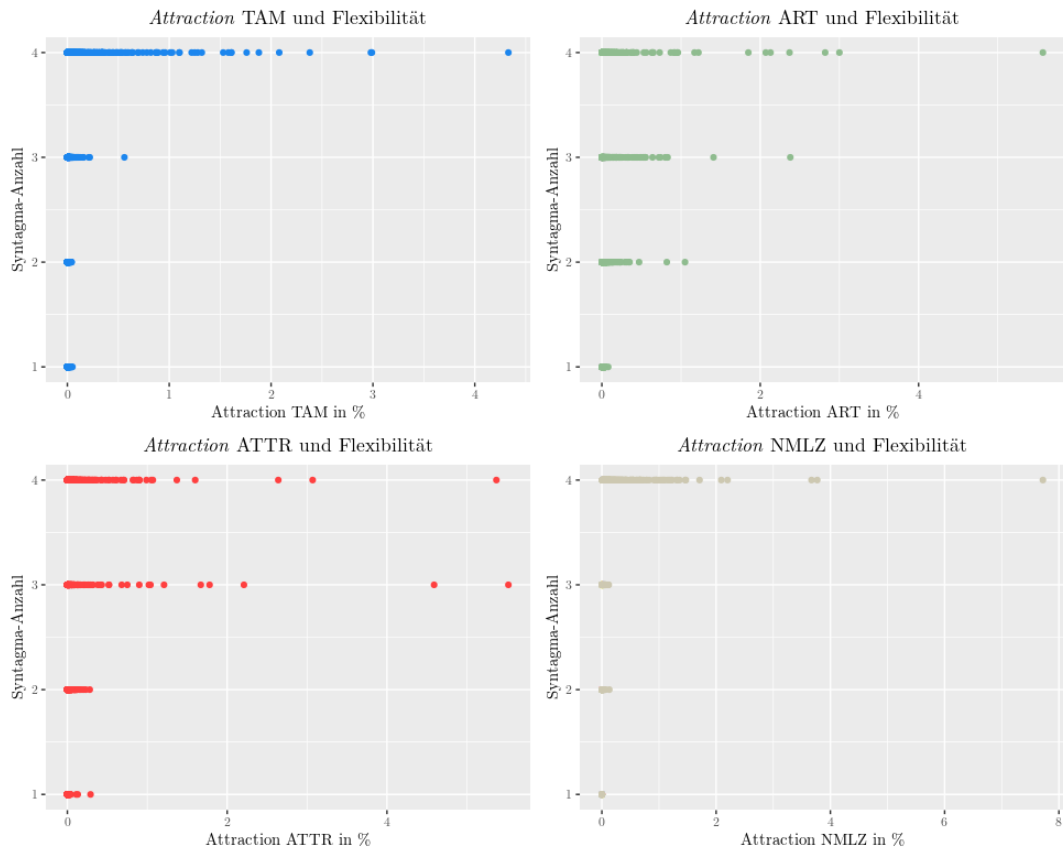


Abbildung 33: *Attraction* und Flexibilität

Der Vollständigkeit halber kann auch wieder der Fall berücksichtigt werden, bei dem die „Nominalisierungs-konstruktion“ zum Artikel-Syntagma gezählt wird (vgl. Abbildung 34, S. 191).

Dabei ergibt sich ein ähnliches Bild: Ein Wort, das in Tempus-Aspekt-Modus-Syntagma, Artikel-Syntagma und attributivem Slot auftritt, müsste anhand der *Attraction*-Werte eher als *Verb*, *Nomen* oder *Adjektiv* klassifiziert werden als ein Wort, das ausschließlich in einem der Syntagmen auftritt.

Aus den Beobachtungen in Abbildung 33 und Abbildung 34 (S. 191) lässt sich schließen, dass alle Wörter im Korpus, auch wenn sie mehrheitlich niedrig frequent vorkommen, tendenziell mit höherer Auftretenshäufigkeit flexibler auftreten. Das bestätigt die Vermutung eines Zusammenhangs von Auftretenshäufigkeit und Flexibilitätsgrad, die bereits anhand der höheren Gesamtfrequenzen (Tokenanzahl) von flexibleren Wörtern im Vergleich zu weniger flexiblen Wörtern aufgestellt wurde (vgl. Kapitel 4.1.1, S. 111).

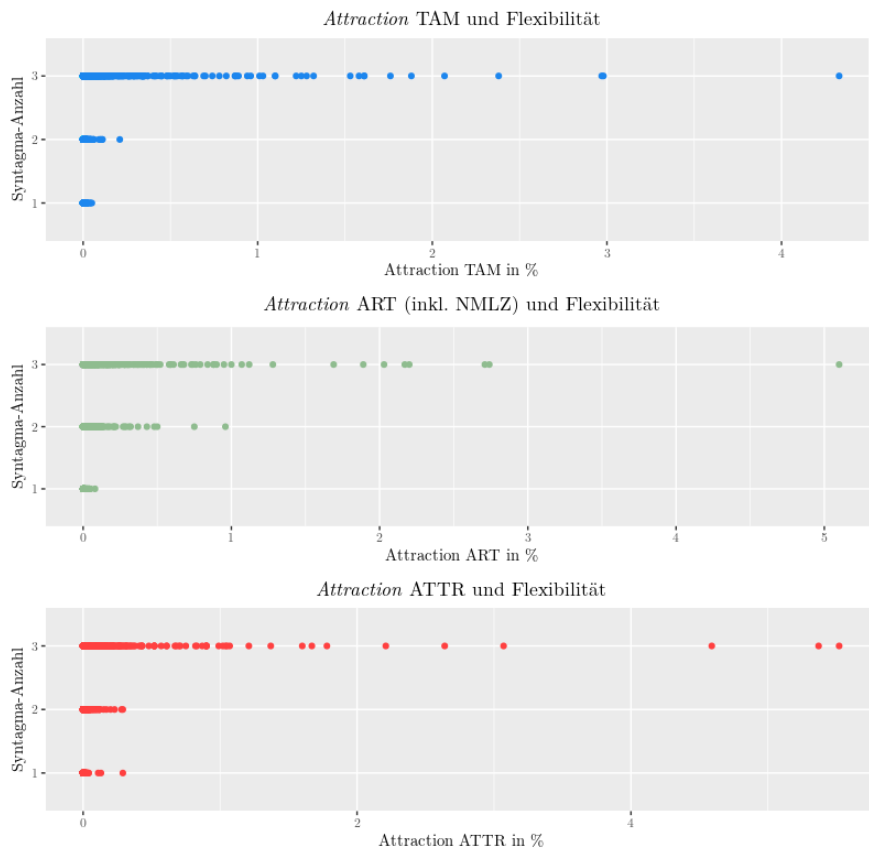


Abbildung 34: *Attraction* und Flexibilität

Wie gezeigt, spricht ein solcher Zusammenhang nun aber nicht nur gegen die Möglichkeit der Bildung von Distributionsklassen, sondern ebenso von Frequenzklassen, zumindest unter Verwendung des *Attraction*-Maßes.

Zusammenfassend kann festgehalten werden, dass auch das *Attraction*-Maß keine besseren Ergebnisse bei der Bildung von Frequenzklassen als der Auftretensanteil bzw. das *Reliance*-Maß liefert, weil auf der einen Seite nur wenige Wörter im Korpus hohe Auftretenshäufigkeiten und damit hohe *Attraction*-Werte erreichen und auf der anderen Seite diejenigen Wörter, die in mehr als einem Syntagma auftreten, eher als *Verb*, *Nomen* oder *Adjektiv* klassifiziert werden. Außerdem ist wie bereits beim *Reliance*-Maß auch beim *Attraction*-Maß unklar, zu welcher Klasse Wörter gehören sollen, die ähnliche Werte für verschiedene Syntagmen aufweisen.

Nach Betrachtung des *Reliance*- und des *Attraction*-Maßes soll abschließend mit der *Odds Ratio* (relative Chance) ein weiteres von Schmid und Küchenhoff

(2013: 552–555) diskutiertes Zusammenhangsmaß für die Bildung von Frequenzklassen im Korpus in Erwägung gezogen werden (vgl. auch Fahrmeir et al. 2007: 119–121).

Das Maß der *Odds Ratio* basiert auf der Berechnung von *Odds* (Chancen), der wiederum die Überlegung zugrunde liegt, dass die mit dem *Reliance*-Maß berechneten Verhältnisse Wahrscheinlichkeiten repräsentieren. Beispielsweise repräsentiert der *Reliance*-Wert von 0,53 des Wortes *kaupaona* in Tempus-Aspekt-Modus-Syntagma die Wahrscheinlichkeit, dass *kaupaona* zu 53% in Tempus-Aspekt-Modus-Syntagma auftritt. Die *Odds* (Chance) für das Auftreten von *kaupaona* in Tempus-Aspekt-Modus-Syntagma werden berechnet, indem die Wahrscheinlichkeit, dass *kaupaona* in Tempus-Aspekt-Modus-Syntagma auftritt durch die Wahrscheinlichkeit geteilt wird, dass *kaupaona* nicht in Tempus-Aspekt-Modus-Syntagma auftritt, oder einfacher, indem die Häufigkeit mit der *kaupaona* in Tempus-Aspekt-Modus-Syntagma auftritt durch die Gesamthäufigkeit, mit der *kaupaona* in anderen Syntagmen auftritt, geteilt wird. Da *kaupaona* im Korpus 16 mal in Tempus-Aspekt-Modus-Syntagma, 4 mal in Artikel-Syntagma, 6 mal in attributivem Slot und 4 mal in „Nominalisierungs-konstruktion“, i.e. insgesamt 14 mal nicht in Tempus-Aspekt-Modus-Syntagma auftritt, ergibt sich:⁴⁶

$$\text{Odds}_{(\text{TAM}, \neg \text{TAM} \mid \text{kaupaona})} = \frac{16}{14} = 1,14$$

Dieser *Odds*-Wert gibt genau genommen die Chance an, dass irgendein Syntagma im Korpus, in dem *kaupaona* auftritt, ein Tempus-Aspekt-Modus-Syntagma ist (vgl. Schmid und Küchenhoff 2013: 553).

Mit dem Maß der *Odds Ratio* werden nun die Chancen für Syntagmen, die ein Wort enthalten, in Relation zu den Chancen für Syntagmen, die dieses Wort nicht enthalten, gesetzt. Zur Berechnung der *Odds Ratio* von *kaupaona* in Tempus-Aspekt-Modus-Syntagma muss also zunächst die Chance berechnet werden, mit der ein Syntagma im Korpus Tempus-Aspekt-Modus ist, aber nicht *kaupaona* enthält. Das geschieht in analoger Weise zur Berechnung der *Odds* von *kaupaona* mit Tempus-Aspekt-Modus-Syntagma, das heißt, es wird die Frequenz der Wörter, die in Tempus-Aspekt-Modus-Syntagma auftreten, aber nicht *kaupaona* sind (63.434), durch die Frequenz der Wörter geteilt, die

⁴⁶Ich folge hier der Notation von Fahrmeir et al. (2007: 119 und passim) und nicht Schmid und Küchenhoff (2013: 552–555).

nicht in Tempus-Aspekt-Modus-Syntagma auftreten und nicht *kaupaona* sind (179.602), vgl.:

$$\text{Odds Ratio}_{(\text{TAM}, \neg \text{TAM} \mid \neg \text{kaupaona})} = \frac{63434}{179602} = 0,35$$

Die *Odds Ratio* gibt nun das Verhältnis dieser beiden Chancen an und wird im Beispiel von *kaupaona* entsprechend berechnet, indem die Chance für das Auftreten von *kaupaona* mit Tempus-Aspekt-Modus-Syntagma durch die Chance für das Auftreten von \neg *kaupaona* mit Tempus-Aspekt-Modus-Syntagma geteilt wird, vgl.:

$$\text{Odds Ratio}_{(\text{TAM}, \neg \text{TAM} \mid \text{kaupaona}, \neg \text{kaupaona})} = \frac{1,14}{0,35} = 3,24$$

Dieser *Odds Ratio*-Wert gibt an, dass ein Syntagma, in dem *kaupaona* auftritt 3,24 mal wahrscheinlicher ein Tempus-Aspekt-Modus-Syntagma ist als ein Syntagma, in dem nicht *kaupaona* auftritt (vgl. Schmid und Küchenhoff 2013: 554–555).

Für die Annahme von über den Parameter Frequenz gemessene prototypische Auftretensweisen von Wörtern stellt sich dann die Frage, ob anhand der relativen Chancen des Auftretens von Wörtern mit bestimmten Syntagmen Klassen gebildet werden können.

In Abbildung 35 (S. 194) sind die *Odds Ratio*-Werte aller im Korpus auftretenden Wörter in jeweils einem Diagramm mit Tempus-Aspekt-Modus-Syntagma, Artikel-Syntagma, attributivem Slot und „Nominalisierungs konstruktion“ zusammengefasst.

In den Diagrammen in Abbildung 35 (S. 194) ist auf der x-Achse der Index eines Wortes angegeben und auf der y-Achse der *Odds Ratio*-Wert des Wortes. Die Werte sind entlang der x-Achsen in abnehmender Reihenfolge angeordnet. Ähnlich wie bei den *Attraction*-Werten ist zu erkennen, dass Wörter in allen Syntagmen mehrheitlich niedrige *Odds Ratio*-Werte erreichen. Das bedeutet, dass die Chancen bei den meisten Wörtern, in einem bestimmten Syntagma aufzutreten, nicht deutlich höher im Vergleich zu anderen Wörtern liegen. Damit entsteht auch das gleiche Problem wie bereits bei den *Attraction*-Werten: Während wenige Wörter zu Klassen höherer Chancen gehören, können viele Wörter nicht eindeutig einer Klasse zugeordnet werden.

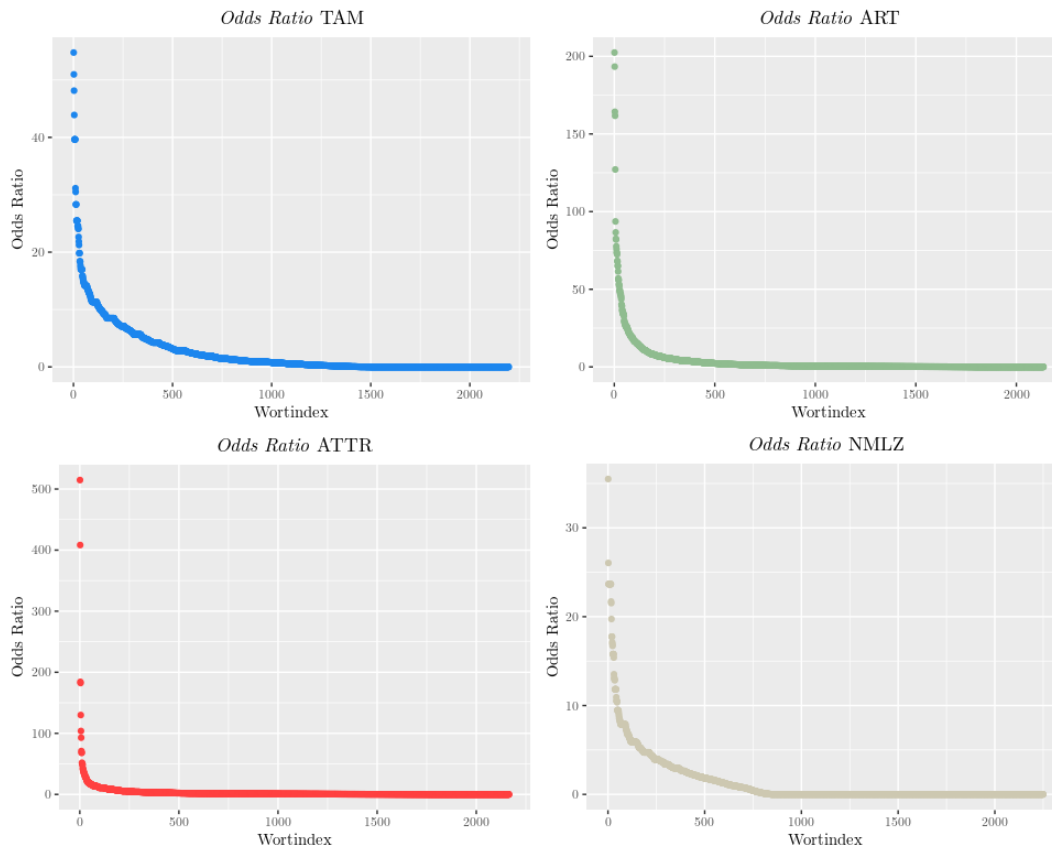


Abbildung 35: Relative Chancen Wörter und Syntagmen

Betrachtet man nun wieder den Zusammenhang mit der Flexibilität von Wörtern, lässt sich allerdings ein Unterschied zwischen den *Odds Ratio*-Werten und den *Attraction*-Werten feststellen (vgl. Abbildung 36, S. 195).

In Abbildung 36 (S. 195) sind in den jeweiligen Diagrammen die *Odds Ratio*-Werte der Wörter auf der x-Achse und die Anzahl der Syntagmen, in denen das entsprechende Wort auftritt, auf der y-Achse eingetragen.

Anders als bei den *Attraction*-Werten steigen die *Odds Ratio*-Werte nicht mit der Anzahl der Syntagmen, in denen ein Wort auftritt (vgl. Tabelle 16, S. 83). Wörter, die in vier Syntagmen auftreten erreichen nicht so hohe Werte wie Wörter, die in zwei oder drei Syntagmen auftreten. Beim Artikel-Syntagma erreichen mehr Wörter, die in zwei Syntagmen auftreten, hohe *Odds Ratio*-Werte als Wörter, die in drei und vier Syntagmen auftreten.

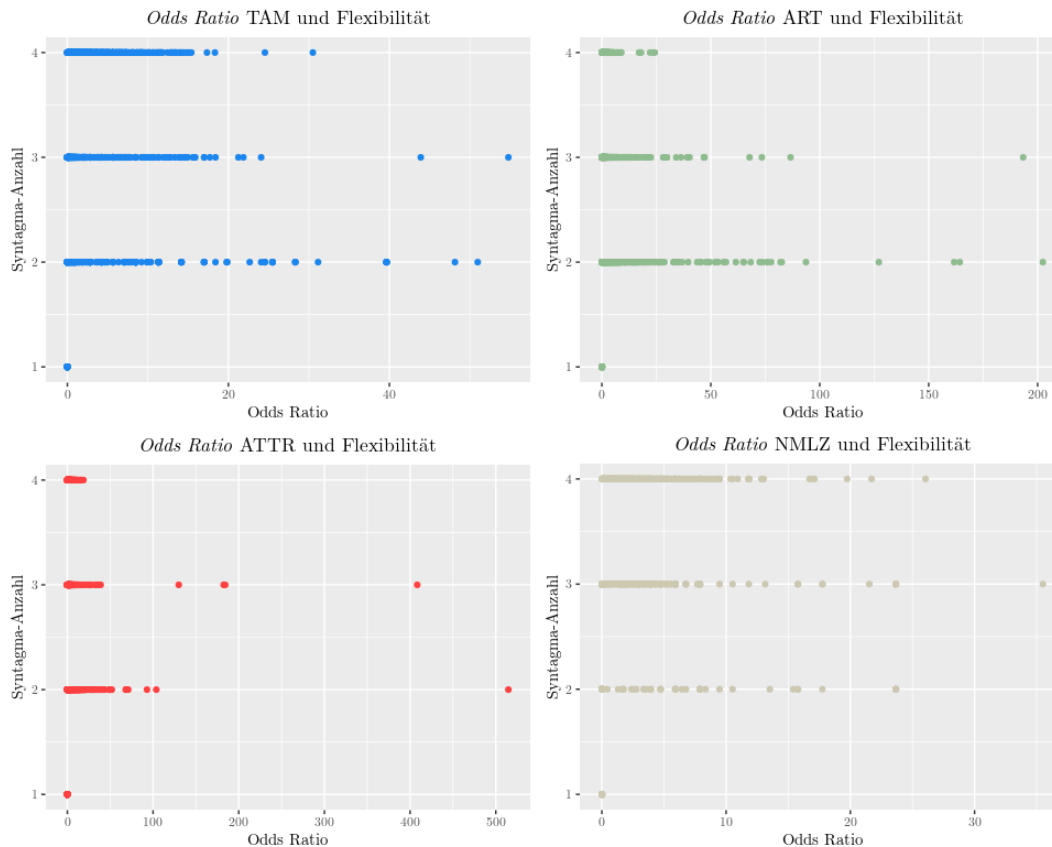


Abbildung 36: *Odds Ratio* und Flexibilität

Ähnliches tritt auch hervor in dem Fall, in dem die „Nominalisierungskonstruktion“ zum Artikel-Syntagma gezählt wird (vgl. Abbildung 37, S. 196).

Unter Beachtung der *Odds Ratio*-Werte würden also eher Wörter, die in zwei Syntagmen auftreten, einer *Worart* zugeordnet werden müssen als Wörter, die ausschließlich in einem Syntagma oder in drei oder vier Syntagmen auftreten. Bei den meisten Wörtern bleibt es aber unmöglich, in klarer Weise eine prototypische Auftretensweise über die *Odds Ratio*-Werte zu bestimmen.

Nachdem mehrere Maße zur Identifikation von Frequenzklassen betrachtet wurden, sollen die Ergebnisse im folgenden Abschnitt noch einmal zusammengefasst werden.

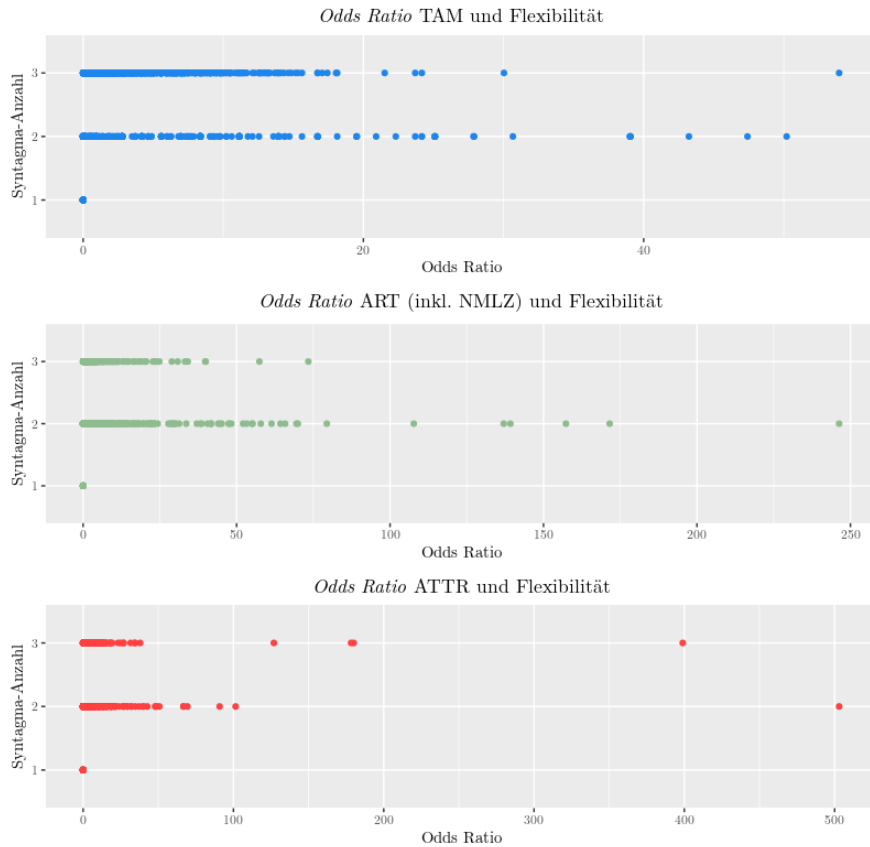


Abbildung 37: Odds und Flexibilität

4.2.3 Zusammenfassung Frequenzklassen

Nachdem die Häufigkeiten, mit denen Wörter im Korpus in Tempus-Aspekt-Modus-Syntagma, Artikel-Syntagma, „Nominalisierungskonstruktion“ und attributivem Slot auftreten, genauer analysiert wurden, muss die Annahme, im Hawaiischen ließen sich Frequenzklassen anstelle von Distributionsklassen bilden, abgelehnt werden.

Zunächst einmal kann festhalten werden, dass sich der Nachweis des prototypischen Gebrauchs von Wörtern über die Auftretenshäufigkeit keineswegs in so direkter Weise durchführen lässt, wie man auf den ersten Blick vermuten könnte. So besteht die Möglichkeit, verschiedene Maße, in denen die Wortfrequenz berücksichtigt wird, für die Stärke des Zusammenhangs von Wörtern und Syntagmen zu verwenden.

Hier wurden die Anteile, mit denen Wörter in einem bestimmten Syntagma auftreten (*Reliance*), die Anteile von Wörtern in einem bestimmten Syntagma

an der Auftretenshäufigkeit des Syntagmas im Korpus (*Attraction*) und die relativen Chancen, mit denen Wörter in einem bestimmten Syntagma auftreten (*Odds Ratio*), betrachtet.

Erstens führen diese Maße in Zusammenhang mit der Flexibilität von Wörtern zu unterschiedlichen Ergebnissen: Während bei den *Reliance*-Werten keine Unterschiede zwischen der Anzahl an Syntagmen, in denen ein Wort auftritt, erkennbar sind, müssten anhand der *Attraction*-Werte Wörter, die in mehr als einem Syntagma auftreten, eher als eine prototypische *Wortart* klassifiziert werden als Wörter, die ausschließlich in einem Syntagma auftreten. Anhand der *Odds Ratio*-Werte hingegen müssten eher Wörter, die in zwei Syntagmen auftreten, als eine prototypische *Wortart* klassifiziert werden als Wörter, die in einem Syntagma, drei oder vier Syntagmen auftreten.

Insgesamt scheint ein Zusammenhang zwischen der Häufigkeit, mit der ein Wort im Korpus auftritt, und der Flexibilität, das heißt der Anzahl an verschiedenen Syntagmen, in denen das Wort auftritt, zu bestehen.

Zweitens lassen die beobachteten Werte keines dieser Maße eindeutig die Bildung von Klassen zu. Wörter treten im Korpus in allen möglichen Häufigkeiten in den verschiedenen Syntagmen auf. Insgesamt kann eher von einem Frequenzspektrum als von klar abgrenzbaren Frequenzklassen gesprochen werden.

Damit ist Frequenz auch nicht als Parameter für die Prädisposition von Wörtern im Sinne des prototypischen Auftretens brauchbar. Wenn das prototypische Auftreten eines Wortes einfach daran gemessen werden soll, dass das Wort häufiger in einem bestimmten Syntagma auftritt, kann letztlich nicht bestimmt werden, ab welchem Wert dieses „häufiger“ auch häufiger genug ist. Ist ein Wort, das zu 90% in Artikel-Syntagma auftritt ein *Nomen*, ein Wort, das zu 60% in Artikel-Syntagma auftritt, hingegen nicht?

Im Kern führt der Versuch, im Hawaiischen über die Auftretenshäufigkeit von Wörtern in einzelnen Syntagmen *Wortarten* zu identifizieren zu keinem überzeugenderen Resultat als der Versuch, über das Auftreten in bestimmten (Kombinationen von) Syntagmen *Wortarten* zu klassifizieren.

Daraus ergibt sich auch die Frage, inwieweit es noch möglich ist, die semantische Klassifizierung von Wörtern zur Bestimmung von *Wortarten* im Hawaiischen heranzuziehen. Im folgenden Kapitel soll dies genauer betrachtet werden.

4.3 semantische Klassen

Dass sich weder Distributionsklassen noch Frequenzklassen im Hawaiischen nachweisen lassen, stellt für die Hypothese, *Wortarten* könnten über die Prädisposition von in ihren lexikalischen Bedeutungsklassen bestimmbar Wörtern definiert werden, eine ernsthafte Herausforderung dar.

Wenn im Hawaiischen potentiell alle Wörter in allen möglichen Häufigkeiten in allen (Kombinationen von) Syntagmen auftreten, stellt sich die Frage, was die semantische Klassifizierung von Wörtern überhaupt zur Identifikation von *Wortarten* beitragen kann.

Im Folgenden werden als Erstes noch einmal unter Berücksichtigung der Ergebnisse aus den bisherigen Untersuchungen die zugrundeliegenden Annahmen bei der semantischen Bestimmung von *Wortarten* dargestellt, um schließlich zwei sich daraus ergebende Definitionsversuche genauer zu analysieren.

4.3.1 Grundannahmen

Betrachtet man nach den vorangegangenen Untersuchungen zu Distributions- und Frequenzklassen (vgl. Kapitel 4.1–4.2, S. 107–196) erneut die in den Kapiteln 2.2.1 (S. 11–14) und 2.2.4 (S. 25–55) besprochenen Arbeiten, lassen sich zusammenfassend zwei verschiedene Ansatzpunkte festmachen, von denen aus die Bestimmung von *Wortarten* über semantische Klassen in den polynesischen Sprachen respektive dem Hawaiischen ausgeht.

Zunächst wird, wie in vielen einzelsprachlichen Beschreibungen, in der hawaiischen Grammatik von Elbert und Pukui (1979) davon ausgegangen, dass primär *Verb* und *Nomen* als Distributionsklassen identifizierbar sind. Die Bestimmung der semantisch-ontologischen Klasse von Wörtern ist dann zwar für die Konstituierung von *Wortarten* nur zweitrangig und wird auch nicht weiter elaboriert. Zumindest implizit liegt aber die Annahme zugrunde, dass sich Distributionsklassen mit semantischen Klassen decken. Die Ansetzung von Distributionsklassen führt dazu, dass Wörter, die sowohl in Tempus-Aspekt-Modus-Syntagma als auch Artikel-Syntagma auftreten, als eigenständige Distributionsklasse behandelt und mit dem Etikett *Verb-Nomen* versehen werden (vgl. Kapitel 4.1.1, S. 112). Aus der grundlegenden Unterscheidung von *Verb* und *Nomen* ergeben sich für die Bestimmung der semantischen Klasse von *Verb-Nomen* zwei Alternativen: Entweder wird aus der jeweiligen semantischen Klasse von *Verb* und *Nomen* eine ausgewählt und *Verb-Nomen* zugeschrieben, was allerdings in Konflikt mit der Ansetzung einer eigenständigen

Distributionsklasse *Verb-Nomen* steht (vgl. Kapitel 4.1.1, S. 113–114), oder die Bedeutung eines *Verb-Nomens* wird ausschließlich im aktuellen Auftreten in einem bestimmten Syntagma als *verbal* oder *nominal* bestimmt. In letzterem Fall wäre die jeweilige semantische Klasse eines *Verbs* und eines *Nomens* definiert, die semantische Klasse eines *Verb-Nomens* als *Wortart* hingegen nicht. Dass überhaupt die Bedeutungsklasse eines Wortes im Auftreten in Tempus-Aspekt-Modus-Syntagma als *verbal* und die Bedeutungsklasse eines Wortes im Auftreten in Artikel-Syntagma als *nominal* bezeichnet werden kann, hängt von der Möglichkeit der distributionalen Unterscheidung zwischen für Tempus-Aspekt-Modus-Syntagma prädisponiertem *Verb* und für Artikel-Syntagma prädisponiertem *Nomen* ab. Wenn nun aber jedes Wort im Hawaiischen potentiell in allen Syntagmen auftreten kann und sich somit keine Distributionsklassen bilden lassen (vgl. Kapitel 4.1.1–4.1.3, S. 108–167), muss auch die Annahme der semantischen Bestimmbarkeit einer Klasse *Verb* und *Nomen* verworfen werden. Einzig über die Identifizierung der semantischen Klassen eines Wortes im aktuellen Auftreten in einem bestimmten Syntagma kann die semantische Klasse einer *Wortart* nicht bestimmt sein, weil dies die Möglichkeit von rein syntaktisch definierten *Wortarten* erfordern würde (vgl. Kapitel 2.2.2, S. 14–16).

Die Untersuchung semantisch-ontologischer Klassen liefert also keinen wesentlichen Beitrag zur Frage der Bestimmung von *Wortarten* im Hawaiischen, solange die Unterscheidung der Distributionsklassen *Verb* und *Nomen* vorausgesetzt ist.

Ein anderer Ansatzpunkt wird dann in den Versuchen gewählt, *Wortarten* unter Verzicht auf die Bildung eindeutiger Distributionsklassen über die prototypischen Auftretensweisen von Wörtern zu definieren. Entscheidend ist dabei die Annahme, dass über die Bestimmung semantischer Klassen sogar das Problem fehlender Distributionsklassen gelöst werden kann, weil die Prädisposition eines Wortes nicht über das ausschließliche Auftreten in bestimmten (Kombinationen von) Syntagmen nachzuweisen ist, sondern über das entsprechend der semantischen Klasse prototypische Auftreten in einem ganz bestimmten Syntagma.

Nun stehen im Hawaiischen aber gar keine unabhängigen Parameter für die Unterscheidung von prototypischem und nicht-prototypischem Fall zur Verfügung. Weder lassen sich unmarkierter und markierter Fall gegenüberstellen (vgl. Kapitel 2.2.4, S. 36–37) noch ist es möglich, Frequenzklassen zu analysieren (vgl. Kapitel 4.2.2, S. 179–197). Die Bestimmung semantisch-ontologischer Klassen

spielt für die Definition von *Wortarten* also auch keine Rolle, solange der Nachweis prototypischer Auftretensweisen vorausgesetzt ist, zumindest wenn dieser über Formklassen oder Frequenzklassen erfolgen soll.

Möchte man weiterhin an einer semantischen Bestimmbarkeit von *Wortarten* im Hawaiischen festhalten, führt das im Kern zu der Annahme, es sei potentiell möglich, über die Klassifizierung der Bedeutung von Wörtern im aktuellen Auftreten in bestimmten Syntagmen auf in der Klassifizierung ihrer Grundbedeutung unterscheidbare Wörter zu schließen.

Die Definition von *Wortarten* zeigt sich dann in zwei Spielarten: Entweder werden die verschiedenen Bedeutungen im aktuellen Auftreten eines Wortes auf eine einzige Grundbedeutung und entsprechend eine einzige semantische Klasse reduziert, sodass jedes Wort, das in mehr als einem Syntagma auftritt, genau einer *Wortart* angehört (z.B. Völkel 2017). Oder die verschiedenen Bedeutungen im aktuellen Auftreten eines Wortes werden jeweils einer Grundbedeutung und entsprechend einer anderen semantischen Klasse zugeschrieben, sodass ein Wort je nach Auftreten in einem bestimmten Syntagma auch einer anderen *Wortart* angehört (z.B. Croft 2001).

Diese Definitionen basieren letztlich auf der Bemühung, jedem Wort eine inhärente, ontologisch klassifizierte Bedeutung zuzuschreiben und die Identifikation von *Wortarten* dadurch zu begründen, dass die jeweilige ontologische Klasse der inhärenten Bedeutung eines Wortes auf ein ganz bestimmtes Syntagma abgedeckt ist bzw. nur eine von mehreren Auftretensweisen eines Wortes seiner inhärenten Bedeutung entspricht.

Für die semantische Definition von *Wortarten* im Hawaiischen stellt sich also die Frage, ob eine inhärent klassifizierte Bedeutung von Wörtern analysiert werden kann, die sich in ganz bestimmter Weise im Auftreten in den verschiedenen Syntagmen zeigt.

Nachfolgend sollen die beiden Spielarten der semantischen Definition von *Wortarten* im Hawaiischen genauer untersucht werden, zunächst der Versuch, bei Wörtern, die in mehreren Syntagmen auftreten, eine einzige inhärente semantische Klasse zu bestimmen, daraufhin der Versuch, je nach Auftretensweise mehrere inhärente semantische Klassen anzusetzen.

4.3.2 Bestimmung einer semantischen Klasse eines Wortes

Die Annahme, in den polynesischen Sprachen könnten Wörter, die in mehr als einem Syntagma auftreten, semantisch auf eine ontologisch klassifizierte

Grundbedeutung reduziert werden, ist insbesondere von Don und Van Lier (2013) und von Völkel (2017) hervorgebracht worden (vgl. Kapitel 2.2.4, S. 40–52), findet sich in ähnlicher Weise aber auch in unversalistischen Ansätzen (vgl. z.B. Haspelmath 2012) und anderen in einzelsprachlichen Arbeiten (z.B. Kieviet 2017 zum Rapanuischen).

Die grundlegende Idee ist dabei, dass entweder die Klasse der inhärenten Bedeutung eines Wortes im Auftreten in verschiedenen Syntagmen unverändert bleibt oder aus verschiedenen Bedeutungen eines Wortes im Auftreten in verschiedenen Syntagmen, eine als die inhärente Bedeutung dieses Wortes identifiziert werden kann, während die anderen Bedeutungen auf diese inhärente Bedeutung zurückführbar sind. Die Zurückführung auf eine inhärente Bedeutung wird dann als „syntactic derivation“ (Don und Van Lier 2013: 57 und passim) bezeichnet. Das bedeutet, dass die Bedeutung eines Wortes regelhaft über die Kombination der Grundbedeutung und den Gebrauch in einem Syntagma beschreibbar sein soll und nicht etwa über Wortbildungsprozesse wie Null-Derivation.

Als inhärente Bedeutungsklassen stellen Don und Van Lier (2013: 59 und passim) *action-denoting-roots* und *object-denoting-roots* einander gegenüber, während Völkel (2017: 447 und passim) von *action/event-denoting terms*, *object-denoting terms* und *property-denoting terms* spricht.

Die Unterscheidung von *Wortarten* basiert im Wesentlichen darauf, dass die Bedeutung eines Wortes im Auftreten in einem ganz bestimmten Syntagma auf eine ganz bestimmte inhärente Bedeutungsklasse reduziert werden kann (vgl. Abbildung 38, S. 202). Lässt sich die Bedeutung eines Wortes im Auftreten in Tempus-Aspekt-Modus-Syntagma auf eine Grundbedeutung zurückführen, die zur ontologischen Klasse EREIGNIS gehört, handelt es sich bei dem Wort um ein *Verb*. Lässt sich hingegen die Bedeutung im Auftreten in Artikel-Syntagma auf eine Grundbedeutung zurückführen, die zur ontologischen Klasse OBJEKT gehört, handelt es sich bei dem Wort um ein *Nomen* und lässt sich schließlich die Bedeutung im Auftreten in attributivem Slot auf eine Grundbedeutung zurückführen, die zur ontologischen Klasse EIGENSCHAFT gehört, handelt es sich bei dem Wort um ein *Adjektiv*.

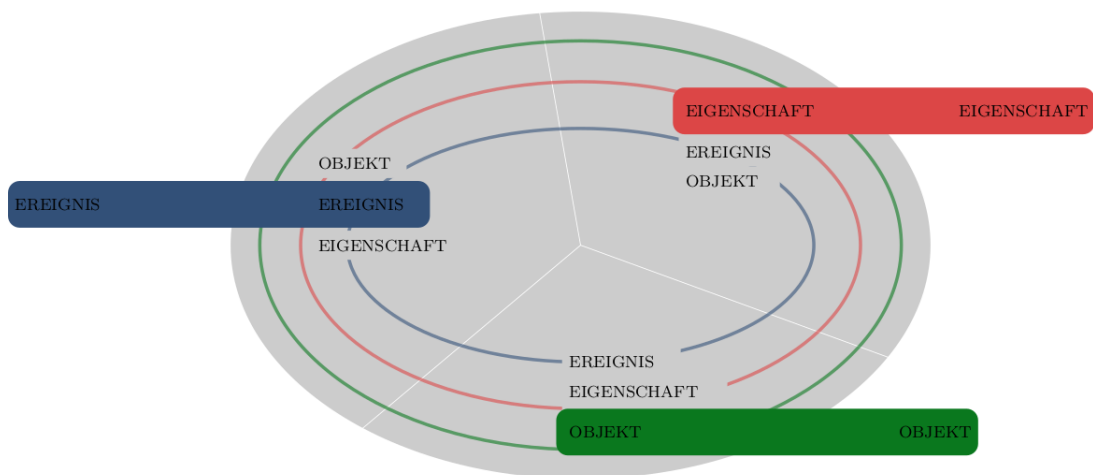


Abbildung 38: Lexikalisierte Grundbedeutung und Auftreten

In Abbildung 38 sind die über die inhärente Bedeutung von Wörtern definierten *Wortarten* zusammengestellt. Während Wörter zwar mit verschiedenen Bedeutungen in jedem Syntagma auftreten, entspricht die Bedeutung im Auftreten in einem bestimmten Syntagma der inhärenten semantischen Klasse eines Wortes.

An dieser Stelle sei angemerkt, dass Don und Van Lier (2013) und Völkel (2017) genau genommen gar nicht von *Wortarten* sprechen. Im Kern versuchen Don und Van Lier (2013) und Völkel (2017) aber die von Croft (2001: 87 und passim) als universell postulierten *typological prototypes* (vgl. Kapitel 2.2.4, S. 26–28) für die polynesischen Sprachen nachzuweisen und zu zeigen, dass Wörter in ihrer inhärent klassifizierten Bedeutung prädisponiert sind für die Bedeutungsprofilierung im Auftreten in verschiedenen Syntagmen. Letztlich sollen sich damit *Wortarten* als rein semantische Klassen unterscheiden lassen, sei es als prototypische Kategorien oder nicht.

Bei genauerer Betrachtung zeigt sich allerdings, dass gar nicht klar ist, wie entschieden werden soll, wann die Bedeutung eines Wortes im Auftreten in einem Syntagma der Grundbedeutung des Wortes entspricht und wann der von einer dem Auftreten in einem anderen Syntagma entsprechenden Grundbedeutung, abgeleiteten Bedeutung, wenn Wörter in allen Bedeutungen in allen Syntagmen auftreten. Daraus geht auch hervor, dass die Grundbedeutung eines Wortes überhaupt nicht bestimmt werden kann.

Möchte man beispielsweise von den Wörtern in Satz (20) die *Wortarten* als semantische Klassen identifizieren, muss man von jedem Wort die inhärente Bedeutung bestimmen und zweifelsfrei entscheiden, ob es sich im jeweiligen Auftreten um die inhärente Bedeutung oder eine von dieser abgeleiteten Bedeutung handelt, vgl.:

(20)	[<i>Ua</i>	<i>hilina'i</i>	<i>aku-la</i>]	' <i>o</i>	<i>Kalani'ōpu'u</i>	
	PERF	vertrauen	DIR-DIST	TOP	Kalani'ōpu'u	
	[<i>i</i>	<i>ke</i>	<i>alaka'i</i>]	[<i>a</i>	<i>kēia</i>	<i>kahuna</i>]
	PREP	DEF	Führung	POSS	DIST	Priester
	[<i>ma</i>	<i>kona</i>	<i>akena</i>	' <i>ana</i>]		
	PREP	3SG	prahlen	NMLZ		
	[<i>e</i>	<i>loa'a</i>	[' <i>i'o</i>]	<i>ana</i>	<i>nō</i>]	
	INTV	besorgen	wirklich	DIST	INTSF	
	[<i>ka</i>	<i>wai</i>	[' <i>ono</i>]].			
	DEF	Wasser	köstlich			

'Kalani'ōpu'u vertraute der Führung dieses Priesters in seinem Prahlen, tatsächlich köstliches Wasser zu besorgen.'

(Desha 1995 [¹1920,1924]:140)

Dabei ist zunächst vorausgesetzt, dass sich die semantische Klasse eines Wortes im aktuellen Auftreten überhaupt eindeutig bestimmen lässt. Soll die Bedeutung eines Wortes im aktuellen Auftreten in einem Syntagma der inhärenten Bedeutung entsprechen, muss man zusätzlich annehmen, dass die Bedeutung dieses Wortes bei jedem Auftreten in diesem Syntagma identisch ist.

Entscheidet man sich nun beispielsweise dafür, die Bedeutung von *hilina'i* in Satz (20) im Auftreten in dem Tempus-Aspekt-Modus-Syntagma *ua hilina'i akula* als EREIGNIS zu interpretieren, stellt sich für die Definition der *Wortart* von *hilina'i* die Frage, ob die angesetzte Bedeutungsklasse in *ua hilina'i akula* auch der inhärenten Bedeutungsklasse von *hilina'i* entspricht oder nicht. Sei weiter angenommen, dass die Bedeutung von *hilina'i* jedes Mal im Auftreten in Tempus-Aspekt-Modus-Syntagma als EREIGNIS klassifiziert werden kann.

Diese Annahme allein reicht nicht aus, um *hilina'i* eine als EREIGNIS klassifizierte Grundbedeutung zuzuschreiben. Dafür muss die Bedeutung von *hilina'i* auch im Auftreten in den anderen Syntagmen berücksichtigt werden, weil ja gerade nicht von vornherein ausgeschlossen sein soll, dass eine als EREIGNIS klassifizierte Bedeutung eines Wortes im Auftreten in Tempus-Aspekt-Modus-Syntagma statt der inhärenten eine abgeleitete Bedeutung darstellt und zur Bestimmung der *Wortart* eben die Identität der aktuellen mit der inhärenten Bedeutung nachzuweisen ist.

Setzt man bei *hilina'i* in Satz (20) ohne Berücksichtigung anderer Auftretensweisen EREIGNIS als inhärente Bedeutungsklasse an, besteht die Gefahr, einfach nur die Kategorien anderer Sprachen, in diesem Fall des Deutschen, als Maßstab der Beschreibung im Hawaiischen heranzuziehen: *hilina'i* lässt sich im Auftreten in Tempus-Aspekt-Modus-Syntagma in Satz (20) in das deutsche Verb *vertrauen* übertragen, also ist es ein *Verb*, zwar nicht über die Flexionsklasse, aber die semantische Klasse definiert (vgl. dazu auch Kapitel 2.2.2, S. 43–46). Würde man *hilina'i* in Satz (20) nun aber mit „zuversichtlich sein“ übersetzen, müsste man die Bedeutung dann in eine Klasse ZUSTAND einordnen (vgl. Henzen 1957: 202–203, Motsch 1999: 181) und eine *nominale* inhärente Bedeutung zugrunde legen?

An dieser Frage soll in etwas provokanter Weise gezeigt werden, dass die Entscheidung über die Klasse der aktuellen Bedeutung eines Wortes keineswegs selbstverständlich ist und mitten in das Problem der Bestimmung der inhärenten Klasse eines Wortes führt.

Zur Erläuterung sei das Auftreten von *hilina'i* in Artikel-Syntagma betrachtet, vgl.:

(21)	<i>'o</i>	<i>ia</i>	<i>ka</i>	<i>hilina'i</i>	<i>nui</i>	<i>o</i>	<i>nā</i>	<i>ali'i</i>
	TOP	3SG	DEF	Vertrauen	groß	POSS	PL	Ali'i
	<i>i</i>	<i>ka</i>	<i>mākaukau</i>	<i>kia'i</i>	<i>o</i>	<i>Kekūhaupi'o</i>		
	PREP	DEF	Fähigkeit	betreuen	POSS	Kekūhaupi'o		

‘Es war das große Vertrauen der Ali'i⁴⁷ in die Vormundschaft von Kekuhapio.’

(Desha 1995 [1920,1924]:79-80)

⁴⁷ *Ali'i* bezeichnet eine herrschende Klasse, in etwa vergleichbar mit dt. „Adelige“.

Für die semantische Analyse von *hilina'i* in Satz (21) (S. 204) im Auftreten in dem Artikel-Syntagma *ka hilina'i nui* gibt es zwei Alternativen. Aus beiden folgt die Ablehnung der Annahme von über die inhärente Bedeutung bestimm- bare *Wortarten*.

Erstens: *hilina'i* wird in Satz (21) (S. 204) mit *hilina'i* in Satz (20) (S. 203) semantisch gleichgesetzt, das heißt die Bedeutung in beiden Auftretensweisen z.B. der Klasse EREIGNIS zugewiesen, und damit auch die inhärente Bedeu- tungsklasse von *hilina'i* als EREIGNIS analysiert. Auf einer solchen Analyse basiert die Annahme, die inhärente Bedeutung eines Wortes könne im Auftre- ten in verschiedenen Syntagmen unverändert bleiben (vgl. S. 201).

Entweder würde dann anders als im Deutschen nicht dem Verb *vertrauen* das durch Konversion abgeleitete *Nomen Actionis Vertrauen* gegenübergestellt, sondern nur ein in seiner Grundbedeutung klassifiziertes *Verb hilina'i* defi- niert, oder auch im Deutschen müsste die Grundbedeutung von *vertrauen* / *Vertrauen* bzw. der Wurzel \sqrt{trau} als inhärent *verbal* und damit *Vertrauen* als semantisches *Verb* analysiert werden.

Diese Behandlung der verschiedenen Auftretensweisen von *hilina'i* führt im Kern wieder zur Definition von *Wortarten* als *typological prototypes* von Croft (2001), einschließlich der inadäquaten Beschreibung des Hawaiischen (vgl. Ka- pitel 2.2.4, S. 25–40): Wenn *hilina'i* in einer als EREIGNIS klassifizierten Grund- bedeutung gleichermaßen in Tempus-Aspekt-Modus-Syntagma und Artikel- Syntagma auftritt, ist eben nicht nachweisbar, dass es sich bei *hilina'i* um ein *Verb* handelt, das aufgrund seiner inhärenten Bedeutung in besonderer Weise in Tempus-Aspekt-Modus-Syntagma gebraucht wird. Man kann dann behaupten, ein „Ereigniswort“ wie *hilina'i* würde *prototypisch* in Tempus- Aspekt-Modus-Syntagma auftreten, Parameter für die *Prototypikalität* lassen sich aber nicht angeben. Und man kann annehmen, die inhärente Bedeutung eines Wortes wie *hilina'i* ließe sich einer bestimmten Klasse zuordnen, für die Bestimmung der *Wortart* von *hilina'i* wäre dies jedoch vollkommen irrelevant. Zweitens: Die Bedeutungsklasse von *hilina'i* im Auftreten in Artikel-Syntagma in Satz (21) (S. 204) wird nicht, oder zumindest nicht vollends, mit der Be- deutungsklasse von *hilina'i* im Auftreten in Tempus-Aspekt-Modus-Syntagma in Satz (20) (S. 203) identifiziert, sondern stattdessen als von der Klasse der inhärenten Bedeutung syntaktisch-regelhaft abgeleitet analysiert. Die Bedeu- tung von *hilina'i* im Auftreten in Artikel-Syntagma kann beispielsweise als Profilierung einer Handlung als Ganzes, als Zustand, als Resultat einer Hand-

lung, oder allgemeiner als *Reifikation* oder Vergegenständlichung, i.e. als OBJEKT (vgl. dazu auch Kapitel 2.2.4, 29–30), betrachtet werden.

Was ist dann aber der primäre Aspekt der Bedeutung von *hilina‘i*? In den verschiedenen Auftretensweisen von *hilina‘i* in Satz (20) (S. 203) und in Satz (21) (S. 204) können unterschiedliche Bedeutungsaspekte profiliert werden, der Begriff an sich muss aber nicht ontologisch klassifiziert sein.

Hinzu kommt, dass *hilina‘i* auch „anlehnen“ oder „Tisch“ bedeuten kann (vgl. Parker 1922: 124, Pukui und Elbert 1986: 70), in den Sätzen (20) (S. 203) und (21) (S. 204) also metaphorisch gebraucht wird. Ob nun *hilina‘i* inhärent die Tätigkeit bezeichnet, die man an einen Tisch vollziehen kann oder das Objekt, an das man sich anlehnen kann, lässt sich nicht entscheiden. Die Alltagserfahrung, in der der Begriff von *hilina‘i* vorsprachlich wurzelt, lässt eben beides zu, ohne inhärent auf eine Lesart festgelegt zu sein. Erdmann (1900: 5 und passim) hat in diesem Zusammenhang den Ausdruck *Popularbegriff* eingeführt und bemerkt dazu: „Worte sind vielmehr Zeichen für ziemlich unbestimmte Complexe von Vorstellungen, die in mehr oder minder loser Weise zusammenhängen. [...] Die Grenzen der Wortbedeutung sind verwaschen, verschwommen, zerfließend.“ (Erdmann 1900: 5). Man kann auch, bemüht man die von Paul (1995 [1880]: 75 und passim) gemachte Unterscheidung von *usueller* und *okkasioneller* Bedeutung, sagen, dass die *usuelle* Bedeutung von *hilina‘i* oder einem anderen Wort, i.e. „der gesamte[] Vorstellungsinhalt, der sich für den Angehörigen einer Sprachgenossenschaft mit einem Worte verbindet“ (Paul 1995 [1880]:75), nicht auf eine einzige ontologische Klasse reduzierbar ist.

Ähnliche grundlegende Überlegungen wie bislang zur inhärenten Bedeutung von *hilina‘i*, können für alle Wörter aus Satz (20) (S. 203) angestellt werden. In Satz (20) (S. 203) tritt z.B. *alaka‘i* in der Bedeutung „Führung“ in dem Artikel-Syntagma *i ke alaka‘i* auf. Potentiell tritt *alaka‘i* aber auch in Tempus-Aspekt-Modus-Syntagma in der Bedeutung „führen“ auf, vgl.:

- (22) *e alaka‘i ai i ka ho‘ohui maoli ‘ana*⁴⁸
 INTV führen ANAPH PREP DEF Annexion tatsächlich NMLZ
 ‘das zur tatsächlichen Annexion führen wird’
 (Silva 2002b:276)

⁴⁸Originaltext: *e alakai ai i ka hoohui maoli ana.*

Versucht man, *alakai'i* über seine semantische Klasse als *Verb* zu definieren, besteht zunächst wieder die Möglichkeit, diese sowohl im Auftreten in Artikel-Syntagma in Satz (20) (S. 203) als auch im Auftreten Tempus-Aspekt-Modus-Syntagma in Satz (22) (S. 206) als EREIGNIS zu bestimmen. Auch hier liegt die Orientierung am Deutschen wieder auf der Hand: *Führung* ist von der verbalen Basis *führen* abgeleitet, also ist *alaka'i*, auch wenn es sowohl in Tempus-Aspekt-Modus-Syntagma als auch Artikel-Syntagma auftritt, „eigentlich“ ein *Verb*, nur eben rein semantisch bestimmbar. Das bedeutet wiederum, dass sowohl bei *Führung* als auch bei *alaka'i* eine inhärente EREIGNIS-klassifizierte Bedeutung zugrunde gelegt werden muss.

Es lässt sich aber weder im Deutschen die Bedeutung von *Führung* noch im Hawaiischen die Bedeutung von *alaka'i* im Auftreten in Artikel-Syntagma auf eine einzige ontologische Klasse reduzieren.

Um dies für das Wort *Führung* zu zeigen, reicht es, eine Auswahl der Bedeutungsangaben im *Digitalen Wörterbuch der deutschen Sprache* (vgl. DWDS) heranzuziehen, vgl.:

1. Leitung, Lenkung

(a) das Führen

(b) diejenigen, die an der Spitze einer Organisation stehen, Leitungsgremium

(c) das Geführtwerden

[...]

3. Leistungsspitze, Vorsprung

[...]

9. [Technik] Einrichtung an einer Maschine, die dafür sorgt, dass ein Maschinenteil, Werkstück sich in einer bestimmten Richtung bewegt; bestimmte Bewegungsrichtung des Werkstücks

Im Deutschen kann *Führung* unter anderem eine nicht weiter spezifizizierte Handlung des Führens bezeichnen, aber auch eine Personengruppe, ein abstraktes Konzept oder, als Fachterminus, eine für bestimmte Zwecke genutzte Einrichtung einer Maschine.

Der Alltagsbegriff *Führung*, oder genauer, der Alltagsbegriff *führen* / *Führung*

ist vorsprachlich aus der Lebenswelt bekannt und kann zwar unterschiedlich semantisch profiliert werden, gehört an sich aber keiner ontologischen Klasse an.

Auch *alaka‘i* im Hawaiischen kann im Auftreten in Artikel-Syntagma unterschiedliche Bedeutungen haben, vgl.:

- (23a) *a hālāwai me ke kuhina ‘o ‘Aiwohikupua,*
 und treffen PREP DEF Minister TOP ‘Aiwohikupua

‘o kona alaka‘i
 TOP 3POSS Berater

‘und traf sich mit (dem) Minister ‘Aiwohikupua, seinem Berater.’
 (Hale‘ole 2001 [1863]:72)

- (23b) *i ho‘oholo ‘ia e nā alaka‘i lāhui*⁴⁹
 PFV entscheiden PASS PREP PL Führer national
 ‘entschieden von den nationalen Führern.’
 (Silva 2003:100)

Anders als in Satz (20) (S. 203) bezeichnet *alaka‘i* in den Sätzen (23a) und (23b) keine Handlung des Führens, sondern Personen, die unterschiedliche Funktionen innehaben. Auch wenn man *alaka‘i* im Auftreten in Artikel-Syntagma in den Sätzen (23a) und (23b) semantisch parallel zum deutschen *Nomens Agentis* analysiert, kann dadurch kein inhärent als EREIGNIS klassifizierter Begriff *alaka‘i* bestimmt und eine „syntaktische Ableitung“ davon in der Bedeutung im Auftreten in Artikel-Syntagma nachgewiesen werden. Berücksichtigt man zudem, dass es sich bei *alaka‘i* um ein Komposition handelt, das zusammengesetzt ist aus *ala* mit den Bedeutungen „Straße“ und „aufwachen, erheben, hervortreten“ (vgl. Parker 1922: 46, Pukui und Elbert 1986: 16) und *ka‘i* mit den Bedeutungen „Richtung, Tanzen, Gesang, Angelhaken“ und „führen, aufheben, das Gehen versuchen oder lernen, tanzen, ziehen“ (vgl. Parker 1922: 241, Pukui und Elbert 1986: 115) wird die Bestimmung der inhärenten Bedeutungsklasse von *alaka‘i* vollends unmöglich.

So wie nicht entscheidbar ist, ob *ala* inhärent ein Objekt bezeichnet, das her-

⁴⁹Originaltext: *i hooholo ia e na alakai Lahui.*

vortritt oder das Hervortreten, das eine Straße charakterisiert und ob *ka‘i* inhärent das Objekt, mit dem geangelt wird, oder die Bewegung mit einem Angelhaken bezeichnet, kann auch *alaka‘i* kein inhärent ontologisch klassifizierter Begriff zugeschrieben werden.

An diesen Beispielen wird deutlich, dass die Annahme, es ließe sich beim Auftreten eines Wortes in verschiedenen Syntagmen zwischen einer der inhärenten Bedeutung des Wortes entsprechenden Bedeutung und einer durch „syntaktische Ableitung“ veränderten Bedeutung unterscheiden, im Grunde vor den selben Problemen steht, wie die Annahme, man könne im Hawaiischen kategorienverändernde Nullderivation analysieren:

Erstens muss die Ableitungsrichtung zwischen den verschiedenen Auftretensweisen eines Wortes eindeutig bestimmt werden können. Einzelne Bedeutungsprofilierungen von Wörtern gegenüberzustellen und z.B. einen Fall wie *hilina‘i* „vertrauen“ → *hilina‘i* „Vertrauen“ analog zur Bildung von *Nomina Actionis* im Deutschen oder einen Fall wie *alaka‘i* „führen“ → *alaka‘i* „Führer“ in Analogie zur Bildung von *Nomina Agentis* im Deutschen zu konstruieren, ist dabei nicht überzeugend.

Bei den meisten Wörtern im Hawaiischen ist es einfach nicht möglich, die vermeintlich richtige Ableitungsrichtung aus mehreren denkbaren Alternativen zu wählen. So kann beispielsweise das Wort *wai*, das in Satz (20) (S. 203) in dem Artikel-Syntagma *ka wai ‘ono* mit der Bedeutung „Wasser“ auftritt, ebenso in verschiedenen Bedeutungen in attributivem Slot und in Tempus-Aspekt-Modus-Syntagma auftreten, ohne dass entscheidbar ist, welche Bedeutung von welcher Bedeutung abgeleitet ist, vgl.:

- (24a) *Eia ho‘i ke ‘ano o ka mahi‘ai ‘ana*
 hier INTSF DEF Methode POSS DEF bewirtschaften NMLZ
- ma [ka ‘āina wai]*⁵⁰
 PREP DEF Land feucht

‘Hier ist die Methode der Bewirtschaftung von feuchtem Land.’
 (Silva 2002a:4)

⁵⁰Originaltext: *eia hoi ke ano o ka mahiai ana ma ka aina wai.*

- (24b) [*wai iho-la*] *ke koko*
 fließen DIR-DIST DEF Blut
 ‘Das Blut floss.’
 (Pukui und Elbert 1986:377)

In Satz (24a) (S. 209) tritt *wai* in attributivem Slot mit der Bedeutung „feucht“ auf. Ist die Ableitungsrichtung nun *wai* „Wasser“ → *wai* „feucht“ oder *wai* „feucht“ → *wai* „Wasser“? Vorstellbar ist beides. Da mit *wai* im Hawaiischen aber alle Flüssigkeiten bezeichnet werden können, die nicht unter den Begriff *kai* „Salzwasser, Flut, am Meer gelegen, salzig, geschmacksarm“ fallen und *wai* in attributivem Slot auch in der Bedeutung „flüssig“ auftritt (vgl. Parker 1922: 612, Pukui und Elbert 1986: 377), muss dann in Analogie zum Deutschen die Ableitungsrichtung *wai* „flüssig“ → *wai* „Flüssigkeit“ analysiert werden? Überzeugend ist dies nicht.

In Satz (24b) tritt *wai* schließlich in Tempus-Aspekt-Modus-Syntagma mit der Bedeutung „fließen“ auf. Daraus ergeben sich noch mehr Möglichkeiten für Ableitungsrichtungen zwischen den verschiedenen Auftretensweisen von *wai*, wie *wai* „fließen“ → *wai* „Flüssigkeit“ → *wai* „flüssig“ oder *wai* „Flüssigkeit“ → *wai* „fließen“ → *wai* „flüssig“. Die Entscheidung über die Ableitungsrichtung ist letztlich nicht möglich.

Zweitens kann die Ableitung einer Kategorie aus einer anderen nicht das konstituierende Merkmal für die Kategorie selbst sein. So wie die Analyse von kategorienverändernder Derivation, setzt auch die Analyse von „syntaktischer Ableitung“ die Bestimmbarkeit von Kategorien voraus.

Hier wird das eigentliche Problem der Prototypikalitätshypothese deutlich. Denn vermutet man zunächst einmal nichts anderes, als dass Wörter im aktuellen Auftreten in einem bestimmten Syntagma eher, i.e. prototypisch, semantisch als EREIGNIS, OBJEKT oder EIGENSCHAFT profiliert werden, sind damit noch keine *Wortarten* bestimmt, da die Prädisposition von Wörtern für das Auftreten in einem bestimmten Syntagma allein dadurch noch nicht nachgewiesen ist. Das Kriterium der Ableitung kann nun aber nicht zur Definition von *Verben* oder *Nomen* herangezogen werden, weil dies schon der Konstituierung einer Klasse *Verben* und einer Klasse *Nomen* bedarf.

Da beispielsweise *Verben* aber nicht als Wörter, die in ihrer Grundbedeutung für das Auftreten in Tempus-Aspekt-Modus-Syntagma prädisponiert sind und auch nicht rein syntaktisch als Wörter, die in Tempus-Aspekt-Modus-Syntagma auftreten, bestimmbar sind, muss die Klasse *Verben* einfach als Klas-

se der prototypischen, aktuellen Wortbedeutungen in Tempus-Aspekt-Modus-Syntagma, vorausgesetzt werden. Ein Wort wie *alaka'i* lässt sich dann aber nicht inhärent als *Verb* klassifizieren, weil *alaka'i* in der Bedeutung „führen“ ja gerade prototypisch in Tempus-Aspekt-Modus-Syntagma und in der Bedeutung „Führer“ prototypisch in Artikel-Syntagma auftritt. Die verschiedenen Auftretensweisen von *alaka'i* ergeben sich eben nicht aus einer wie auch immer klassifizierten Grundbedeutung.

Im Deutschen mag mit einem Verb eher eine Ereignisbedeutung und mit einem Nomen eher eine Objektbedeutung fokussiert werden. Damit ist aber kein zugrunde liegender ontologisch klassifizierter Alltagsbegriff bestimmt, aus dem heraus die Ableitung als Verb oder Nomen gegeben ist.

Im Hawaiischen mögen die Bedeutungen von Wörtern im Auftreten in Tempus-Aspekt-Modus-Syntagma prototypisch als EREIGNIS und von Wörtern im Auftreten in Artikel-Syntagma prototypisch als OBJEKT profiliert sein (vgl. auch Kapitel 2.2.4, S. 29–30). Damit ist aber nicht gezeigt, dass Wörter wie *hilinea'i* oder *alaka'i* inhärent semantisch klassifizierte *Verben* sind. Genau genommen ist damit sogar gezeigt, dass sie es nicht sind, weil die prototypische Bedeutung im Auftreten in Tempus-Aspekt-Modus-Syntagma und die prototypische Bedeutung im Auftreten in Artikel-Syntagma von einer angenommenen inhärenten Bedeutung unabhängig ist.

Im Kern ist die Annahme, die Bedeutung eines Wortes könne auf eine einzige ontologisch klassifizierte Grundbedeutung reduziert werden für die Bestimmung der *Wortart* des Wortes irrelevant, weil mit der wie auch immer klassifizierten inhärenten Bedeutung nichts über das Auftreten des Wortes oder über die Bedeutung des Wortes im Auftreten in bestimmten syntaktischen Slots gesagt ist.

4.3.3 Bestimmung mehrerer semantischer Klassen eines Wortes

Insbesondere Croft (2001) hat die Annahme vertreten, dass das jeweilige Auftreten eines Wortes in einem bestimmten Syntagma einer semantisch-ontologischen Klasse und damit auch genau einer *Wortart* zugewiesen werden kann (vgl. Kapitel 2.2.4, S. 25–40).

Zur Definition von *Wortarten* unterscheidet Croft (2001) zwischen universell prototypischem und universell nicht prototypischem Fall der Auftretensweise von Wörtern, wobei über den prototypischen Fall *Wortarten* bestimmt sind: Ein Wort, dessen Bedeutung im Auftreten in Tempus-Aspekt-Modus-

(2001) also bei einem Wort wie *alaka'i* nicht eine einzige inhärente semantisch-ontologische Klasse und damit eine *Wortart* zu identifizieren, sondern differenziert je nach Auftretensweise mehrere semantische Klassen und damit mehrere prototypische *Wortarten*: *alaka'i* „führen, Führung“ als *Verb* und *alaka'i* „Berater, Führer“ als *Nomen*.

Da aber keine Möglichkeit besteht, im Hawaiischen prototypische und nicht-prototypische Auftretensweise von Wörtern gegenüberzustellen, weil überhaupt keine Parameter dafür zur Verfügung stehen, werden *Wortarten* letztlich einfach als Übereinstimmung der aktuellen Bedeutung mit einer inhärent klassifizierten Bedeutung definiert, ähnlich des Versuchs, jedem Wort eine einzige Grundbedeutung zuzuweisen.

Dabei ist vor allem die unscharfe Trennung zwischen Wort- und Bedeutungsgrenzen bzw. *Wortarten* und *Bedeutungsarten* im Sinne von semantisch-ontologischen Klassen problematisch.

Ist *alaka'i* ein Wort, dessen Bedeutung im Auftreten in verschiedenen Syntagmen unterschiedlich semantisch klassifiziert sein kann oder müssen mit der Unterscheidung zweier *Wortarten* über die unterschiedliche semantische Klassifizierung auch zwei Wörter angesetzt werden? Croft (2001) selbst gibt hier keine klare Antwort. Auf der einen Seite soll ein Wort wie *alaka'i* als ein einziges polysemes Wort angesetzt werden (vgl. Croft 2001: 71 und passim). Auf der anderen Seite soll bei den verschiedenen Auftretensweisen eines Wortes wie *alaka'i* „semantic shift“ analysierbar sein (vgl. Croft 2001: 74 und passim), das heißt, die Ableitung eines als inhärentes Objektwort klassifizierten *Nomens* *alaka'i* aus einem als inhärentes Ereigniswort klassifizierten *Verbs* *alaka'i* oder andersherum.

In einer Polysemie-Analyse von *alaka'i* sehe ich im Wesentlichen keine Gegenposition zur Annahme, dass sich *alaka'i* eben nicht auf einen inhärent klassifizierten Begriff reduzieren lässt, sondern im aktuellen Auftreten semantisch unterschiedlich profiliert wird. Wenn man möchte, kann man die aktuellen Bedeutungen von *alaka'i* im Auftreten innerhalb eines Syntagmas und über die verschiedenen Syntagmen hinweg in semantisch-ontologische Klassen einordnen, mehrere *Wortarten* von *alaka'i* oder mehrere Wörter *alaka'i* mit jeweils zuweisbarer *Wortart* sind damit aber nicht bestimmt.

Genauso kann man im Deutschen die Bedeutungen von *Führung* in verschiedenen Kontexten einmal als EREIGNIS und einmal als OBJEKT klassifizieren. Daraus ergibt sich aber nicht die Möglichkeit, eine *Wortart* *Verb* „führen, Führung“ einer *Wortart* *Nomen* „Führung“ gegenüberzustellen.

Die Analyse von „semantic shift“ ist entweder identisch mit der Analyse von „syntactic derivation“ oder der Analyse von Nullderivation und führt damit wieder zu dem Problem, dass keine Ableitungsrichtung bestimmbar ist und *Wortarten* nicht unabhängig von einer vermeintlichen Ableitung definiert werden können.

Zur Bestimmung von *Wortarten* über semantisch-ontologische Klassen muss die Unterscheidung von Polysemie, „semantic shift“, „syntactic derivation“ und Nullderivation letztlich als hinfällig betrachtet werden, weil im Kern jede klassifizierbare aktuelle Bedeutung ein Wort konstituieren soll. Ein *Wort* als grammatisches Wort ist im Hawaiischen dann nicht über die Form, sondern rein über die Semantik bzw. die aktuelle Bedeutung definiert, also z.B. ein „Ereigniswort“ *alaka‘i* gegenüber einem „Objektwort“ *alaka‘i*.

Unabhängig davon, in welchen Syntagmen eine Wortform auftritt, muss dann für jede einzelne identifizierbare Bedeutungsklasse im aktuellen Auftreten dieser Wortform ein *Wort* und eine *Wortart* differenziert werden.

Diese Analyse kann je nachdem, wie vielfältig die Bedeutungsprofilierungen in den verschiedenen Auftretensweisen ausfallen, zur Unterscheidung sehr vieler *Wörter* und *Wortarten* bei einer Wortform führen.

So z.B. bei ‘*ono*, das in Satz (20) (S. 203) in attributivem Slot mit der Bedeutung „köstlich“ auftritt, dessen aktuelle Bedeutung im jeweiligen Auftreten in einem Syntagma bzw. in mehreren Syntagmen jedoch sehr verschieden sein kann, vgl.:

- (25a) ‘*o* ‘*ono* *nō* *o* *ka* ‘*ai*⁵¹
 TOP Köstlichkeit INTF POSS DEF Essen
 ‘Die Köstlichkeit des Essens’
 (Silva 2002b:178)

- (25b) ‘*a* *me* *kekahi* *mau* ‘*ono* ‘*ē*
 und PREP ART PL Köstlichkeit besonders
 ‘und andere besondere Köstlichkeiten’
 (Desha 1995 [1920,1924]:115)

⁵¹Originaltext: *o ono no o ka ai*

(25c) *ua* **'ono** *kona* *'i'o*⁵²
 PFV köstlich 3SG Fleisch
 'Sein Fleisch ist köstlich.'
 (Silva 2003:74)

(25d) *Pau* *a'e-la* *ho'i* *ka* **'ono** *o* *ka*
 beenden DIR-DIST INSTF DEF Verlangen POSS DEF

wai *a* **'ono** *mai-la* *ho'i* *wau* *i*
 Wasser und begehre DIR-DIST INTSF 1SG PREP

ka *'ai*
 DEF Essen

'Mein Verlangen nach Wasser ist beendet und ich begehre Essen.'
 (Desha 1995 [1920,1924]:299)

In Satz (25a) (S. 214) bezeichnet *'ono* die Eigenschaft des Essens, köstlich zu sein, in Satz (25b) (S. 214) hingegen Dinge, die köstlich sind. In der Analyse von Croft (2001) muss dann sowohl im Hawaiischen als auch Deutschen differenziert werden zwischen einem als „Eigenschaftswort“ bestimmten *Adjektiv* *'ono* bzw. *köstlich* / *Köstlichkeit* und einem als „Objektwort“ bestimmten *Nomen* *'ono* bzw. *Köstlichkeit*. Je nach Interpretation der Bedeutung von *'ono* im Auftreten in Tempus-Aspekt-Modus-Syntagma in Satz (25c) entweder als EIGENSCHAFT „köstlich“ oder als EREIGNIS „köstlich sein“, wird dann noch zusätzlich ein als „Ereigniswort“ bestimmtes *Verb* *'ono* analysiert.

Im Auftreten von *'ono* in Tempus-Aspekt-Modus-Syntagma in Satz (25d) kommt schließlich ein weiteres als „Ereigniswort“ bestimmtes *Verb* *'ono* hinzu, das je nach Interpretation auch das Auftreten in Artikel-Syntagma in Satz (25d) umfasst oder von einem als „Objektwort“ bestimmten *Nomen* *'ono* unterschieden werden kann.

Eine solche Analyse begründet Croft (2001: 72 und passim) letztlich damit, dass jeder Gebrauch eines Wortes bzw. jede voneinander unterscheidbare aktuelle Bedeutung eines Wortes genau eine konventionalisierte und im Lexikon bereitliegende Bedeutung repräsentiert.

Abgesehen von der Frage, ob die verschiedenen Bedeutungsprofilierungen ei-

⁵²Originaltext: *ua ono kona io*

nes Wortes immer in eindeutiger Weise als diskrete Einheiten voneinander getrennt werden können, heißt das konsequent zu Ende gedacht, dass jede einzelne konkretisierte Bedeutung ein *Wort* von einem anderen scheiden muss und jede konkretisierte Bedeutung in der Zuweisung einer ontologischen Klasse eine *Wortart* bestimmt. Dann müsste nicht nur beispielsweise ein *Verb alaka'i* „führen“, ein *Nomen alaka'i* „Führer“ und ein *Nomen alaka'i* „Berater“ oder ein *Verb hilina'i* „vertrauen“, ein *Verb hilina'i* „anlehen“ und ein *Nomen hilina'i* „Tisch“ gegenübergestellt werden, sondern im Einzelnen auch alles, was mit *hilina'i* „Tisch“ bezeichnet werden kann, also ein runder Tisch, ein großer Tisch, ein Holztisch usw.

Geht man stattdessen davon aus, dass nicht jede Bedeutungsprofilierung eines Wortes zugleich mit einer inhärenten ontologischen Klasse gleichzusetzen ist, muss die Annahme einer semantischen Bestimmbarkeit von *Wortarten* im Hawaiischen abgelehnt werden.

4.3.4 Zusammenfassung semantische Klassen

Betrachtet man zusammenfassend die Versuche, semantisch-ontologische Klassen im Hawaiischen zu bestimmen und darüber *Wortarten* zu identifizieren, kann festgehalten werden, dass dabei mehrere Probleme entstehen.

Erstens ist unklar, ob die aktuelle Bedeutung eines Wortes im Auftreten in einem bestimmten Syntagma immer in eindeutiger Weise einer ontologischen Klasse zugeordnet werden kann.

Zweitens lässt sich von der aktuellen Bedeutung eines Wortes nicht auf einen inhärent klassifizierten Begriff schließen, aus dem heraus die *Wortart* eines Wortes determiniert ist.

Drittens konstituieren semantisch-ontologische Klassen alleine keine *Wortarten*.

Man kann versuchen, die aktuelle Bedeutung von Wörtern semantisch zu klassifizieren und man kann sogar davon ausgehen, dass die aktuelle Bedeutung von Wörtern im Auftreten in Tempus-Aspekt-Modus-Syntagma prototypisch als EREIGNIS, die aktuelle Bedeutung von Wörtern im Auftreten in Artikel-Syntagma prototypisch als OBJEKT und die aktuelle Bedeutung von Wörtern im Auftreten in attributivem Slot prototypisch als EIGENSCHAFT klassifiziert ist. Damit ist aber nichts über die inhärent klassifizierte Bedeutung gesagt, sollen die aktuellen Bedeutungen nun auf einen einzigen Begriff reduziert oder jede aktuelle Bedeutung mit einem Begriff gleichgesetzt werden.

Im Kern trägt die Identifikation von semantisch-ontologischen Klassen nichts zur Bestimmung von *Wortarten* im Hawaiischen bei, weil die Prädisposition von Wörtern für das (mehrheitliche) Auftreten in bestimmten Syntagmen darüber nicht nachzuweisen ist.

5 Fazit und Ausblick

In der vorliegenden Arbeit sollte untersucht werden, ob im Hawaiischen, einer flexionslosen Sprache, trotz fehlender Unterscheidbarkeit von Formklassen Wortarten als adäquate Beschreibungsgröße angesetzt werden können.

Die Diskussion um die Generalisierbarkeit des in der griechisch-lateinischen Tradition entwickelten Wortartmodells geht bis auf erste Auseinandersetzungen mit nicht-indogermanischen Sprachen zurück. Zweifel an der Übertragbarkeit dieses Modells auf das Hawaiische und andere polynesishe Sprachen finden sich bereits in frühen Arbeiten aus dem 19. Jahrhundert.

So wenig wie aus der Skepsis in frühen Arbeiten jedoch der Verzicht auf tradierte Begrifflichkeiten folgt, so sehr wird auch in späteren einzelsprachlichen Grammatiken an Wortartkategorien festgehalten.

Neuere typologische Ansätze gehen einen Schritt weiter in dem Anspruch, universell anwendbare Wortartdefinitionen an die Stelle der etablierten Wortarttheorie zu setzen und damit auch *Wortarten* in Sprachen wie dem Hawaiischen bestimmen zu können. Dabei liegt die Annahme zugrunde, es sei möglich, über die Identifizierung ontologisch-semantischer Klassen die prototypische Auftretensweise von Wörtern in bestimmten syntaktisch-funktionalen Slots nachzuweisen. Dass mithilfe dieser Annahme die traditionelle Wortarttheorie überwunden wird, ist nicht überzeugend. Denn im Kern geht es bei der Unterscheidung von Wortarten um die Frage, ob Wörter für den Gebrauch im Satz prädisponiert sind oder nicht. Wenn die Prädisposition von Wörtern aber nicht über die Zugehörigkeit zu morphologischen Paradigmen gezeigt werden kann, müssen unabhängig von der semantischen Klassifizierung andere Parameter zur Anwendung kommen.

Als Erstes wurde überprüft, ob sich im Hawaiischen Distributionsklassen analysieren lassen. Dafür wurde nach der Erarbeitung verschiedener Kriterien und in der Berücksichtigung der Geschichte des Hawaiischen aus bereits vorliegenden Daten in *Ulukau – The Hawaiian Electronic Library* ein 1.121.198 Tokens umfassendes Korpus erstellt und mithilfe von automatisierten Verfahren eine Tabelle erzeugt, in der festgehalten ist, wie häufig Inhaltswörter im Korpus in Tempus-Aspekt-Modus-Syntagma, Artikel-Syntagma, attributivem Slot und „Nominalisierungs konstruktion“ auftreten. Eine erste Auswertung hat dann ergeben, dass Wörter im Korpus in allen möglichen Kombinationen von Syntagmen auftreten und *ad hoc* keine Möglichkeit besteht zu entscheiden, welche genauen Auftretensweisen Distributionsklassen konstituieren. Desweiteren tritt die Mehrheit der Wörter im Korpus nicht ausschließlich in einem Syn-

tagma, sondern in mehreren Syntagmen auf. Damit ist die Vermutung widerlegt, Wörter, die beispielsweise sowohl in Tempus-Aspekt-Modus-Syntagma als auch Artikel-Syntagma auftreten, könnten in Sprachen wie dem Hawaiischen als Sonderfälle behandelt werden.

Nachdem die Wortverteilung im Korpus mit Angaben in Wörterbüchern und dem in der hawaiischen Grammatik von Elbert und Pukui (1979) konstruierten Wortartsystem verglichen wurde, muss schließlich die Annahme, Wortarten ließen sich im Hawaiischen über die Bildung von Distributionsklassen identifizieren, abgelehnt werden, weil jedes Wort potentiell in jedem Syntagma auftreten kann. Ein Wort ist eben nicht allein durch das aktuelle Auftreten in Tempus-Aspekt-Modus-Syntagma als *Verb* oder durch das aktuelle Auftreten in Artikel-Syntagma als *Nomen* klassifiziert, sondern durch die Prädisposition für ein bestimmtes Syntagma. Auch die Rede von *Nominalisierung* ist dann nicht plausibel, wenn gar nicht ausschließlich *Verben* in „Nominalisierungskonstruktion“ und weder derivierte *Nomen* ausschließlich in Artikel-Syntagma, noch *verbale* Basen ausschließlich in Tempus-Aspekt-Modus-Syntagma auftreten.

Als Zweites wurde die Annahme überprüft, dass sich anstelle von Distributionsklassen im Hawaiischen Frequenzklassen analysieren lassen und darüber der prototypische Gebrauch von Wörtern nachweisbar ist. Im Wesentlichen treten dabei ähnliche Probleme wie bei der Bildung von Distributionsklassen auf: Da Wörter im Korpus in allen möglichen Häufigkeiten bzw. Häufigkeitsverhältnissen in verschiedenen Syntagmen vorkommen, können keine klar voneinander abgegrenzten Klassen gebildet werden. Letztlich ist dann auch nicht entscheidbar, welche Auftretensweise den prototypischen Gebrauch eines Wortes repräsentiert, das zu gleichen Anteilen in verschiedenen Syntagmen auftritt. Verschiedene Maße wie *Attraction* oder *Odds Ratio* führen außerdem in Zusammenhang mit der Flexibilität von Wörtern zu unterschiedlichen Ergebnissen. Während mit dem *Attraction*-Maß eher Wörter, die in allen Syntagmen auftreten, als eine *Wortart* klassifiziert werden müssten, gilt dies bei Anwendung der *Odds Ratio* für Wörter, die in zwei Syntagmen auftreten. Insgesamt scheint mit höherer Auftretenshäufigkeit auch der Flexibilitätsgrad von Wörtern im Korpus zu steigen. Das spricht gegen die Annahme, dass Wörter aufgrund einer inhärenten Klassifizierung für das Auftreten in bestimmten (Kombinationen von) Syntagmen prädisponiert sind.

Als Letztes wurde schließlich untersucht, was die Unterscheidung von ontologisch-semantischen Klassen zur Identifizierung von Wortarten noch beitragen

kann, wenn die Prädisposition von Wörtern weder über Formklassen noch über Distributionsklassen oder Frequenzklassen nachweisbar ist. Zunächst muss festgehalten werden, dass mit der Einteilung von Wortbedeutungen in ontologische Klassen nichts über das Auftreten von Wörtern in nur ganz bestimmten Syntagmen gesagt ist. Letztlich kann man zwar versuchen, die aktuelle Bedeutung eines Wortes im Auftreten in einem bestimmten Syntagma einer der ontologischen Klassen EREIGNIS, OBJEKT oder EIGENSCHAFT zuzuweisen – wobei selbst dies ist nicht in jedem Fall eindeutig möglich ist. Auf eine inhärent klassifizierte Bedeutung des Wortes kann so aber nicht geschlossen werden. Die verschiedenen Bedeutungsprofilierungen eines Wortes auf eine einzige lexikalisierte Bedeutungsklasse zu reduzieren ist nicht überzeugend, weil einfach nicht entscheidbar ist, welche der vielen aktuellen Bedeutungen eines Wortes die Grundbedeutung und welche eine davon abgeleitete Bedeutung darstellt. Soll hingegen die jeweilige Bedeutungsprofilierung eines Wortes mit einer inhärenten Bedeutungsklasse gleichgesetzt werden, führt das in der Konsequenz dazu, dass mithilfe jeder bis ins Kleinste differenzierbaren Bedeutungsprofilierung auch ein *Wort* und eine *Wortart* unterschieden werden müsste. Im Kern kann in einer Sprache wie dem Hawaiischen, in der es nicht möglich ist, Wortarten über die Zugehörigkeit zu Flexionsparadigmen zu unterscheiden, die Prädisposition von Wörtern für bestimmte syntaktisch-funktionale Slot auch nicht über Distributionsklassen, Frequenzklassen oder semantisch-ontologische Klassen gezeigt werden. So praktikabel es für einzelsprachliche Beschreibungen sein mag, etablierte Begrifflichkeiten zu verwenden, so wenig führt dies zu einer adäquaten Beschreibung des Hawaiischen oder anderer polynesischer Sprachen. Und so sehr die Typologie einen übereinzelsprachlichen Vergleich zum Ziel hat, so wenig muss jede etablierte Beschreibungsgröße durch den Versuch diverser Modifikationen zu einer universellen Kategorie erhoben werden. Letztlich kann man auch einfach akzeptieren, dass das tradierte Wortartmodell nicht aufgegeben werden muss, die Grenzen seiner Anwendung aber in Sprachen wie dem Hawaiischen liegen.

Literatur

- ABNEY, STEVEN. 1996. Part-of-speech tagging and partial parsing. *Corpus-Based Methods in Language and Speech*, hrsg. von Ken Church, Steve Young, und Gerrit Bloothoof, 118–136. Dordrecht: Kluwer Academic Publishers.
- ALDRIDGE, EDITH. 2009. Minimalist questions for the nominalist analysis of Tagalog syntax. *Theoretical Linguistics* 35.51–62.
- ANDREWS, LORRIN. 1854. *Grammar of the Hawaiian Language*. Honolulu: Mission Press.
- ANDREWS, LORRIN. 1865. *A Dictionary of the Hawaiian Language. To which is append an English-Hawaiian Vocabulary and a chronological Table of remarkable Events*. Honolulu: Henry M. Whitney.
- BAAYEN, HARALD R. 2001. *Word Frequency Distributions*. Dordrecht, Boston, London: Kluwer Academic Publishers.
- BAAYEN, HARALD R. 2008. *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.
- BAKER, MARK. 2003. *Lexical Categories: Verbs, Nouns, and Adjectives*. Cambridge: Cambridge University Press.
- BARNER, DAVID, und ALAN BALE. 2012. No nouns, no verbs: psycholinguistic arguments in favor of lexical underspecification. *Lingua* 112.771–791.
- BAUER, WINIFRED. ⁴2001. *Maori*. London, New York: Routledge.
- BECK, DAVID. 2013. Uni-directional flexibility and the noun-verb distinction in Lushootseed. In Rijkhoff & Van Lier, 185–221.
- BENHAM, MAENETTE K. P., und RONALD H. HECK. 1994. Political Culture and Policy in a State-Controlled Educational System: The Case of Educational Politics in Hawaii. *Educational Administration Quarterly* 30.419–450.
- BESNIER, NIKO. 2000. *Tuvaluan. A Polynesian Language of the Central Pacific*. London, New York: Routledge.
- BESTGEN, YVES. 2014. Inadequacy of the chi-squared test to examine vocabulary differences between corpora. *Literary and Linguistic Computing* 29.164–170.

- BIBER, DOUGLAS. 1993. Representativeness in corpus design. *Literary and Linguistic Computing* 8.243–257.
- BIBER, DOUGLAS, und K. JONES, JAMES. 2009. Quantitative methods in corpus linguistics. *Corpus Linguistics. An International Handbook*, hrsg. von Anke Lüdeling und Merja Kytö, Ausg. 2, 1287–1305. Berlin, New York: Walter de Gruyter.
- BIGGS, BRUCE. 1960. Morphology-Syntax in a Polynesian Language. *The Journal of the Polynesian Society* 69.376–379.
- BIGGS, BRUCE. 1961. The structure of New Zealand Maori. *Anthropological Linguistics* 3.1–54.
- BIGGS, BRUCE. 1969. *Let's Learn Maori*. Wellington: A. H. and A. W. Reed.
- BIGGS, BRUCE. 1971. The Languages of Polynesia. *Linguistics in Oceania*, hrsg. von J. Donald Bowen, 467–505. The Hague, Paris: Mouton.
- BIRD, STEVEN; EWAN KLEIN; und EDWARD LOPER. 2009. *Natural Language Processing with Python*. Sebastopol: O'Reilly.
- BLAISDELL, KEKUNI. 1989. “Hawaiian” versus “Kanaka Maoli” as metaphors. *Hawai'i Review* 13.77–79.
- BLOOMFIELD, LEONARD. ¹¹1973. *Language*. George Allen and Unwin LTD. [1933].
- BLOOMFIELD, LEONARD. 1917. Tagalog texts with grammatical analysis. *University of Illinois Studies in Languages and Literature*, Ausg. 3, 134–316. Urbana, Illinois: University of Illinois.
- BOAS, FRANZ. 1911. Introduction. *Handbook of American Indian Languages*, hrsg. von Franz Boas, Ausg. 1, 1–83. Washington D.C.: Bureau of American Ethnology.
- BOSSONG, GEORG. 1992. Reflections on the history of the study of universals: the example of the *partes orationis*. *Meaning and grammar: Cross-Linguistic Perspectives*, hrsg. von Michel Kefer und Johan van der Auwera, 3–16. Berlin, New York: Mouton de Gruyter.
- BOWKER, LYNNE, und JENNIFER PEARSON. 2002. *Working with Specialized Language. A practical guide to using corpora*. London, New York: Routledge.

- BREZINA, VACLAV. 2018. *Statistics in Corpus Linguistics. A Practical Guide*. Cambridge: Cambridge University Press.
- BRILL, ERIC. 1995. Transformation-based error-driven learning and natural language processing: A case study in part-of-speech tagging. *Computational Linguistics* 21.543–565.
- BROSCHART, JÜRGEN. 1995. *Why Tongan does it differently. Categorical distinctions in a language without nouns and verbs*. (Arbeiten des Sonderforschungsbereichs 282). Köln: Institut für Sprachwissenschaft.
- BROSCHART, JÜRGEN. 1997. Why Tongan does it differently. Categorical distinctions in a language without nouns and verbs. *Linguistic Typology* 1.123–165.
- BRÉAL, MICHEL. 1897. *Essai de sémantique*. Paris: Hachette.
- BUSCHMANN, JOHANN CARL EDUARD. 1839. Vergleichende Grammatik der Südsee-Sprachen und beiläufig des Malayischen Sprachstammes überhaupt. *Über die Kawi-Sprache auf der Insel Java nebst einer Einleitung über die Verschiedenheit des menschlichen Sprachbaues und ihren Einfluß auf die geistige Entwicklung des Menschengeschlechts*, hrsg. von Wilhelm von Humboldt, Dritter Band. Südsee-Sprachen als östlicher Zweig des Malayischen Sprachstammes, 569–1028. Berlin: Druckerei der Königlichen Akademie der Wissenschaften.
- BUSE, J. E. 1963. The structure of Rarotongan nominal, negative, and conjunctival pieces. *Bulletin of the School of Oriental and African Studies* 26.393–419.
- CABLITZ, GABRIELE. 2000. Nominalisation of verbal clauses in Marquesan (Oceanic, French Polynesia). *Proceedings of AFLA 7. The Seventh Meeting of the Austronesian Formal Linguistics Association*, hrsg. von Mairan Klammer, 1–15. Amsterdam: Vrije Universiteit Amsterdam. Department of Linguistics.
- CABLITZ, GABRIELE. 2006. *Marquesan: A grammar of space*. New York: Mouton de Gruyter.
- CANTOS GÓMEZ, PASCUAL. 2013. *Statistical Methods in Language and Linguistic Research*. Sheffield, Bristol: Equinox Publishing Ltd.

- CAPELL, ARTHUR. 1964. Verbal System in Philippine Languages. *Philippine Journal of Science* 93.231–249.
- CARSTENSEN, KAI-UWE ET AL. (Hg.) ³2010. *Computerlinguistik und Sprachtechnologie. Eine Einführung*. Heidelberg: Spektrum Akademischer Verlag.
- CHAPIN, HELEN GERACIMOS. 1984. Newspapers of Hawai'i 1934 to 1903. From "He Liona" to the Pacific Cable. *The Hawaiian Journal of History* 18.47–86.
- CHAPIN, HELEN GERACIMOS. 2000. *Guide to Newspapers of Hawai'i*. Honolulu: Hawaiian Historical Society.
- CHOMSKY, NOAM A. 1957. *Syntactic Structures*. 's-Gravenhage: Mouton.
- CHUNG, SANDRA. 1974. The Syntax of Nominalizations in Polynesian. *Oceanic Linguistics* 12.641–686.
- CHURCHWARD, MAXWELL C. 1953. *Tongan Grammar*. Nuku'alofa: Vava'u Press.
- CHURCHWARD, MAXWELL C. 1959. *Dictionary Tongan - English, English -Tongan*. Nuku'alofa: Government of Tonga Printing Press.
- CHURCHWARD, SPENCER. 1928. On the Origin of the Polynesian Passive. *The Journal of the Polynesian Society* 3.300–3005.
- CLARK, ROSS. 1981. Inside and outside Polynesian nominalizations. *Studies in Pacific Languages and Cultures in Honour of Bruce Biggs*, hrsg. von J. Hollyman und Andrew Pawley, 65–81. Auckland: Linguistic Society of New Zealand.
- COMRIE, BERNARD, und SANDRA A. THOMPSON. ²2007. Lexical nominalization. *Language Typology and Syntactic Description. Volume III: Grammatical Categories and the Lexicon*, 334–381. Cambridge: Cambridge University Press.
- COOK, KENNETH WILLIAM. 1996. The Cia Suffix as a Passive Marker in Samoan. *Oceanic Linguistics* 35.57–76.
- CORMEN, THOMAS H.; CHARLES E. LEISERSON; und RONALD L. RIVEST. 2000. *Introduction to Algorithms*. Cambridge: MIT Press.

- CROFT, WILLIAM. 2001. *Radical Construction Grammar. Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.
- CROFT, WILLIAM. 2005. Word classes, parts of speech, and syntactic argumentation. *Linguistic Typology* 9.431–441.
- DAVIES, JOHN. 1823. *A grammar of the Tahitian dialect of the Polynesian Language*. Tahiti: Mission Press.
- DAVIES, JOHN. 1853. *A Tahitian and English Dictionary with introductory remarks on the Polynesian Language, and a short grammar of the Tahitian dialect*. Tahiti: London Missionary Society Press.
- DAVIS, LISA, HENRY UND MATTHEWSON. 1999. On the functional determination of lexical categories. *Revue québécoise de linguistique* 27.29–69.
- DBEDT STATE OF HAWAII. 2016. *Statistical Report. Detailed Languages Spoken at Home in the State of Hawaii*. Hawaii: Hawaii State Data Center. Research and Economic Analysis Division. Department of Business, Economic Development and Tourism. State of Hawaii.
- DBEDT STATE OF HAWAII. 2018. *Demographic, Social, Economic, and Housing Characteristics for Selected Race Groups in Hawaii*. Hawaii: Hawaii State Data Center. Research and Economic Analysis Division. Department of Business, Economic Development and Tourism. State of Hawaii.
- DESAGULIER, GUILLAUME. 2017. *Corpus Linguistics and Statistics with R. Introduction to Quantitative Methods in Linguistics*. Cham: Springer.
- DESHA, STEPHEN LANGHERN. 1996. *He Mo‘olelo Ka‘ao No Kekūhaupio*. Hilo: Hale Kuamo‘o. [1920,1924].
- DIBBLE, SHELDON (Hg.) 2005. *Ka Moolelo Hawaii*. Hawaiian Language Reprint Series. Buke III. Honolulu: Hawaiian Historical Society. [1838].
- DIK, SIMON. 1978. *Functional Grammar*. Amsterdam: North-Holland.
- DIXON, R. M. W. 2004. *Basic Linguistic Theory. Volume 2 Grammatical Topics*. Oxford: Oxford University Press.
- DIXON, R. M. W, und ALEXANDRA Y AIKHENVALD (Hg.) 2004. *Adjective Classes: A cross-linguistic typology*. Oxford: Oxford University Press.

- DON, JAN, und EVA VAN LIER. 2013. Derivation and categorization in flexible and differentiated languages. In Rijkhoff & Van Lier, 56–89.
- DORDILLON, RENÉ ILDEFONSE. 1904. *Grammaire et Dictionnaire de la Langue des Iles Marquises*. Paris: Institute D’Ethnologie.
- DU FEU, VERONICA. 1996. *Rapanui*. London, New York: Routledge.
- DWDS – DIGITALES WÖRTERBUCH DER DEUTSCHEN SPRACHE. hrsg. v. d. Berlin-Brandenburgischen Akademie der Wissenschaften, <<https://www.dwds.de/d/wb-dwdswb>>, abgerufen am 15.02.2022.
- EBERHARD, SIMONS, DAVID M.; GARY F.; und CHARLES D. FENNING (Hg.) 2021. *Ethnologue: Languages of the World*. Dallas, Texas: SIL International. Online version: <http://www.ethnologue.com/language/haw>.
- ELBERT, SAMUEL H. 1988. *Echo of a Culture. A Grammar of Rennell and Bellona*. Honolulu: University of Hawaii Press.
- ELBERT, SAMUEL H., und MARY KAWENA PUKUI. 1979. *Hawaiian Grammar*. Honolulu: University of Hawaii Press.
- ERDMANN, KARL OTTO. 1900. *Die Bedeutung des Wortes*. Leipzig: Eduard Avenarius.
- EVANS, NICHOLAS, und TOSHIKI OSADA. 2005. Mundari: The myth of a language without word classes. *Linguistic Typology* 9.351–390.
- EVERT, STEFAN. 2006. How Random is a Corpus? The Library Metaphor. *Zeitschrift für Anglistik und Amerikanistik* 54.177–190.
- FAHRMEIR, LUDWIG ET AL. ⁶2007. *Statistik. Der Weg zur Datenanalyse*. Berlin, Heidelberg: Springer.
- FILLMORE, CHARLES J. 1982. Frame semantics. *Linguistics in the Morning Calm*, hrsg. von The Linguistic Society of Korea, 111–137. Seoul: Hanshin.
- FLEISCHER, WOLFGANG, und IRMHILD BARZ. ⁴2012. *Wortbildung der deutschen Gegenwartssprache*. Berlin, Boston: Walter de Gruyter.
- FORBES, DAVID W. 2000. *Hawaiian National Bibliography. 1780 - 1900*, Ausg. 2. Honolulu und Sydney: University of Hawai’i Press und Modern House.

- FORNANDER, ABRAHAM. 1996. *Ancient History of the Hawaiian People to the Times of Kamehameha I*. Honolulu: Mutual. [1880].
- FREGE, GOTTLIB. 1892. Über Begriff und Gegenstand. *Vierteljahresschrift für wissenschaftliche Philosophie* 16.192–205.
- GEERAERTS, DIRK. 1993. Vagueness’s puzzles, polysemy’s vagaries. *Cognitive Linguistics* 4.223–272.
- GIL, DAVID. 2000. Syntactic categories, cross-linguistic variation and universal grammar. In Vogel & Comrie, 173–216.
- GOLDBERG, ADELE E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.
- GOODYEAR-KA‘ŌPUA, NOELANI. 2014. Introduction. *A nation rising: Hawaiian movements for life, land, and sovereignty*, hrsg. von Noelani Goodyear-Ka‘ōpua, Ikaika Hussey, und Erin Kahunaawaika‘ala Wright, 1–33. Durham and London: Duke University Press.
- GREENBERG, JOSEPH H. 1966. *Language Universals, With Special Reference to Feature Hierarchies*. The Hague: Mouton.
- GREENHILL, S. J., und R. CLARK. 2011. POLLEX-Online: The Polynesian Lexicon Project Online. *Oceanic Linguistics* 50.551–559, Online: <https://pollex.shh.mpg.de/>.
- GRIES, STEFAN TH. 2010. Corpus linguistics and theoretical linguistics. A love–hate relationship? Not necessarily... *International Journal of Corpus Linguistics* 15.327–343.
- GRIES, STEFAN TH., und ANATOL STEFANOWITSCH. 2004. Extending collocation analysis: A corpus-based perspective on “alternations”. *International Journal of Corpus Linguistics* 9.97–129.
- HAGENBRUCH, ANDRÉ. 2010. Flache Satzverarbeitung. In Carstensen, 264–280.
- HALE, KENNETH. 1968. Review of Hohepa 1967. *Journal of the Polynesian Society* 77.83–99.
- HALE‘OLE, S. N. 2001. *Ke Ka‘ao Lāieikawai*. Hilo: Hale Kuamo‘o. [1863].

- HARLOW, ROY. 2006. *Māori. A Linguistic Introduction*. Cambridge: Cambridge University Press.
- HASPELMATH, MARTIN. 2007. Pre-established categories don't exist: Consequences for language description and typology. *Linguistic Typology* 11.119–132.
- HASPELMATH, MARTIN. 2012. How to compare major word-classes across the world's languages. *Theories of Everything: In Honor of Ed Keenan*, hrsg. von Thomas Graf et al., (= UCLA Working Papers in Linguistics 17), 109–130. Los Angeles: University of California.
- HENGEVELD, KEES. 1992. *Non-verbal Predication: Theory, Typology, Diachrony*. Berlin, New York: Mouton de Gruyter.
- HENGEVELD, KEES. 2013. Parts-of-speech systems as a basic typological determinant. In Rijkhoff & Van Lier, 31–56.
- HENGEVELD, KEES, und JAN RIJKHOFF. 2005. Mundari as a flexible language. *Linguistic Typology* 9.406–431.
- HENGEVELD, KEES; JAN RIJKHOFF; und ANNA SIEWIERSKA. 2004. Part-of-speech systems and word order. *Journal of Linguistics* 40.527–570.
- HENZEN, WALTER. ²1957. *Deutsche Wortbildung*. Tübingen: Max Niemeyer.
- HIMMELMANN, NIKOLAUS P. 1997. *Deiktikon, Artikel, Nominalphrase. Zur Emergenz syntaktischer Strukturen*. Tübingen: Max Niemeyer Verlag.
- HIMMELMANN, NIKOLAUS P. 2008. Lexical categories and voice in Tagalog. *Voice and grammatical functions in Austronesian languages*, hrsg. von Peter K. Austin und Simon Musgrave, 247–293. Stanford: CSLI.
- HOFFMANN, JOHN. 1903. *Mundari Grammar*. Calcutta: The Secretariat Press.
- HOPPER, PAUL, und SANDRA A THOMPSON. 1984. The Discourse Basis for Lexical Categories in Universal Grammar. *Language* 60.703–752.
- HUA'ŌLELO, KŌMIKE. 1996. *Māmaka Kaiāo. A Modern Hawaiian Vocabulary*. Honolulu: University of Hawai'i Press.
- HUMBOLDT, WILHELM. 1836. *Über die Verschiedenheit des menschlichen Sprachbaues und ihren Einfluß auf die geistige Entwicklung des Menschengeschlechts*. Berlin: Dümmler.

- HUNSTON, SUSAN. 2008. Collection strategies and design decisions. *Corpus Linguistics. An International Handbook*, hrsg. von Anke Lüdeling und Merja Kytö, Ausg. 1, 154–168. Berlin, New York: Walter de Gruyter.
- JACKENDOFF, RAY. 1983. *Semantics and Cognition*. Cambridge, Massachusetts: MIT Press.
- JACOBSEN, WILLIAM H. 1979. Noun and verb in Nootkan. *The Victoria Conference on Northwestern Languages*, hrsg. von Barbara S. Efrat, 83–155. Victoria: British Columbia Provincial Museum.
- JUDD, BERNICE; JANET E. BELL; und CLARE G. MURDOCH. 1978. *Hawaiian Language Imprints 1822 - 1899. A Bibliography*. Honolulu: The Hawaiian Missions Children's Society and The University Press of Hawaii.
- JUDD, GERRIT. 1838. *Anatomia. He Palapala Ia E Hoike Ai I Ke Ano O Ke Kanaka Kino*. Oahu: Mea Paipalapala A Na Misionari.
- JURAFSKY, DANIEL, und JAMES H. MARTIN. 2009. *Speech and Language Processing. An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition*. Englewood Cliffs: Prentice-Hall.
- KALTZ, BARBARA. 2000. Wortartensysteme in der Linguistik. *Morphologie. Ein internationales Handbuch zur Flexion und Wortbildung*, hrsg. von Geert Booij et al., 1. Halbband, 693–708. Berlin, New York: Walter de Gruyter.
- KAMAKAU, SAMUEL MĀNAIAKALANI. 1992. *Ruling Chiefs of Hawaii*. Revised Ed. Honolulu: Kamehameha School Press. Translation of: Ka moololo o Kamehameha I and Ka moololo o na Kamehameha. [1866-71].
- KAMANĀ, KAUANOE, und WILLIAM H WILSON. 2012. *Nā Kai 'Ewalu. Beginning Hawaiian Lessons*. Hilo: Hale Kuamo'o.
- KANAHELE, GEORGE S. 1979. Haw'n Renaissance Grips, Changes Island History. *Ha'ilono Mele* 5.1–12.
- KANAHELE, GEORGE S. 1982. *Hawaiian Renaissance*. Honolulu: Project WAIAHA.
- KANAHELE, GEORGE S. 1986. *Kū kanaka: stand tall*. Honolulu: University of Hawai'i Press.

- KARTTUNEN, LAURI. 2003. Oxford Handbook of Computational Linguistics. In Mitkov, 339–358.
- KAUFMAN, DANIEL. 2009. Austronesian nominalism and its consequences: A Tagalog case study. *Theoretical Linguistics* 35.1–49.
- KAUHI, EMMA KAPŪNOHU‘ULAOKALANI. 1996. *He Mo‘olelo no Kapa‘ahu*. Hilo: Pili Productions.
- KIEVIET, PAULUS. 2017. *A grammar of Rapa Nui*. Berlin: Language Science Press.
- KILGARRIFF, ADAM. 2005. Language is never, ever, ever, random. *Corpus Linguistics and Linguistic Theory* 1/2.263–276.
- KIMURA, LARRY LINDSEY KAUANOE. 1972 - 1988. *Kani‘āina, ulukau.org, Digital Repository of Ka Haka ‘Ula O Ke‘elikōlani College of Hawaiian Language*. University of Hawai‘i at Hilo.
- KINKADE, M. D. 1983. Salish evidence against the universality of noun and verb. *Lingua* 60.25–39.
- KLABUNDE, RALF. 2010. Automatentheorie und formale Sprachen. In Carstensen, 66–93.
- KOPLINIG, ALEXANDER. 2017. Against statistical significance testing in corpus linguistics. *Corpus Linguistics and Linguistic Theory* 15.321–346.
- KREHL, AUGUST (Hg.) 1819-20. *Prisciani caesariensis grammatici opera. i-ii*. Weidmannsche Buchhandlung.
- KRUPA, VIKTOR. 2005. Syntax of Verbal Nouns in Marquesan. *Oceanic Linguistics* 44.505–516.
- KUIPERS, A. H. 1968. The categories verb-noun and transitive-intransitive in English and Squamish. *Lingua* 21.610–26.
- KUYKENDALL, RALPH S. 1938. *The Hawaiian Kingdom. Volume I. 1778-1854. Foundation and Transformation*. Honolulu: University of Hawai‘i Press.
- KUYKENDALL, RALPH S. 1953. *The Hawaiian Kingdom. Volume II. 1854-1874. Twenty critical Years*. Honolulu: University of Hawai‘i Press.

- KUYKENDALL, RALPH S. 1967. *The Hawaiian Kingdom. Volume III. 1874-1893. The Kalakaua Dynasty*. Honolulu: University of Hawai'i Press.
- KÜRSCHNER, WILFRIED. 1996. Die Lehre des Grammatikers Dionysios (Dionysios Trax, Tékhne grammatiké - deutsch [mit griechischem Paralleltext]). *Ancient grammars. Contents and contexts*, hrsg. von Pierre Swiggers und Alfons Wouters, 177–215. Löwen, Paris: Peeters.
- LANGACKER, RONALD W. 1987a. *Foundations of Cognitive Grammar. Volume I: Theoretical prerequisites*. Stanford: Stanford University Press.
- LANGACKER, RONALD W. 1987b. Nouns and Verbs. *Language* 63.53–94.
- LANGACKER, RONALD W. 1991. *Foundations of Cognitive Grammar, Volume II: Descriptive applications*. Stanford: Stanford University Press.
- LAZARD, GILBERT, und LOUISE PELTZER. 2000. *Structure de la Langue Tahitienne*. Paris, Louvain: Peeters.
- LEVIN, BETH. 1993. *English Verb Classes and Alternations: A preliminary investigation*. Chicago: University of Chicago Press.
- LIJFFIJT, JEFREY ET AL. 2016. Significance testing of word frequencies in corpora. *Digital Scholarship in the Humanities* 31.374–397.
- LUUK, ERKKI. 2010. Nouns, verbs and flexibles: implications for typologies of word classes. *Language Sciences* 32.349–365.
- LYONS, JOHN. 1977. *Semantics*. Volume 2. Cambridge: Cambridge University Press.
- MALO, DAVID. 1903. *Hawaiian Antiquities. (Moolelo Hawaii)*. Honolulu: Hawaiian Gazette Co., Ltd. Translated from the Hawaiian by Dr. N. B. Emerson.
- MANNING, CHRISTOPHER D., und HINRICH SCHÜTZE. 1999. *Foundations of Statistical Natural Language Processing*. Cambridge: MIT Press.
- MCENERY, TONY, und ANDREW HARDIE. 2012. *Corpus Linguistics. Methods, Theory and Practice*. Cambridge: Cambridge University Press.
- MIKHEEV, ANDREI. 2003. Text segmentation. In Mitkov, 201–218.

- MITHUN, MARIANNE. 2000. Noun and verb in Iroquoian languages: Multi-categorisation from multiple criteria. In Vogel & Comrie, 397–420.
- MITKOV, RUSLAN (Hg.) 2003. *Oxford Handbook of Computational Linguistics*. Oxford: Oxford University Press.
- MOOKINI, ESTHER K. 1974. *The Hawaiian Newspapers*. Topgallant Publishing Company, Ltd.
- MOSEL, EVEN, ULRIKE UND HOVDHAUGEN. 1992. *Samoan Reference Grammar*. Oslo: Scandinavian University Press.
- MOTSCH, WOLFGANG. 1999. *Deutsche Wortbildung in Grundzügen*. Berlin, New York: Walter de Gruyter.
- MOYSE-FAURIE, CLAIRE. 2016. Referential markers in Oceanic nominalized constructions. *Finiteness and Nominalization*, hrsg. von Claudine Chamoreau und Estrada-Fernández, 171–204. Amsterdam, Philadelphia: John Benjamins.
- MUMM, PETER-ARNOLD. 1995. Generische Bezeichnung. Onomasiologische Aufgaben und ihre Lösungen durch das neuhochdeutsche Artikelsystem. *Sprachwissenschaft* 20.420–467.
- MUTU, MARGARET. 1989. *Aspects of the structure of the ‘Ua Pou dialect of the Marquesan language*. University of Auckland Dissertation.
- MÜLLER, MAX. 1861-64. *Lectures on the Science of Language*. London: Longman, Green, Longman, and Roberts.
- NELSON, MIKE. 2010. Building a written corpus: what are the basics? In O’Keeffe & McCarthy, 53–66.
- NE SMITH, KEAO. 2002. *Tūtū’s Hawaiian and the Emergence of a Neo-Hawaiian Language*. Masterarbeit, University of Hawai‘i.
- NE SMITH, KEAO. 2003. Tutu’s Hawaiian and the Emergence of a Neo-Hawaiian Language. *‘Ōiwi: A Native Hawaiian Journal* 3.67–75.
- NLTK: NATURAL LANGUAGE TOOLKIT. Online: <https://www.nltk.org/>.
- NOGELMEIER, PUAKEA M. 2005. Introduction. In Dibble, xvii–xxi. [1838].

- NOGELMEIER, PUAKEA M. 2010. *Mai Pa‘a I Ka Leo: Historical Voice in Hawaiian Primary Materials, Looking Forward and Listening Back*. Honolulu: Bishop Museum Press and Awaiulu.
- NORDHOFF, SEBASTIAN. 2004. *Nomen/Verb-Distinktion im Guarani*. Arbeitspapiere Köln. (Neue Folge) 48. Köln: Universität zu Köln.
- OSORIO, JONATHAN KAY KAMAKAWIWO‘OLE. 2002. *Dissembling Lāhui: A History of the Hawaiian Nation to 1887*. Honolulu: University of Hawai‘i Press.
- O’KEEFFE, ANNE, und MICHAEL MCCARTHY. 2010. *The Routledge Handbook of Corpus Linguistics*. New York: Routledge.
- PARKER, HENRY H. 1922. *A Dictionary of the Hawaiian Language*. Honolulu: Board of Commissioners of Public Archives.
- PAUL, HERMANN. ¹⁰1995. *Prinzipien der Sprachgeschichte*. Tübingen: Max Niemeyer Verlag. [¹1880].
- PAWLEY, ANDREW. 1966. Samoan Phrase Structure: The Morphology-Syntax of a Western Polynesian Language. *Anthropological Linguistics* 8.1–66.
- PERKINS, JACOB. 2014. *Python 3 Text Processing with NLTK 3 Cookbook*. Birmingham: Packt Publishing.
- PLANK, FRANS. 1984. 24 grundsätzliche Bemerkungen zur Wortart-Frage. *Leuense Bijdragen* 73.489–520.
- POPESCU, IOAN-IOVITZ. 2009. *Word Frequency Studies*. Berlin, New York: Mouton de Gruyter.
- PRATT, GEORGE. 1893. *A Grammar and Dictionary of the Samoan Language*. London Missionary Society.
- PUKUI, MARY, und SAMUEL H ELBERT. ⁶1986. *Hawaiian Dictionary. Hawaiian-English. English-Hawaiian*. Honolulu: University of Hawai‘i Press.
- PUSTEJOVSKY, JAMES, und AMBER STUBBS. 2012. *Natural Language Annotation for Machine Learning*. Sebastopol: O’Reilly.
- PUSTET, REGINA. 2004. Zipf and his heirs. *Language Sciences* 26.1–15.
- PYTHON. Online: <https://www.python.org/>.

- R. Online: <https://www.r-project.org/>.
- REPPEN, RANDI. 2010. Building a corpus: what are the key considerations? In O’Keeffe & McCarthy, 31–38.
- RICHARDS, NORVIN. 2009. Nouns, verbs, and hidden structure in Tagalog. *Theoretical Linguistics* 35.139–152.
- RIJKHOFF, JAN. 2002. Verbs and nouns from a cross-linguistic perspective. *Rivista di Linguistica* 14.115–147.
- RIJKHOFF, JAN, und EVA VAN LIER (Hg.) 2013. *Flexible Word Classes*. Oxford: Oxford University Press.
- ROBINS, R. H. 1966. The Development of the Word Class System of the European Grammatical Tradition. *Foundations of Language* 2.3–19.
- ROBINS, R. H. 1986. The *techne grammaticke* of Dionysius Thrax in Its Historical Perspective: The Evolution of the Traditional European Word Class Systems. *Mot et Parties du Discours*, hrsg. von Pierre Swiggers und Willy Van Hoëcke, 9–37. Leuven: Leuven University Press.
- ROSCH, ELENOR. 1978. Principles of categorization. *Cognition and Categorization*, hrsg. von Elenor Rosch und Barbara Lloyd, 27–48. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- ROSS, R, JOHN. 1972. The Category Squish: Endstation Hauptwort. *Papers from the Eight Regional Meeting of the Chicago Linguistic Society*, 316–328.
- SANDERS, GERALD. 1991. Levelling and Reanalysis in the History of Polynesian Passive Formations. *The Journal of the Polynesian Society* 100.71–90.
- SANTORINI, BEATRICE. 1991. *Part-of-Speech Tagging Guidelines for the Penn Treebank Project*.
- SAPIR, EDWARD. 1921. *Language. An Introduction to the Study of Speech*. New York: Harcourt, Brace.
- SASSE, HANS-JÜRGEN. 1993. Das Nomen - eine universale Kategorie? *Sprachtypologie und Universalienforschung* 46.187–221.
- SASSE, HANS-JÜRGEN. 2001. Scales between nouniness and verbiness. *Language Typology and Language Universals. An International Handbook*, hrsg. von

- Martin Haspelmath et al., Volume 1, 495–509. Berlin, New York: Mouton de Gruyter.
- SCHACHTER, PAUL. ²2007. Parts-of-speech systems. *Language Typology and Syntactic Description. Volume I: Clause Structure*, hrsg. von Timothy Shopen, 3–61. Cambridge: Cambridge University Press.
- SCHEERER, OTTO. 1924. On the Essential Difference between the Verbs of the European and the Philippine Languages. *Philippine Journal of Education* 7.1–10.
- SCHILLER, ANNE ET AL. 1999. *Guidelines für das Tagging deutscher Textcorpora mit STTS. (Kleines und großes Tagset)*.
- SCHLEICHER, AUGUST. 1865. Die Unterschuchung von Nomen und Verb in der lautlichen Form. *Abhandlungen der philologisch-historischen Classe der Königlich Sächsischen Gesellschaft der Wissenschaften*, Band 4, 497–589. Leipzig: S. Hirzel.
- SCHMID, HANS-JÖRG. 2000. *English abstract nouns as conceptual shells. From corpus to cognition*. Berlin, New York: Mouton de Gruyter.
- SCHMID, HANS-JÖRG. ³2016. *English morphology and word formation*. Berlin: Erich Schmidt Verlag.
- SCHMID, HANS-JÖRG, und HELMUT KÜCHENHOFF. 2013. Collostructional analysis and other ways of measuring lexicogrammatical attraction: Theoretical premises, practical problems and cognitive underpinnings. *Cognitive Linguistics* 24.531–577.
- SCHMITT, ROBERT C. 1968. *Demographic statistics of Hawaii: 1778-1965*. Honolulu: The University of Hawai'i Press.
- SCHULZE, WOLFGANG. 2012. Perception, experience and the structure of event images: Modeling the emergence of Parts of Speech from cognitive processes. Vortrag gehalten auf dem 17. LIPP-Symposium, LMU München.
- SCHÖNBERGER, AXEL. 2008. *Die Ars minor des Aelius Donatus*. Frankfurt am Main: Valentia.
- SCHÖNBERGER, AXEL. 2009. *Die Ars maior des Aelius Donatus*. Frankfurt am Main: Valentia.

- SCHÜTZ, ALBERT J. 1994. *The Voices of Eden. A History of Hawaiian Language Studies*. Honolulu: University of Hawai'i Press.
- SEARLE, JOHN R. 1969. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.
- SIDDALL, JOHN WILLIAM (Hg.) 1921. *Men of Hawaii. A Biographical Reference Library, Complete and Authentic, of the Men of Note and Substantial Achievement in the Hawaiian Islands*, Ausg. 2. Revised Ed. Honolulu: Honolulu Star Bulletin Limited.
- SIEWIERSKA, ANNA. 1991. *Functional Grammar*. London: Routledge.
- SILVA, KALENA (Hg.) 2002a. *Ka Ho'olina. The Legacy. Puke Pai 'Ōlelo Hawai'i. Journal of Hawaiian Language Sources*, Ausg. 1(1). Honolulu: Kamehameha Schools Press und University of Hawai'i Press.
- SILVA, KALENA (Hg.) 2002b. *Ka Ho'olina. The Legacy. Puke Pai 'Ōlelo Hawai'i. Journal of Hawaiian Language Sources*, Ausg. 1(2). Honolulu: Kamehameha Schools Press und University of Hawai'i Press.
- SILVA, KALENA (Hg.) 2003. *Ka Ho'olina. The Legacy. Puke Pai 'Ōlelo Hawai'i. Journal of Hawaiian Language Sources*, Ausg. 2. Honolulu: Kamehameha Schools Press und University of Hawai'i Press.
- SILVA, KALENA (Hg.) 2004a. *Ka Ho'olina. The Legacy. Puke Pai 'Ōlelo Hawai'i. Journal of Hawaiian Language Sources*. Honolulu: Kamehameha Schools Press und University of Hawai'i Press.
- SILVA, NOENOE K. 2004b. *Aloha Betrayed. Native Hawaiian Resistance to American Colonialism*. Durham, London: Duke University Press.
- SIMONE, RAFFAELE. 2003. Maṣ dar, 'ismu al-marrati et la frontière verbe/nom. *Estudios ofrecidos al Profesor José Jesús de Bustos Tovar*, hrsg. von José Luis Girón Alconchel et al., Ausg. 1, 901–918. Madrid: Editorial Complutense.
- SINCLAIR, JOHN. 2005. Corpus and Text. Basic Principles. *Developing Linguistic Corpora: A Guide to Good Practice*, hrsg. von Martin Wynne, 1–16. Oxford: Oxbow Books.
- STANNARD, DAVID E. 1989. *Before the horror: the population of Hawai'i on the eve of Western contact*. Honolulu: Social Science Research Institute.

- STEFANOWITSCH, ANATOL. 2020. *Corpus linguistics. A guide to the methodology*. Berlin: Language Science Press.
- STEFANOWITSCH, ANATOL, und STEFAN TH. GRIES. 2003. Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics* 8.209–243.
- STEINTHAL, HEYMANN. 1860. *Charakteristik der hauptsächlichsten Typen des Sprachbaues*. Berlin: Ferd. Dümmler's Verlagsbuchhandlung.
- STEWART WILLIAMS, JULIE. 1993. *Kamehameha the Great*. Honolulu: Kamehameha School Press.
- STEWART WILLIAMS, JULIE. 1996. *Kamehameha Nui*. Honolulu: Kamehameha School Press. Übersetzt von Hana Pau.
- STUBBLEBINE, TONY. 2003. *Regular Expression Pocket Reference*. Sebastopol: O'Reilly.
- SWADESH, MORRIS. 1939. Nootka internal syntax. *International Journal of American Linguistics* 9.77–109.
- TAUMOEOFOLAU, MELENAITE. 1998. *Problems in Tongan lexicography*. Auckland: University of Auckland Dissertation.
- TCHEKHOFF, CLAUDE. 1981. *Simple sentences in Tongan*. Canberra: Pacific Linguistics.
- TCHEKHOFF, CLAUDE. 1984. Une langue sans opposition verbo-nominale: le tongien. *Modèles linguistiques* 6.125–132.
- THOMPSON, RICHARD, und 'OFA THOMPSON. 1992. *The student's English-Tongan and Tongan - English dictionary*. Nuku'alofa: Friendly Islands Bookshop.
- TOGNINI-BONELLI, ELENA. 2001. *Corpus Linguistics at Work*. Amsterdam, Philadelphia: John Benjamins Publishing.
- TRUSK, HAUNANI-KAY. 1993. *From a native daughter; colonialism and sovereignty in Hawai'i*. Honolulu: University of Hawai'i Press.
- TRYON, DARRELL TREVOR. 1970. *Conversational Tahitian*. Canberra: Australian National University Press.

- TUGGY, DAVID. 1993. Ambiguity, polysemy, and vagueness. *Cognitive Linguistics* 4.273–290.
- TUMMERS, JOSE; KRIS HEYLEN; und DIRK GEERAERTS. 2005. Usage-based approaches in Cognitive Linguistics: A technical state of the art. *Corpus Linguistics and Linguistic Theory* 1-2.225–261.
- UHLIG, GUSTAV. 1883. *Dionysii Thracis Ars Grammatica qualem exemplaria vetustissima exhibent. Subscriptis discrepantiis et testimoniis, quae in codicibus recentioribus scholiis erotematis apud alios scriptores interpretem armenium reperiuntur*. Leipzig: Teubner.
- ULLMANN, STEPHEN. 1967. *Grundzüge der Semantik. Die Bedeutung in sprachwissenschaftlicher Sicht*. Berlin: Walter de Gruyter.
- ULUKAU - THE HAWAIIAN ELECTRONIC LIBRARY. Online: <https://ulukau.org>.
- VAN EIJK, J.P., und T HESS. 1986. Noun and Verb in Salish. *Lingua* 69.319–331.
- VAN LIER, EVA. 2016. Lexical flexibility in Oceanic languages. *Linguistic Typology* 20.197–232.
- VOGEL, PETRA M., und BERNARD COMRIE (Hg.) 2000. *Approaches to the Typology of Word Classes*. Berlin: Mouton de Gruyter.
- VON CHAMISSE, ADALBERT. 1837. *Über die Hawaiische Sprache*. Leipzig: Weidmannsche Buchhandlung.
- VONEN, ARNFINN MURUVIK. 2000. Polynesian multifunctionality and the ambitions of linguistic description. In Vogel & Comrie, 479–487.
- VÖLKEL, SVENJA. 2017. Word classes and the scope of lexical flexibility in Tongan. *Studies in Language* 41.445–495.
- WILLIAMS, WILLIAM, HERBERT (Hg.) 1862. *First lessons in Maori*. Wellington: Whitcombe and Tombs.
- WINTNER, SHULY. 2010. Formal Language Theory. *The Handbook of Computational Linguistics and Natural Language Processing*, hrsg. von Alexander Clark, Chris Fox, und Shalom Lappin, 11–43. Singapore: Wiley-Blackwell.
- WONG, LAIANA. 27. Oktober 2002. Kauakūalahale. *Honolulu Star Bulletin*.

- ZIEGENBALG, JOCHEN. 1996. *Algorithmen. Von Hammurapi bis Gödel*. Heidelberg, Berlin: Spektrum Akademischer Verlag.
- ZINSMEISTER, HEIKE. 2010. Korpora. In Carstensen, 482–491.
- ZIPF, G. K. 1935. *The Psycho-Biology of Language*. Boston: Houghton Mifflin.
- ZIPF, G. K. 1949. *Human Behavior and the Principle of the Least Effort. An Introduction to Human Ecology*. New York: Hafner.