

Flexion im Koreanischen

- Allomorphie und Konkatenation -

Soora Kim (Chung-Ang Uni.)

1. Einführung

Die Vielförmigkeit eines Wortes kommt durch die *Allomorphie*¹⁾ einerseits und *Zusammengesetztheit* andererseits zustande. Allomorphie bezieht sich auf die Tatsache, daß ein Morphem in verschiedenen Kontexten seine Gestalt ändert. Unter Zusammengesetztheit versteht man die Tatsache, daß ein Wort aus mehr als einem Bestandteil besteht. Den Prozeß der Zusammensetzung von Bestandteilen bezeichnet man als *Konkatenation*.

Der Flexion koreanischer Verben²⁾ liegen diese zwei Phänomene zugrunde, nämlich die Allomorphie und Konkatenation. Verbstämme werden mit verschiedenen Flexionssuffixen kombiniert, die jeweils die Kategorien subjektbezogene Honorifikation, Tempus, Modus und adressatbezogene Höflichkeit ausdrücken (Konkatenation). Dabei haben

-
- 1) Dieser Artikel behandelt die Allomorphie auf orthographischer Basis, d.h. er betrachtet die Allomorphie-Phänomene, die in der Orthographie wiedergespiegelt werden. Zwar ist hier strenggenommen der Begriff der Allomorphie nicht korrekt; im Rahmen der Aufgabenstellung und in Anbetracht der Tatsache, daß die Orthographie einer Sprache als eine Abbildung der lautlichen Gestalt dieser Sprache betrachtet wird, erscheint die Benutzung dieses Begriffs dennoch angemessen.
- 2) Unter Verb versteht man die Wortart, die im Satz als Prädikat fungiert. Koreanische Adjektive fallen auch unter diese Kategorie, vor allem aufgrund der Tatsache, daß sie auch als Prädikate verwendet und wie die Verben flektiert werden. Sie bilden die semantisch definierte Untergruppe der Zustandsverben.

manche Verbstämme unterschiedliche Formen (Allomorphie). Die Flexion ist sowohl hinsichtlich der Allomorphie als auch der Konkatenation bestimmten Regeln unterworfen.

Dieser Artikel beschreibt die Flexion koreanischer Verben und bildet daraus die Regeln bzw. Muster, denen die Allomorphie-Phänomene und Konkatenation der Allomorphe unterworfen sind, was als die theoretische Grundlage für die Erstellung eines automatischen Wortformererkennungssystems³⁾ genutzt werden.

Eine Voraussetzung zur Erstellung eines Wortformererkennungssystems ist ein Grammatikformalismus, der der adäquaten Beschreibung linguistischer Fragestellungen dient. Um diesen für die Computerlinguistik verwenden zu können, muß es weiterhin möglich sein, daraus effiziente Algorithmen für die Datenverarbeitung zu entwickeln. Diese Anforderungen erfüllt die *Linksassoziative Grammatik* (kurz LAG).

Die LAG ist ein von Hausser (1985) entwickelter Grammatikformalismus, mit dessen Hilfe wohlgeformte sprachliche Äußerungen formal definiert werden können. Die LAG besteht aus einem Lexikon aus Segmenten (in der Morphologie Allomorphe, in der Syntax Wortformen) und Konkatenationsregeln, die die Segmente zu sprachlichen Äußerungen zusammenfügen.

Die Segmente werden dabei immer *linksassoziativ*, also von links nach rechts bzw. in Schreibrichtung der griechisch-lateinischen Sprachen sowie der modernen koreanischen Sprache zusammengefügt. Der Name LAG bezieht sich auf die linksassoziative Ableitungsordnung, die diesem Grammatikformalismus zugrundeliegt. Daher ist die Zusammensetzung

3) Die Aufgabenbereiche und Methoden der automatischen Wortformererkennungssystems werden im nächsten Abschnitt diskutiert.

zeitlinear, d.h. die Segmente werden in genau der Reihenfolge zusammengefügt, in der sie entlang der Zeitachse (also von links nach rechts) auftauchen.

Die Anwendung einer Konkatenationsregel ist dabei an das Einlesen (Anfügen) genau eines Segments gebunden; daher ist LAG *oberflächenkompositional*. Somit ist die Anzahl der Ableitungsschritte durch die Anzahl der Segmente beschränkt, wodurch die Komplexität der Syntax und der Semantik auf der niedrigsten Stufe steht.

Innerhalb dieses Grammatikformalismus wurde eine morphologische Anwendung, nämlich der LA-MORPH-Ansatz (R. Hausser, 1989) entwickelt, der auf der Allomorph-Methode zur Wortformererkennung beruht.

2. LA-MORPH

Aufgrund der Beobachtung, daß die Wörter einer natürlichen Sprache sich konkret in verschiedenen Formen manifestieren, können wir den Begriff *Wort* als eine abstrakte lexikalische Grundeinheit, die in verschiedenen grammatikalischen Formen realisiert sein kann, und den Begriff *Wortform* als die konkrete Realisierung eines Wortes unterscheiden. Dies hat auch den Vorteil, die teilweise uneinheitliche und widersprüchliche Verwendung des Begriffs *Wort* zu vermeiden. Als Repräsentant eines Wortes dient die sog. Grundform, die mit deren Wörterbuchform zu identifizieren ist.

Die automatische Wortformererkennung beschäftigt sich mit der automatischen Erkennung der Form und des Inhalts von Wortformen. Ihr Aufgabenbereich vereint drei Aspekte der Betrachtung von Wortformen: die grammatische Kategorisierung (Zuweisung der Wortart und der spezifischen morphosyntaktischen Eigenschaften an die

Wortform), Lemmatisierung (Zuordnung der Wortform zu ihrer Grundform(en)) und Segmentierung (Zerlegung der Wortform in ihre Bestandteile), wobei ein System zur automatischen Wortformerkenung nicht zwangsläufig alle drei Aspekte berücksichtigen muß.

Ganz allgemein werden für die automatischen Wortformerkenung folgende Komponenten benötigt:

- Lexikon: Für jedes Element der natürlichen Sprache muß in dem Lexikon eine lexikalische Analyse definiert und elektronisch gespeichert werden.
- Erkennungsalgorithmus: Jede Wortform muß von dem System automatisch kategorisiert und lemmatisiert werden.

Methoden der automatischen Wortformerkenung können danach unterschieden werden, ob ihr Analyselexikon aus Wortformen, Morphemen oder Allomorphen besteht. Betrachten wir folgende Analyse.

- 1) a. 고크오 / ㅏ서 ← Zerlegung in die Allomophe
- b. 곱 / 아서 ← Rekonstruktion des Allomorps 고크오 zu dessen Morphem 곱 und des Allomorps ㅏ서 zu dessen Morphem 아서⁴⁾
- c. 곱(VERB) / 아서(ENDING) ← Lemmatisierung und Kategorisierung
- d. 곱(VERB) + 아서(ENDING) ← Konkatenation der Morpheme mit Hilfe von grammatischen Regeln⁵⁾

4) In diesem Beispiel ist provisorisch angenommen, daß -아서 für alle Allomorphe -서, -아서, -어서, -ㅏ서, -ㅑ서 und -여서 als Repräsentant dient.

5) Genauer gesagt, Überprüfung ob es eine Regel gibt, die diese Konkatenation ermöglicht: z. B. bei Wortformen wie *보았어 scheidet die Analyse in diesem 4. Schritt und kann somit dieser Wortform als Ganzes keine morphosyntaktischen Eigenschaften zuweisen und/oder sie als ungrammatisch erkennen, weil es keine Regel gibt, die die Kombination von einem Verbstamm mit einer offenen Silbe auf ㅏ mit einer Endung mit einem Anlaut 어 ermöglichen würde, obwohl die Analyse bis zu den ersten drei Schritten gelungen ist.

Dieser Analyse liegt implizit die Morphem-Methode zugrunde. Diese Methode verwendet ein Analyselexikon aus analysierten Morphemen. In dieser Methode erfordert die Analyse einer unbekanntes Oberfläche *während der Laufzeit* die Schritte (1) Segmentierung, (2) Rekonstruktion, (3) Erkennung der Morpheme und (4) Konkatenation, was einen sehr komplexen Erkennungsalgorithmus zur Folge hat, der ein Nachteil dieser Methode ist. Hier haben Allomorphe nicht den Status voll analysierter grammatikalischer Einheiten, sondern existieren nur als quasi verfälschte Manifestationen der 'eigentlichen' Morpheme. Die Konkatenation findet auf der Ebene der zugeordneten Morpheme - und nicht auf der Ebene der konkret gegebenen Allomorphe - statt. Deshalb verletzt die Morphem-Methode das Prinzip der Oberflächenkompositionalität. Außerdem werden die Morpheme soweit wie möglich zu Konstituenten zusammengesetzt, weshalb diese Methode das Prinzip der Zeitlinearität verletzt. Mathematisch und programmiertechnisch ist diese Methode von einer hohen Komplexität, weil das System alle möglichen Allomorphievorkommen überprüfen muß⁶⁾.

In der Allomorph-Methode wird im Gegensatz dazu das Analyselexikon mit Hilfe eines Regelpakets automatisch *vor dem Beginn der morphologischen Analyse* aus einem Elementarlexikon abgeleitet, das aus den analysierten elementaren Grundformen der offenen bzw. geschlossenen Wortklassen sowie den grammatischen Elementen wie

6) Jedesmal wenn sie ein ㅁ in einer Wortform (etwa in ㅁ와, ㅁ아, ㅁ니) findet, muß sie überprüfen ob das entsprechende Morphem ebenfalls ㅁ- (als reguläres Verb) oder vielleicht ein ㅁ- (als b-irreguläres Verb) oder ㅁ- (als l-irreguläres Verb) enthält, um aus diesem Allomorph - genau genommen einem Kandidat für ein passendes Allomorph - dessen Morphem zu rekonstruieren. Erst wenn diese Frage für alle auf eine Oberfläche potentiell anwendbaren Allomorphie-Phänomene geklärt ist, kann mit den Schritten 3 und 4 begonnen werden.

Suffixen besteht.

Während der Laufzeit stellt das Analyselexikon somit sämtliche Allomorphe⁷⁾ als vorberechnete Formen zur Verfügung. Sie bilden die Grundlage für eine maximal einfache Segmentierung, indem auf die unbekanntenen Wortformen von links nach rechts die passenden Allomorphe abgepaßt werden - *ohne eine Reduktion zu Morphemen*. Die Konkatenation findet auf der Ebene der analysierten Allomorphe mit Hilfe eines zweiten Regelpakets statt.

Die Teile, aus denen die Wortformen zusammengesetzt sind, sind nach der Betrachtungsweise der Allomorph-Methode Allomorphe (und nicht Morpheme)⁸⁾. Der Begriff des Morphems ist eine linguistische Abstraktion, die sich konkret in der Form endlich vieler Allomorphe manifestiert. Ein Morphem wird als Name der zugehörigen Allomorph-Menge definiert. Indem die Konkatenation von links nach rechts stattfindet, d.h. die Allomorphe in genau der Reihenfolge zusammengefügt werden, in der sie entlang der Zeitachse auftauchen, genügt die Allomorph-Methode dem Prinzip der Zeitlinearität. Bei dieser Methode findet die Konkatenation auf der Ebene der analysierten Allomorphe mit Hilfe von Kombi-Regeln statt. Dies entspricht dem Prinzip der Oberflächenkompositionalität. Sie ist somit aus der Sicht der LAG am besten für die automatische Wortformererkennung geeignet.

Die Allomorph-Methode wurde von Hausser (1989) im Rahmen der LAG entwickelt und unter dem Namen LA-MORPH implementiert.

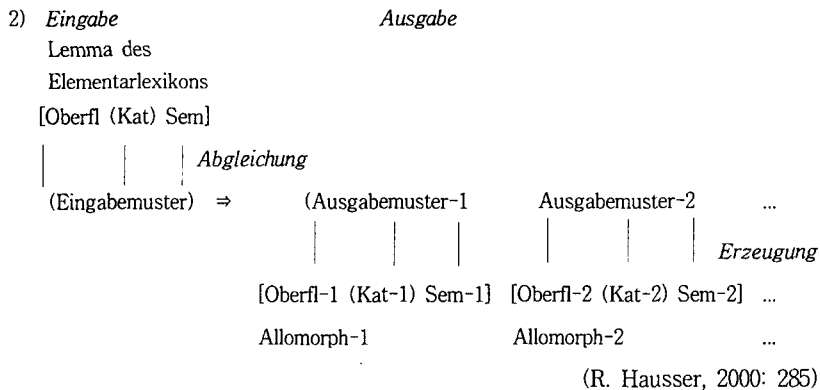
7) d.h. alle Allomorphe des Verbs 갈-, also 갈-, 가- und des Verbs 잇- also, 잇-,이 - usw. sowie die Endungen wie -니, -으니 oder aber auch der Partikeln, Substantive, Pronomen, Adverbien usw., die in den meisten Fällen nur ein einziges Allomorph haben.

8) Dies wird deutlich, wenn wir die Wortformen wie 4) und 5), die in ihre Bestandteile zerlegt sind, betrachten.

LA-MORPH besteht aus folgenden Komponenten:

- einem Elementarlexikon;
- einer Menge von Allo-Regeln, die vor der Laufzeit aus dem Elementarlexikon ein zugehöriges Allomorphlexikon generieren;
- einer Menge von Kombi-Regeln, die die Konkatenation der analysierten Allomorphe während der Laufzeit kontrollieren.

Eine Allo-Regel nimmt ein Lemma des Elementarlexikons als Eingabe und leitet daraus Allomorphe ab. Die Eingabe und Ausgabe werden über Muster definiert. Wenn die Allo-Regelkomponente auf ein Elementarlexikon angewendet wird, durchlaufen alle Lemmata die Allo-Regeln. Wenn ein Lemma auf das Eingabemuster paßt, akzeptiert sie es und leitet die Allomorphe ab. Die Ausgaben der Allo-Regeln werden in eine Datei geschrieben. Auf diese Weise wird aus dem Elementarlexikon automatisch ein Lexikon analysierter Allomorphe abgeleitet. Die Grundstruktur der Allo-Regel ist wie folgt:



Eine Kombi-Regel besteht aus dem Regelname, der Angabe der

kategorialen Operation und dem Regelpaket. Das allgemeine Schema dieser Regeln ist wie im folgenden.

- 3) $r_i: cat_1(\text{Muster des Anfangs}) cat_2(\text{Muster des Nächsten Allomorphs}) \Rightarrow cat_3(\text{Muster des neuen Anfangs}) r_i$

Das obige Regelschema ist zu paraphrasieren als

“Wenn eine Regel r_i ein Eingabepaar erfolgreich in einen neuen Wortformanfang abgebildet hat, wird ein neues ‘nächstes Allomorph’ gesucht. Das nächste Allomorph bildet mit dem gerade abgeleiteten Wortformanfang ein neues Eingabepaar, auf das alle Regeln in dem Regelpaket r_i angewendet werden...” (R. Hausser, 2000: 298)

Wie im ersten Abschnitt erwähnt, zeigt die Flexion sowohl hinsichtlich der Allomorphie als auch der Konkatenation bestimmte Regelmäßigkeiten, die mit Hilfe von Allo-Regeln und Kombi-Regeln implementiert werden sollen. Der nächste Abschnitt betrachtet diese Aspekte der Flexion und ermittelt die Regeln, denen sie unterworfen sind.

3. Flexion der Verben

Koreanische Verben kommen in zahlreichen Formen vor. Die Vielförmigkeit des koreanischen Verbs ist einerseits darauf zurückzuführen, daß Verbstämme mit verschiedenen Endungen kombiniert werden können. Ein anderer Grund ist daß manche Verbstämme und Endungen bei der Konkatenation miteinander unterschiedliche Formen haben.

In der traditionellen Grammatik wie Lee (1989) werden die Verbstämme in regelmäßige und unregelmäßige unterteilt, je nachdem, ob sie bei der Konkatenation mit Endungen unterschiedliche Formen haben. Bei den "unregelmäßigen" Verbstämmen kann man jedoch bestimmte Regelmäßigkeiten finden, die zwischen der Oberfläche der Grundform und der ihrer Allomorphe bestehen. In diesem Sinne spricht man von der Regularitätsgrad der Allomorphie.

3. 1. Allomorphieregularität der Verben

Hinsichtlich der Allomorphie existieren vier Regularitätsgrade, die von drei Fragen abhängen. Diese Fragen sind:

1. Wie viele Allomorphe gibt es zu dieser Grundform? Gibt es überhaupt mehr als eins? (Anzahl der Allomorphe)

2. Wenn es mehr als ein Allomorph gibt: Kann man allein aus der Oberfläche der Grundform entscheiden, wie die Allomorphe gebildet werden? (Universalmuster)

3. Wenn nein: gibt es überhaupt ein Muster? Gibt es andere Grundformen, die ihre Allomorphe auch nach demselben Muster erzeugen? (Klassenmuster)

Tabelle 1 zeigt die Regularitätsgrade, die nach diesen Fragen beurteilt werden können:

	Frage 1 (Anzahl)	Frage 2 (Universalmuster)	Frage 3 (Klassenmuster)
regulär	1	-	-
semi-regulär	>1	ja	-
semi-irregulär	>1	nein	ja
irregulär	>1	nein	nein

Tabelle 1. Regularitätsgrade der Allomorphie-Phänomene

Betrachten wir folgende Wortformen. In runden Klammern sind jeweils die Allomorphe des Verbs aufgelistet. Die Grundform wird als erstes Allomorph (Allo1), und die anderen Allomorphe, die aus der Grundform abgeleitet sind, als zweites (Allo2), drittes (Allo3) usw. bezeichnet. Innerhalb derselben Alternationsklasse bezeichnet die gleiche Nummerierung der Allomorphe das gleiche Alternationsmuster, so daß beispielsweise alle zweite Allomorphe der “b-irregulären” Verben durch die Anwendung der gleichen Alternationsregel abgeleitet werden, die etwa besagt, den letzten Buchstaben ㅂ wegzulassen und 오 einzufügen:

4) Wortform	Allo1	Allo2	Allo3
a. 자/다 자/니 자/서	(자-	-	-)
b. 잡/다 잡/으니 잡/아서	(잡-	-	-)

In beiden Wortformen in 4) erscheint jeweils nur ein einziges Allomorph, das die gleiche Oberfläche hat wie die Grundform: ihre Allomorphie ist also regulär. Tabelle 2 zeigt diese Allomorphie:

Alternationstyp	Muster der Grundform	Muster der Allomorphbildung			
		Grundform	Allo1	Allo2	Allo3
Regulär (keine Alternation)	-	잡-	잡-	-	-

Tabelle 2. Reguläre Allomorphie der Verbstämme

In folgenden Wortformen kommt ein anderes Allomorphie-Phänomen vor:

5) Wortform	Allo1	Allo2	Allo3
a. 갈/다 가/니 갈/아서	(갈-	가-	-)
b. 빨강/다 빨가/니 빨기/서	(빨강-	빨가-	빨기-)

- c. 하/다 하/니 하/여서 흥/내서 (하- 흥- -)
- d. 쓰/다 쓰/니 쓰/켜서 (쓰- 쓰- -)

In obigen Wortformen gibt es mehrere Allomorphe zu jeder Grundform, die aus der Oberfläche der Grundform gebildet werden können: sie zeigen also semi-reguläre Allomorphie. Das jeweilige Muster der Allomorphbildung kann man als l/NULL-Alternation (kurz l(-)S(emi-)R(egulär))(5a), h/NULL/ay-Alternation (kurz h(-)S(emi-)R(egulär)) (5b), ha/h-Alternation (kurz ha(-)S(emi-)R(egulär))(5c), u/NULL-Alternation (kurz u(-)S(emi-)R(egulär))(5d) bezeichnen. Die jeweiligen Alternationstypen sind in der Tabelle 3 dargestellt.

Betrachten wir jetzt folgende Wortformen:

Alternationstyp	Muster der Grundform	Muster der Allomorphbildung			
		Grundform	Allo1	Allo2	Allo3
ISR	*ㄹ	갈-	갈-	가- (Wegfallen des letzten Buchstaben ㄹ)	-
haSR	하	하-	하-	흥- (Wegfallen des letzten Buchstaben ㅏ)	-
hSR	*ㅎ (Zustandsverb)	노랑-, 빨강-	노랑-, 빨강-	노라, 빨거 (Wegfallen des letzten Buchstaben ㅎ)	노르-, 빨거- (Wegfallen der letzten zwei Buchstaben)
uSR	*ㅡ	쓰-	쓰-	쓰- (Wegfallen des letzten Buchstaben ㅡ)	-

Tabelle 3. Semi-reguläre Allomorphie der Verbstämme

- 6) Wortform Allo1 Allo2 Allo3
- a. 돕/다 도우/니 도오⁹/켜서 (돕- 도우- 도오-)
- b. 덩/다 더우/니 더우/켜서 (덩- 더우- -)

c. 잣/다 자/으니 자/아서	(잣- 자- -)
d. 다르/다 다르/니 달르/ㅓ서	(다르- 달르- -)
e. 걸/다 걸/으니 걸/어서	(걸- 걸- -)

Bei den obigen Wortstämmen existieren jeweils mehrere Allomorphe, wobei man anhand der Oberfläche der Grundform selbst nicht vorhersagen kann, wie die Allomorphe gebildet werden¹⁰⁾. Es gibt jedoch ein Muster für den jeweiligen Allomorphbildungstyp und mehrere Grundformen bilden ihre Allomorphe nach diesem Muster: sie zeigen also semi-irreguläre Allomorphie.

Den jeweiligen Allomorphbildungstyp kann man b/wu/o- bzw. b/wu-Alternation (kurz b(-)S(emi-)I(irregulär))(6a-b), s/NULL-Alternation (kurz s(-)S(emi-)I(irregulär))(6c), lu/l-Alternation (kurz lu(-)S(emi-)I(irregulär))(6d), d/l-Alternation (kurz d(-)S(emi-)I(irregulär))(6e) nennen. Diese Allomorphie stellt die Tabelle 4 dar.

9) Die Zerlegung auf diese Art und Weise kann sprachhistorisch begründet werden. Der intervokalische Auslaut ㅃ in dem Verbstamm wurde im frühen Mittelkoreanischen vor einer darauffolgenden vokalisch anlautenden Endung zu einem stimmhaften, bilabialen Reibelaut β, danach graphisch mit dem Zeichen ㅃ fixiert. Im Neukoreanischen tritt dieser Laut nur noch als Halbvokal /w/ im Verbstamm, aber graphisch nicht in Erscheinung. Nur in Verbindung mit dem folgenden Vokal - hier kommt allein die Endung -아(서)/-어(서) in Betracht - wird das bilabiale /w/ graphisch zum Endungsanlaut gezogen und durch 오 bzw. 우 wiedergegeben.

10) Dies bedeutet im konkreten Fall, daß nicht alle Grundformen, die gleiche Muster haben, ihre Allomorphe auf gleiche Art und Weise bilden. Diese Tatsache kann man feststellen, wenn man die Allomorphbildung von Verben der regulären Allomorphie wie 잡-, 벗-, 받- einerseits und die von semi-irregulären Verben wie 돕- (6a), 잇- (6b), 들- (6d) andererseits vergleicht.

Aternationstyp	Muster der Grundform			Muster der Allomorphbildung			
				Grundform	Allo1	Allo2	Allo3
bSI	letzter Vokal	ㅏ, ㅑ	ㅂ	뽁-	뽁-	도우- (Wegfallen des letzten Buchstaben ㅂ und Hinzufügen von 우)	도오- (Wegfallen des letzten Buchstaben ㅂ und Hinzufügen von 오)
		restliche		땀-	땀-	더우- (Wegfallen des letzten Buchstaben ㅂ und Hinzufügen von 우)	-
sSI	*ㅅ			좃-	좃-	자- (Wegfallen des letzten Buchstaben ㅅ)	-
luSI	*르			다르-	다르-	달르- (Wegfallen des letzten Buchstaben ㄹ und Hinzufügen des 르)	-
dSI	*ㄷ			견-	견-	겉- (Wegfallen des letzten Buchstaben ㄷ und Hinzufügen des 르)	-

Tabelle 4 Semi-irreguläre Allomorphie der Verbstämme

Wie bis jetzt beschrieben wurde, unterliegt das Allomorphie-Phänomen bestimmten Regeln bzw. Mustern. Nicht nur die Allomorphie sondern auch die Konkatenation der Verbstämme mit den Endungen folgt bestimmten Mustern. Die Ermittlung der Muster der Konkatenation wird dadurch verkompliziert, daß alle Endungen zu einem bestimmten Allomorphbildungs- und Anschlußtyp gehören, nach dem sie einerseits Allomorphe bilden und andererseits an den Verbstamm angefügt werden.

3. 2. Anschlußtypen der Endungen

Beobachten wir die folgenden Wortformen, wobei die Allomorphe der Endungen in Klammern aufgelistet sind:

- | | |
|----------------------|----------------|
| 7) a. 하/다, 들/다, 잇/다 | (-다) |
| b. 하/니, 들/으니, 이/으니 | (-니, -으니) |
| c. 가/서, 자/아서, 도오/아서, | (-서, -아서, -아서, |
| 들/어서, 더우/아서, | -어서, -아서, |
| 하/여서 흥/아서 | -여서, -아서) |

Die Bestandteile, die in den obigen Wortformen nach dem Verbstamm stehen, sind aufgrund ihrer Funktion und Bedeutung in gemeinsamen Gruppen zu sortieren: sie sind Allomorphe einer Endung. Wenn wir den Begriff der Grundform als den Repräsentant mehrerer Allomorphe erweitern, können wir sagen, daß die Allomorphe der Endung innerhalb einer Gruppe auch aus einer "Grundform"¹¹⁾ abzuleiten sind. Bei den zerlegten Wortformen werden vor allem die graphischen Varianten -아서 von -아서 und -아서 von -어서 beobachtet(7c)¹²⁾. Diese graphischen Varianten sind durch die Eigenschaft der Silbenbildung im Koreanischen entstanden, die im folgenden kurz beschrieben wird.

Silbenstruktur kann unter zwei Aspekten betrachtet werden, dem graphischen Aspekt und dem phnologischen. Unter graphischem Aspekt sind folgende Silbenstrukturen möglich, wobei K einen Konsonanten und

11) Bei den Endungen ist die Frage, welcher von verschiedenen Allomorphen einer Gruppe als deren Repräsentant bzw. deren Name zu sehen ist, mehr oder weniger künstlich. Diese Problematik wird dadurch vermieden, daß man das nach dem vokalisch auslautenden Stamm auftretende Allomorph als Grundform nimmt und die andere durch das Anfügen des Bindevokals vor der Grundform ableitet, wie wir später sehen werden.

12) Die traditionelle Grammatik achtet nicht auf die Existenz dieser Variante und vermeidet diese Problematik dadurch, indem sie eine (meistens implizit vorgenommene) Transformation von -아서 bzw. -어서 zu -아서 bzw. -아서 bei der Kombination mit entsprechenden Stämmen annimmt, was aber das Oberflächenkompositionalitätsprinzip verletzt.

V einen Vokal bezeichnet:

- 8) a. KV: 너/ne/, 소/so/
 b. KVK: 산/san/, 물/mwul/
 c. KVKK: 녀/negs/, 얹/yalb/

Unter phonologischem Aspekt sind folgende Silbentypen vorhanden:

- 9) a. V: /i/, (keine entsprechende graphische Silbenstruktur)
 b. KV: 너/ne/
 c. VK: /os/ (keine entsprechende graphische Silbenstruktur)
 d. KVK: 산/san/, 물/mwul/

Wie in den obigen Beispielen erwähnt wurde, zeigt die Gegenüberstellung mit der graphischen Silbenstruktur bzw. den graphischen Silbentypen Nicht-Übereinstimmung im folgenden Punkt: Unter phonologischem Aspekt existieren vokalisch anlautende Silbentypen V und VK, unter graphischem Aspekt aber nicht. Dies hat historische Ursachen und erklärt sich daraus, daß vokalisch anlautende Silben den Konsonanten ㅇ/ng/, jedoch nicht als Lautzeichen, sondern nunmehr als obligatorisches und stummes Rechtschreibungszeichen, zum Anfangsgraphem haben(10a). Das Zeichen "ㅇ", das als Anfangsgraphem verwendet wird, fällt weg, wenn das darauffolgende Vokalzeichen auf ein konsonantisches (Anfangs-)Graphem tritt(10b). So gesehen sind ㅇ| und | beide graphische Variante zur Wiedergabe des Phonems /i/:

- 10) a. ㅇ|i/'dies' (graph. Silbenstruktur: KV, phon. Silbenstruktur: V)
 b. 바|/bi/'Regen'(graph. Silbenstruktur: KV, phon. Silbenstruktur: KV)

In den Wortformen wie 들르/ㄹ vs. 접/어 einerseits und 갈르/ㅏ vs. 잡/아 andererseits sind also ㄹ und 아 bzw. ㄹ und 어 graphische Varianten zur Wiedergabe des gleichen Phonems /e/ bzw. /a/. Dies führt dazu, daß bei bestimmten Endungen graphische Varianten innerhalb phonologisch bedingten Allomorphen unterschieden werden müssen.

Dieser Artikel geht von dem Ansatz aus, daß bei Endungen mit mehr als einem Allomorph diejenige als Grundform angenommen wird, die ohne Bindevokal 아, 어 usw. auftritt. Aus dieser Grundform werden verschiedene Allomorphe mit und ohne Bindevokal abgeleitet. Die Endungen, die den gleichen Bindevokal haben, haben auch beim Anschluß an den Verbstamm gleiche Distribution. Mit anderen Worten decken sich die Allomorphbildungstypen und Anschlußstypen miteinander. Man kann 3 Allomorphbildungs- und Anschlußstypen (kurz AT)¹³⁾ unterscheiden.

AT 1: Ein Teil der zahlreichen Verbalendungen des Koreanischen darf nur der Grundform folgen, und zwar unabhängig davon, ob sie auf Vokal oder Konsonant endet. Ein Beispiel dafür ist die Satzschlußendung g¹⁴⁾ -다:

- 11) a. 쓰(Allo1)¹⁵⁾-다
b. 빨강(Allo1)-다

13) Ein ähnlicher Ansatz ist in Lewin (1970) zu finden. Er klassifiziert auch die Endungen nach ihrer Distributionsumgebung in 3 Anschlußstypen, in die sich alle Endungen bis auf wenige Ausnahmen einordnen lassen.

14) Die Klassifikation der Endung bzw. deren Bezeichnung in diesem Artikel folgt der traditionellen Grammatik.

15) In den Beispielen wird bei dem ersten Allomorph, also dem Allomorph, das die gleiche Form hat wie dessen Grundform, die Angabe der Grundform unterlassen.

Tabelle 5 und 6 stellen jeweils die Allomorphbildungsmuster dieser Endungen und ihre Distribution beim Anschluß an den Verbstamm dar.

Grundform	Allomorph
-다	-다

Tabelle 5. Allomorphbildung der Endngen des AT1 und AT1a

Alternationstyp	Muster der Grundform	Allo1	Allo2	Allo3
Alle	beliebig	*/다		

Tabelle 6. Tabelle Kombinationstabelle der Endung des AT 1

Eine andere Gruppe von Endungen wie die Attributendung -는 ist identisch mit den Endungen des AT 1 bis auf den Umstand, daß sie bei ISR-Verbstämmen ausnahmslos an das Allo2 antritt. Diese Gruppe nennen wir AT 1a. Tabelle 7 stellt ihre Distribution dar.

Alternations-typ	Muster der Grundform	Allo1	Allo2	Allo3
ISR	*ㄹ	달	다/는	-
restliche	beliebig	걸/는	-	-

Tabelle 7. Kombinationstabelle der Endng des AT 1a

AT 2: Dieser Teil der Verbalendungen hat zwei unterschiedliche Formen bzw. Allomorphe, nämlich die Grundform und eine mit einem Bindevokal 으- davor. Die Wahl der Form hängt von der Offen- bzw. Geschlossenheit der letzten Stammsilbe ab. Die Konjunktivendung -ㄴ이 gehört z.B. zu dieser Gruppe. Tabelle 8 und 9 zeigen jeweils die Allomorphbildungsmuster dieser Endungen und ihre Distribution beim Anschluß an den Verbstamm.

Grundform	Allomorph	
	1	2
-니	-니	-으니

Tabelle 8. Allomorphbildung der Endngen des AT 2 and AT 2a

Alternationstyp			Grundform	Allo1	Allo2	Allo3
regulär	Muster der Grundform					
	letzte Silbe	offen geschlossen	가-	가/니	-	-
			잡-	잡/으니	-	-
	uSR		쓰-	쓰/니	쓰-	-
	haSR		하-	하/니	ㅎ-	-
	luSI		다르-	다르/니	달르-	-
	haSR		하-	하/니	ㅎ-	-
	ISR		달-	달-	다/니	-
	bSI		돕-	돕-	도우/니	도오-
	hSR		빨강-	빨강-	빨가/니	빨ㄱ-
	sSI		잇-	잇-	이/으니	-
	dSI		견-	견-	걸/으니	-

Tabelle 9. Tabelle Kombinationstabelle der Endungen des AT 2

Aus praktischen Gründen kann man innerhalb dieser Gruppe eine andere Gruppe als Anschlußtyp 2a aussondern. Die Endungen dieses Anschlußtyps verhalten sich genauso wie die des Anschlußtyps 2 bis auf den Umstand, daß sie an die Grundform der l-semi-reguläre Verben treten, wie in der Tabelle 10 gezeigt wird.

Alternationstyp		Grundform	Allo1	Allo2	Allo3	
regulär	Muster der Grundform					
	Silbe	letzte	offen	가-	가/면	-
		geschlossen	잡-	잡/으면	-	-
uSR			쓰-	쓰/면	쓰-	-
haSR			하-	하/면	ㅎ-	-
luSI			다르-	다르/면	달르-	-
haSR			하-	하/면	ㅎ-	-
ISR			달-	달/면	다-	-
bSI			돕-	돕-	도우/면	도오-
hSR			빨강-	빨강-	빨가/면	빨ㄱ-
sSI			잇-	잇-	이/으면	-
dSI			걸-	걸-	걸/으면	-

Tabelle 10. Kombinationstabelle der Endungen des AT 2a

AT3: Die Endungen dieses Anschlußtyps haben mehrere Formen. Charakteristisch bei der Bildung dieser unterschiedlichen Formen ist, daß sie beim Anschluß an den Verbstamm vor allem mit konsonantischem Auslaut einen Bindevokal vor sich brauchen und die Wahl des Bindevokals abhängig von dem Vokalqualität und der Offen- bzw. Geschlossenheit der letzten Silbe des Verbstamms und dessen Allomorphbildungstyp ist. - 아/ㅏ, -어/ㅓ¹⁶⁾, -여 und -히 werden jeweils als Bindevokal verwendet. Die Tabelle 11 und 12 zeigen jeweils die Allomorphbildungsmuster dieser Endungen und deren Distribution.

Grundform	Allomorph						
	1	2	3	4	5	6	7
-서	-서	-아서	-ㅏ서	-어서	-ㅓ서	-여서	-히서

Tabelle 11. Allomorphbildung der Endungen des AT 3

16) Die Vokalalternation zwischen 아/ㅏ und 어/ㅓ verkörpert einen Rest von Vokalharmonie (als einem Merkmal der altaischen Sprachen) im Koreanischen, die noch im Mittelkoreanischen weit stärker ausgeprägt war.

Alternationstyp			Grundform	Allo1	Allo2	Allo3	
Muster der Grundform							
regulär	letzte Silbe: offen	letzter Vokal	ㅏ, ㅑ	가- 서-	가/서 서/서	-	-
			(o) + ㅓ	오-	오/ㅓ	-	-
			(beliebige Zeichen außer (o)+ ㅓ	보-	보/아 /ㅓ	-	-
			ㅣ	기-	기/어서	-	-
				지-	지/ㄱ서	-	-
				피-	피/어서 /ㄱ서	-	-
			ㅞ	개-	개/어서 /서	-	-
	ㅓ	부수-	부수/어서 /서	-	-		
	letzte Silbe: offen	letzter Vokal	ㅏ, ㅑ	잡- 속-	잡/아서 속/아서	-	-
			restliche	업- 집-	업/어서 집/어서	-	-
	ISR	letzter Vokal	ㅏ, ㅑ	달-	달/아서	-	-
			restliche	질-	질/어서	-	-
	uSR	letzter Vokal	ㅏ, ㅑ	담그-	담그-	담ㄱ/ㅓ서	-
			restliche	슬프-	슬프/ㄱ서	슬ㅍ/ㄱ서	-
bSI	letzter Vokal	ㅏ, ㅑ	돕-	돕-	도우-	도우/ㅓ서	
		restliche	덥-	덥-	더우/ㄱ서	-	
sSI	letzter Vokal	ㅏ, ㅑ	жат-	жат-	자/아서	-	
		restliche	잇-	잇-	이/어서	-	
luSI	letzter Vokal	ㅏ, ㅑ	다르-	다르-	달르/ㅓ서	-	
		restliche	저르-	저르-	걸르/ㄱ서	-	
dSI	letzter Vokal	ㅏ, ㅑ	달-	달-	달/아서	-	
		restliche	걸-	걸-	걸/어서	-	
hSR			빨강-	빨강-	빨가-	빨ㄱ/ㅓ서	
haSR			하-	하-	ㅎ/ㅓ서		

Tabelle 12 Kombinationstabelle der Endngen des AT 3

Bis jetzt wurde die Ableitung der Allomorphe der Endungen geschildert, die durch das Anfügen verschiedener Bindevokale geschieht, und deren Distribution. Es gibt jedoch einige Endungen, die eine andere Art von Allomorphie zeigen: die Endungen -다 und -랴 sind ein Beispiel dafür: sie sind beide Allomorphe der Deklarativendung. Die zweite Ausprägung tritt nur an die präfinale Endung -더¹⁷⁾, die erste kommt in den restlichen Umgebungen vor.

- 12) a. 가 - 다
 b. 가 - 겠 - 다
 c. 가 - 더 - 라

Es wäre eventuell möglich, eine von den beiden als Grundform anzunehmen und daraus die andere abzuleiten, was aber vergleichsweise künstlich und auch unpraktisch ist, weil diese Allomorphie nur auf wenige Endungen beschränkt ist. Solche Allomorphe werden also nicht durch die Allo-Regel erzeugt sondern stehen im Grundformlexikon.

Die einfachste Form einer Verbflexion besteht aus zwei obligatorischen Elementen: einem Verbstamm und einer finalen Endung. Im Koreanischen werden jedoch sehr häufig zwischen einem Verbstamm und der finalen Endung sog. präfinale Endungen angefügt, die jeweils semantisch-pragmatische Funktion wie subjektbezogene Honorifikation, Tempus, Modus und adressatbezogene Höflichkeit haben. Die Beschreibung der Konkatenation aller Endungen aneinander benötigt eine Diskussion über Tempus, Modus und Sprechstufen, die über den Rahmen dieses Artikels hinausgeht. Hier sind stattdessen die allgemeinen Aspekte aufgelistet, die bei der Konkatenation beachtet

17) Die Diskussion der Konkatenation der Endungen folgt später.

werden müssen:

- Reihenfolge des Vorkommens: die allgemeine Reihenfolge der einzelnen Endungen ist wie folgt:

(Verbstamm) - Subjektbezogener Honorifikationsmarker - Tempusmarker - Modusmarker - Adressatbezogener Höflichkeitsmarker - Finale Endung

13) 가-시-었-겠-습니-다

- der Allomorphbildungs- und Anschlußtyp der Endung: je nach dem Anschlußtyp der Endung muß abhängig von der Offen- bzw. Geschlossenheit und der Vokalqualität der letzten Silbe der vorangehenden Endung die passende Form (Allomorph) verwendet werden:

14) a. 가 - 시 - 었 - 습니-다

b. 가 - 쓰 - 을 - 까

c. 가 - 쓰 - 더 - ㄴ

- Kombinatorik der Endungen: manche Tempus- und Modusmarker treten nur in Verbindung mit bestimmten finalen Endungen auf, während andere sich gegenseitig ausschließen, was größtenteils semantisch-pragmatisch und/oder auch phonologisch bedingt ist:

15) a. 가 - 더 - 라

b. *가 - 더 - 다

c. 가 - 더 - ㄴ

d. *가 - 쓰 - 은

4. Schlußbemerkungen

Die Flexion im Koreanischen ist ein höchst komplizierter morphologischer Prozeß, dessen Implementierung eine der schwierigsten Aufgaben bei der Erstellung der Morphologiekomponente für die automatische Wortformererkennung ist. Die Vielförmigkeit der Verben ist auf zwei Aspekte zurückzuführen: die Allomorphie und Konkatenation der Allomorphe. Bei der Betrachtung der Flexion geht dieser Artikel von dem LA-MORPH-Ansatz aus, der auf folgenden Komponenten beruht: i) einem Elementarlexikon, ii) einer Menge von Regeln, die aus allen Einträgen des Grundformlexikons die zugehörigen Allomorphe erzeugen (Allo-Regeln) und iii) einer Menge von Regeln, die die Allomorphe zu Wortformen zusammensetzen (Kombi-Regeln).

Zuerst wurde das Allomorphie-Phänomen beschrieben. Von den Wortformen ausgehend, in denen ein Verb sich manifestiert, wurden die Allomorphe der Verbstämme und Endungen ermittelt und die Regeln bzw. Mustern jeweiliger Allomorphbildung vorgestellt. Dann wurden die Konkatenationen (der Allomorphe) der Verbstämme und Endungen aus dem Perspektiv der Allomorphie neu untersucht und Regeln aufgestellt, die die Konkatenation der Allomorphe kontrollieren, so daß sie später bei der Erstellung der Morphologiekomponente für eine automatische Wortformererkennung implementiert werden können, die auf dem LA-MORPH-Ansatz beruht.

Durch die Beachtung der Eigenarten der koreanischen graphischen Silbenstruktur, die in der traditionellen Grammatik unbeachtet bleiben, konnten die Konkatenationsmuster des Verbstamms und der Endung vollständig ermittelt werden. Weiterhin wurde gezeigt, daß in Flexionsparadigmen, die in den traditionellen Grammatiken als irregulär angesehen werden, auch (verschiedene Grade von) Regelmäßigkeiten auftreten.

Literaturverzeichnis

- 고영근(1989) 국어형태론연구.서울. 서울대학교 출판부.
- 김차균(1999) 우리말의 시제 구조와 상 인식. 서울. 태학사.
- 서상규 외(1999) 언어 정보의 탐구. 서울. 연세대학교 언어정보개발연구원.
- 서정수(1997) 국어문법. 서울. 한양대학교 출판부.
- 이희자(1999) 텍스트 분석적 국어 어미의 연구. 서울. 한국문화사.
- Lewin, Bruno(1970) Morphologie des Koreanischen Verbs. Wiesbaden.
Hessische Druckerei GmbH.
- Hausser, Roland(2000) Grundlagen der Computerlinguistik. Berlin.
Springer.
- _____ (1989) "Principles of Computational Morphology". Laboratory
of Computational Linguistics. Carnegie Mellon University.
- Herrmann, Wilfried(1994) Lehrbuch der modernen koreanischen Sprache.
Hamburg. Buske.
- Lee, Hansol H. B.(1989) Korean Grammar. New York. Oxford University
Press.
- Martin, Samuel(1992) Reference Grammar for Korean. Tokyo. Charles E.
Tuttle Company.
- Sohn, Ho-Min(1999) The Korean Language. Cambridge. Cambridge
University Press.

국문요약

한국어의 굴절: 변이와 상호결합

김수라(중앙대)

필자는 이 논문에서 LA-MORPH 개념에 바탕을 둔 한국어 어형 분석 시스템을 구현하기 위한 사전 작업으로서 한국어 동사의 활용 현상을 기술하였다. 한국어 동사 활용의 특징인 동사의 다형성은 크게 두 가지 측면으로 나뉘어서 관찰할 수 있다. 하나는 동사 및 어미의 변이형 형성과 관련된 측면이고 다른 하나는 이들 동사 및 어미 변이형의 상호 결합에 관련된 측면이다.

관찰의 출발점은 동사가 구체적으로 실현되어 나타나는 단어 형태이다. 규칙 활용 및 불규칙 활용 동사의 실현 형태를 어간 및 어미로 분절하고 이 작업을 통해 얻어진 동사 및 어미의 변이형과 그 기본형 사이에 존재하는 규칙성을 토대로 변이형 유도 모형을 유형별로 제시하였다. 어미를 어간과의 결합 유형에 따라 세 가지로 분류하고 각각에 대해 동사 변이형과의 결합 양상을 제시하였다. 한국어 정서법의 특성을 고려함으로써 동사 활용 현상 중 특히 동사 변이형 및 어미 변이형들 간의 결합 양상 및 규칙 등 기존 문법 및 사전이 완전히 기술하지 못하는 부분을 보완할 수 있었다. 더 나아가 ‘ㄷ’ 불규칙 활용, ‘으’ 불규칙 활용 등 기존 문법에서 불규칙적이라고 간주되는 활용 현상에도 일정한 정도의 규칙성이 존재함을 보였다.