

QPATR를 이용한 독일어 파생어의 어휘구조 분석¹⁾

- 통합에 기반한 어휘통사론적 접근 -

최병진(목포대)

0. 이끄는 말

전산언어학은 전산학과 언어학의 인접학문으로, 언어학의 여러 이론을 전산처리에 의해 그 타당성을 검토하고 이를 바탕으로 새로운 이론이 발전해 나갈 수 있는 기반을 제공한다. 따라서 일반 언어학자와의 긴밀한 협조가 필수적이다. 그러나 전산언어학에 대하여 대부분의 언어학자들은 상당히 복잡하고 어렵다는 선입관을 가지고 이 분야에 대한 연구를 꺼리는 경향이 있다. 또한 전산언어학을 공부하기 위해서는 프로그래밍언어를 습득하고 프로그래밍에 능숙해야만 한다고 생각한다. 이러한 생각 때문인지 많은 사람들이 언어학의 응용분야로서 전산언어학에 대한 학문적 이해를 원하면서도, 쉽게 접근을 하지 못하고 심지어는 전산언어학에 대한 거부감조차 느끼는 경우도 있다. 전산언어학의 연구를 위해서 반드시 프로그래밍 지식이 필요한 것은 아니다. 전문적인 프로그래밍 지식을 가지고 프로그램을 개발하기도 하지만, 개발된 프로그램(언어분석도구)을 가지고 언어학적 이론을 검증할 수도 있다.

이러한 이유로 이 논문에서는 전산언어학에 대한 언어학자들의 이해를 용이하게 하고 상호연구가 긴밀하게 연관되어 이루어질 수 있도록 문장분석에 이용할 수 있는 PATR-II 형식의 구문분석기를 소개하고자 한다. 이미 국외에서는 PATR-II 형식을 언어분석에 다양하게 이용하고 있지만, 국내에서는 그렇지 못한 실정이다²⁾. 여기서 소개하고자하는 PATR-II 구문분

1) 이 논문은 전산언어학에 대한 개략적인 이해를 돕고자 2000년 한국독어학회 정기학술대회의 특강내용을 수정 보완한 것임. 논문의 완성을 위해 부족한 점을 지적해 주신 심사위원 두 분께 감사함을 드린다.

2) PATR-II형식을 독일어 언어분석에 이용한 논문으로는 조자경(1995, 2000)이 있다.

석기는 통합에 기반을 두고 있는 전산언어학적인 도구로, 이론적으로 매우 간결하고 표현력이 뛰어난 통합기반의 문법기술도구이며, 문장분석을 하는데 유용하게 사용될 수 있다.

본 소고는 우리 독어학자들이 언어학적 이론이나 가설의 타당성을 검증하기 위하여 어떻게 PATR-II 구문분석기를 이용할 수 있는지, 그 구체적인 방법을 제시하고자 한다. 먼저 1장에서는 통합기반문법의 형식과 주요개념을 간단히 살펴보고, 2장에서는 통합기반문법형식을 어떻게 PATR-II 형식으로 표현할 수 있는지를 설명한다. 3장에서는 문법기술방법의 이해를 위하여 독일어의 조어구조와 생성가능성을 어휘통사론적 입장에서 분석한다³⁾.

1. 통합기반문법의 기본적 이해

최근 언어학연구에서 통사론의 흐름은 크게 두 가지로 볼 수 있다. 그 하나는 촘스키의 변형생성문법이론을 바탕으로 해서 계속적으로 발전해 나가는 문법이론과 다른 하나는 자연언어의 전산처리를 목적으로 80년대 초부터 발전되어 온 통합기반문법 Unifikationsbasierte Grammatiken의 두 가지 흐름으로 나뉘어진다. 통합기반문법은 다시 브레스난/카플란 Bresnan/Kaplan(1982)의 어휘기능문법 Lexical Functional Grammar, 게즈다 Gazdar(1985)의 일반구조문법 Generalized Phrase Structure Grammar, 폴라드/색 Pollard/Sag(1987)의 머리주도구조문법 Head-driven Phrase Structure Grammar 등의 이론 지향적인 문법이론과 한정절문법 Definite Clause Grammar, 케이 Kay(1983)의 함수통합문법 Functional Unification Grammar, 쉬버 Shieber(1986)의 PATR-II 문법형식과 같은 시스템 지향적 문법으로 구분할 수 있다(Shieber, 1986: 8f.). 그런데 이처럼 다양한 문법이론을 통합문법이론이라고 부르는 이유는 이 문법이론의 틀 안에서 가장 중

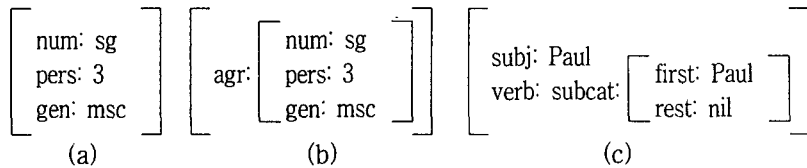
3) 독일어의 조어현상을 전산언어학적인 관점에서 기술한 연구로는 이민행(1999)이 있다. 이민행에서는 wert-합성어를 중심으로 동사부정형과 형용사 'wert'가 결합하여 형성된 wert-합성어의 형태통사적, 의미적인 특성과 합성어 형성 제약들을 고찰하고, wert-합성어의 어휘기재항 생성기를 프롤로그 PROLOG로 구현하였다.

요한 기능을 담당하고 있는 연산이 바로 “통합 Unifikation”이기 때문이다.

지금까지 자연언어의 문장분석을 위해서, 문장구성소들이 결합하기 위한 경로를 문맥자유문법으로 설정하였다. 그리고 문장분석의 정확성을 높이기 위하여, 문맥자유문법과 함께 문장구성소들이 결합하기 위한 제약을 부가하는 방법이 점차로 사용되기 시작하였다. 이때 규칙에 제약을 부가하는 방법으로 도입된 것이 통합이라는 연산이다. 통합을 통하여 문장분석의 정확성은 높아졌으나, 분석과정과 경제성이라는 측면에서는 비효율적이라는 지적도 있다. 그러나 최근에 와서는 이러한 비효율성에 대한 지적은 그리 크게 문제시되고 있지 않다.

이 단락에서는 먼저 다양한 통합문법이론의 이해를 위해 통합이라는 연산이 어떻게 이루어지는지 살펴보도록 한다.

“통합”은 “자질구조 Merkmalstruktur”를 바탕으로 이루어진다. 언어정보를 표현하는 방법에는 여러 가지가 있으나, 요즈음은 언어정보를 표상하는데 주로 자질구조를 이용하고 있다. 자질구조란 “속성”과 “값”으로 이루어진 쌍 *Attribut-Wert-Paare*의 집합으로 속성은 원자적인 기호 *atomare Symbole*이며, 값은 원자적 *atomar*일 수[예: (a)]도 있거나, 또는 복잡한 자질구조일수도 있다. [예: (b), (c)]⁴⁾



<도표 1> 자질 구조의 예

이러한 자질구조를 바탕으로 통합이 이루어지는데, 여기서 통합이란 자질구조의 통합을 말한다. 그리고 통합이라는 연산은 “포함 Subsumption”이라는 관계를 통해서 정의할 수 있다. <도표 2>는 자질구조의 종류와 이들 개별 자질구조사이의 포함관계⁵⁾를 보여주고 있다.

4) 빈 자질구조([])도 존재한다.

[]	(D _{VAR}) ⁶⁾ : 빈 자질구조(변수)
[<i>cat: NP</i>]	(D _{NP}): 명사구
[<i>cat: NP</i> <i>agreement: [number: sg]</i>]	(D _{NPsg}): 단수 명사구
[<i>cat: NP</i> <i>agreement: [number: sg</i> <i>person: third]</i>]	(D _{NP3sg}): 3인칭단수명사구
[<i>cat: NP</i> <i>agreement: [number: sg</i> <i>person: third]</i> <i>subject: [number: sg</i> <i>person: third]</i>]	(D _{NP3sgSubj}): 주어3인칭단수명사구
[<i>cat: NP</i> <i>agreement: [number: sg</i> <i>person: third]</i> <i>subject: 1</i>]	(D' _{NP3sgSubj}): 주어3인칭단수명사구

적은 정보 $D_{VAR} \subseteq D_{NP} \subseteq D_{NPsg} \subseteq D_{NP3sg} \subseteq D_{NP3sgSubj} \subseteq D'_{NP3sgSubj}$ 많은 정보

<도표 2> 자질구조의 종류와 포함관계

<도표 2>에서 알 수 있듯이 왼쪽에 있는 자질 구조일수록 정보의 양은 적은 대신에 더 일반적이고 보편적인 개체의 자질구조를 나타내고 있으며,

- 5) 자질구조의 포함관계 표시는 “ \subseteq ”를 사용한다.
 6) 포함관계의 수식화에서 D(f)란, 자질구조 D에서 속성 f를 말한다. 예를 들어 다음의 D3sg의 자질구조에서 D3sg (number)의 값은 “sg”로서 ‘D3sg(number)= sg’ 로 표시할 수 있다.

[<i>number: sg</i> <i>person: third]</i>	(D3sg): 3인칭단수
--	---------------

오른쪽에 있는 자질 구조일수록 정보의 양은 많아지면서 구체적이고 세분화된 개체의 자질구조를 표시하고 있다. 이러한 자질구조의 포함관계를 바탕으로 통합이 이루어진다.⁷⁾

예를 들어 A와 B라는 두 개의 자질 구조를 통합해야 할 경우, 개개의 자질구조에 정의되어 있는 속성-값에서 모순이 없을 경우에 통합이 이루어진다.(<도표 3>) 통합이 이루어지면서 두 개의 개별 자질 속에 있던 속성-값의 정보는 합쳐지게 되어 통합된 자질 구조는 보다 더 자세한 정보를 가지게 된다. 그러나 통합하려는 자질구조의 개별 속성-값들 사이에 모순이 생기면 통합은 실패한다.⁸⁾(<도표 4>)

$$\begin{array}{ccc}
 \left[\begin{array}{l} \text{cat: NP} \\ \text{agreement: [number: sg]} \end{array} \right] & \sqcup & \left[\begin{array}{l} \text{cat: NP} \\ \text{agreement: [person: third]} \end{array} \right] = \left[\begin{array}{l} \text{cat: NP} \\ \text{agreement: [number: sg} \\ \text{person: third]} \end{array} \right] \\
 \text{A} & \sqcup^9 & \text{B} \\
 \text{D}_{NPsg} & & \text{D}_{NP3} \qquad \qquad \text{D}_{NP3sg} \\
 & & \text{<도표 3>}
 \end{array}$$

$$\begin{array}{ccc}
 \left[\begin{array}{l} \text{cat: NP} \\ \text{agreement: [number: sg]} \end{array} \right] & \sqcup & \left[\begin{array}{l} \text{cat: NP} \\ \text{agreement: [number: plural]} \end{array} \right] \Rightarrow \text{실패} \\
 \text{A} & \sqcup & \text{B} \\
 & & \text{<도표 4>}
 \end{array}$$

- 7) 통합의 여러 과정에 대한 자세한 설명은 Shieber(1986) 17ff. 참조
- 8) 통합은 집합개념의 합집합으로 생각할 수 있으나, 수학에서의 합집합개념과는 차이가 있음에 유의해야 한다. 실제로 수학에서는 R이라는 관계 하에 생겨난 다음과 같은 2개의 집합을 가정할 수 있다.
- $$A = \{(a,1), (b,2)\}, \quad B = \{(a,2), (c,4)\}$$
- 이 때, A와 B의 합집합 $A \cup B = \{(a,1), (a,2), (b,2), (c,4)\}$ 가 된다. 그러나 A와 B를 통합한다고 할 경우, 두 집합의 원소간에 모순이 있는지를 살핀 후, 모순이 없을 경우만 통합된다. 따라서 위의 A, B집합 사이에서는 R이라는 관계 속에, 순서쌍에서 개별 원소 'a'에 해당하는 값은 모순되므로[$(a,1) \in A$ 와 $(a,2) \in B$], 통합은 실패한다.
- 9) 통합의 부호로 '∪'를 사용한다.

2. QPATR를 이용한 문법기술방법

QPATR는 쉬버의 PATR-II 문법형식을 바탕으로 독일 뒤셀도르프 대학에서 개발된 구문분석기로서, 통합기반문법개발을 위한 보편적인 도구이며, 문법을 선언적으로 기술함으로써 문장분석에 이용할 수 있다.

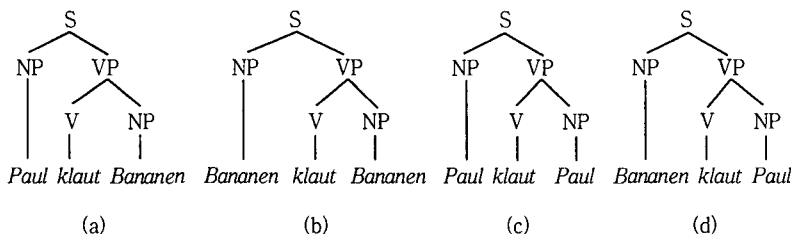
QPATR를 이용하여 문법을 기술하기 위해서는 구구조규칙과 자질구조의 미명세정보를 명시적으로 규정하여야 한다. 구구조규칙은 문맥자유문법의 형식으로 표현하며, 하위구성성분이 상위구성성분으로 결합하는 방법을 나타낸다. 이 규칙에 의하여 문장의 생성이나 분석을 위한 문장구조의 수형도가 만들어진다. 그리고 미명세정보에는 하위구성성분간의 결합을 위한 제약조건을 자질구조의 통합형식으로 표현한다. 즉, 구구조규칙의 오른쪽에 있는 하위구성성분들의 미명세정보에 선언된 통합이 성공하여야만, 왼쪽의 상위구성성분으로의 감축이 일어난다. QPATR는 통합기반의 다양한 통제 원칙들을 쉽게 표현할 수 있는 표현 도구를 제공하고 있기 때문에 다양한 문법이론을 쉽게 표현할 수 있다.

QPATR의 문법을 기술하기 이전에, 우선 구구조규칙과 미명세정보가 어떻게 유용하게 쓰일 수 있는지를 다음의 예를 통해 살펴보고자 한다.

우리는 간단한 문장분석을 위한 구구조규칙을 어휘부와 함께 다음과 같이 나타낼 수 있다.

$$\begin{aligned} S &\rightarrow NP VP \\ VP &\rightarrow V NP \\ NP &\rightarrow Paul \\ NP &\rightarrow Bananen \\ V &\rightarrow klaut \end{aligned}$$

이러한 구구조규칙과 어휘부를 바탕으로 생성되거나 분석될 수 있는 문장은 <도표 5>에서와 같이 4개이다. 물론 여기에는 문장의 뜻이 옳다고 볼 수 없는 비문도 있다.



<도표 5> 문장 수형도

이 중에서 비문법적인 문장의 생성이나 분석을 제한하기 위해서 우리는 미명세정보를 설정할 수 있다. 즉 S의 직접지배를 받는 NP는 주어로 사용되고 '주격(nom)'의 형태를 취하며 VP의 직접지배를 받는 NP는 목적어로서 '목적격(akk)'의 형태를 지닌다는 제약조건을 미명세정보로 정의할 수 있다. 그리고 S의 직접지배를 받는 NP와 VP가 결합할 때, 주어 NP와 동사구 VP사이에는 인칭과 수의 일치관계가 지켜져야 한다는 내용도 미명세정보를 통하여 정의할 수 있다. 이러한 내용을 포함하는 미명세정보는 자질구조의 경로통합등식 Pfadgleichung을 이용하여 나타냄으로써, 구구조규칙은 다음과 같이 확장될 수 있다.

1. $S \rightarrow NP VP$
 $S/head \neq VP/head,$
 $NP/head/agr \neq VP/head/agr,$
 $VP/subcat/first \neq NP,$
 $VP/subcat/rest \neq nil,$
 $NP/head/cas \neq nom.$
2. $VP \rightarrow V NP$
 $VP/head \neq V/head,$
 $V/subcat/first \neq VP/subcat/first,$
 $V/subcat/rest \neq NP,$
 $V/subcat/rest/rest \neq nil.$
3. $NP \rightarrow 'Paul'$
 $NP/head/agr/num \neq sg,$
 $NP/head/agr/pers \neq 3,$
 $NP/head/agr/gen \neq msc,$
 $NP/head/cas \neq nom.$
4. $NP \rightarrow 'Bananen'$
 $NP/head/agr/num \neq pl,$
 $NP/head/agr/pers \neq 3,$
 $NP/head/agr/gen \neq fem,$
 $NP/head/cas \neq acc.$
5. $V \rightarrow 'klaut'$
 $V/head/agr/num \neq sg,$
 $V/head/agr/pers \neq 3,$
 $V/subcat/rest/first \neq Obj,$
 $Obj/head/cas \neq acc.$

규칙 1과 규칙 2는 추상적인 구성성분구조를 형성하고, 이것을 통해서 어휘삽입을 위한 자리들이 마련된다. 규칙 3에서부터 규칙 5를 통하여 어휘삽입이 이루어지는데, 이러한 과정은 통합을 통하여 이루어진다. 규칙 2의 VP는 규칙 1의 VP와 통합하고, 여기에서 $V \leftrightarrow VP \leftrightarrow S$ 사이의 머리자질의 계승이 이루어진다.

미명세정보를 포함하는 위의 구구조규칙을 통하여, <도표 5>에서 제시한 4가지 분석 중에서 하나("Paul klaut Bananen")만을 허용한다. 그것은 우리가 규칙 1에서 명사구 NP와 동사구 VP사이의 자질구조에 대하여 제한을 설정하였기 때문이다. 즉, 문장에서 주어 명사구 NP와 동사구 VP 사이의 일치현상이 머리자질의 통합에 의해 정의되어 있다. 또한 VP의 좌측에 있는 NP는 주어로서 머리자질 'cas:nom(주격)'을 가진다고 정의되어 있어서, 구문분석기는 사전목록에서 이 제한을 충족하는 NP로 "Bananen"이 아닌 "Paul"을 선택하게 된다¹⁰⁾. 계속해서 동사의 하위범주화들에서 "klaut"는 하위범주화목록의 두 번째 요소가 목적어로서 'cas:acc' 라는 머리자질을 가져야 한다는 것을 보여준다. 따라서 여기서 어휘 "Bananen"이 목적어로 선택될 수 있는 것이다.

문법기술에 있어서 주목할 만한 점은 하위범주화목록의 구성이다. 규칙 1에서 VP의 하위범주화 목록이 첫 번째 요소 바로 뒤에서 닫혀 있다(VP/subcat/rest *= nil). 우리는 한 눈에 이 정의가 단 하나의 문장성분, 즉 주어만을 필요로 하는 자동사의 문장만 분석하는데 기여한다는 것을 알 수 있다. 그러나 우리는 어휘부에 2개의 보족어를 필요로 하는 'klaut'라는 타동사를 가지고 있다. 여기서 우리는 통합을 할 때에 모순에 부딪히는 것은 아닌가?

이에 대한 해답은 규칙 2에서의 미명세정보 속에서 찾을 수 있다. 규칙 2의 정의를 자세히 살펴보면, V의 범주에 고유의 하위범주목록을 열어놓고 있으며, 그 첫 번째 요소가 VP의 하위범주목록의 첫 번째 요소와 공지표를 가지고 있다. 여기서는 완전한 목록이 통합되는 것이 아니라 단지 부분적인 정보만 통합이 이루어진다. 이러한 방법으로 우리는 규칙 1을 각각

10) 예제 문법에서는 구문분석과정의 설명을 위해 규칙과 어휘목록을 단순화시켰다. 동사와 관련된 주어의 선택제한 Selektionsrestriktion에 관한 내용은 포함시키지 않았으나 미명세정보를 통하여 규칙을 손쉽게 확장할 수 있다.

다른 하위범주화 목록을 가지는 동사의 문장에도 적용할 수 있다.

이와 같이 개별 어휘에 미명세정보로 정의된 제한은 규칙 1과 규칙 2에 의한 통합을 시도하면서 문법적인 문장을 판별하는데 기여한다. 이 과정에서 정보전달과정은 구문분석기에 의해서 자동으로 이루어지지만, 정보전달 방법은 선언적으로 정의되어야 하며, 이것이 바로 문법개발자가 담당해야 할 일이다. 따라서 QPATR를 이용하여 문장분석을 하기 위해서 우리가 해야 할 일은 바로 규칙 내에서 통합이 어떻게 이루어져야 할지를 명시적으로 규정하는 일이다. 이러한 점에서 이론 중심의 문법학자들이 컴퓨터를 통하여 자신의 이론을 검증하는데 QPATR가 유용하게 사용될 수 있다.

한편, 미명세정보의 기술에 있어서 문장주어에 대한 주격의 제한은 통사적인 차원에서 도입된 반면, 동사의 목적어에 대한 4격 제한은 어휘적인 차원에서 이루어졌다. 이러한 결정은 문법개발자에 의해 이루어진 것이며, 이러한 정의는 QPATR가 이론적으로 어떠한 제한도 받지 않으며, 따라서 이론적으로 가능한 문법의 모델설정을 위해 매우 다양한 가능성을 제시하고 있다는 것을 알 수 있다.

3. 독일어 파생어의 구조 분석

이 논문에서는 독일어 어휘구조의 유형 설정의 정당성과 타당성을 논하기보다는, QPATR를 이용하여 문장을 분석할 수 있는 것과 마찬가지로 어휘구조를 분석할 수 있음을 보여주고자 한다. 이를 위하여 먼저 어휘구조 규칙과 미명세정보를 선언하고, 이를 바탕으로 분석된 어휘구조를 자질구조와 함께 수형도로 표상한다.

물론 이 논문에서 독일어의 어휘구조에 대하여 어휘구조규칙을 설정하고 그 구조를 분석하는 것이 한편으로는 무모하다는 생각도 든다. 이는 독일어의 조어 가능성이 매우 생산적이고, 형태소간의 결합가능성도 매우 다양하여, 어휘구조에 대한 일반화된 규칙을 설정하고 그 기준에 맞추어 어휘를 분석한다는 것이, 문장분석의 차원과 비교할 때 매우 큰 차이가 있기 때문이다.

그럼에도 불구하고 이 논문에서 어휘구조의 분석을 다루고자 하는 것은

문장통사론과 마찬가지로 어휘통사론에서 개별 어휘의 구조나 조어 유형에 대한 구조를 컴퓨터로 파악하기 위해 어떠한 규칙과 제한을 설정해야 할지를 고려하고, 어떻게 그것을 컴퓨터로 분석이 가능하도록 표상할 수 있는지의 예를 보여 주는 것이다. 그리고 이를 통하여 독어학자들로 하여금 전산언어학에 대한 충분한 지식 없이도 언어학적인 현상을 파악하고 분석하는데 QPATR를 도구로 이용할 수 있도록 하는데 초점을 맞추고 있기 때문이다.

이 논문에서 다루는 어휘구조분석의 대상은 파생어의 구조를 보여주어야 하는 것이기 때문에, 명시적 파생어 *explizite Ableitung*를 대상으로 삼으며, 품사의 전환을 통한 파생(*treffen-Treffen*)이나, 통시적 고찰을 전제로 하는 모음교체를 통한 파생어(*Wurf-werfen*)는 제외한다. 또한 'Er-verb'나 'Be-such'와 같은 단어의 파생도 관찰대상에서 제외한다. 'Erwerb'나 'Besuch'는 접두사 'er'나 'be'가 단독으로 명사와 결합하여 파생어를 만들었다고 볼 수 없다. 오히려 'Er-verb'나 'Be-such'는 접두사 조어 'er-verb-en'이나 'be-such-en'을 기초로 무형형태소(\emptyset)에 의해 생겨난 파생명사로 보아야 할 것이다. 이와 같이 형태소의 삭제에 의해 생겨난 파생어 역시 이 논문에서 고려하지 않는다.

한편, 파생의 특별한 형태로 어군 *Wortgruppe*에 접미사가 결합되어 파생이 이루어지는 공성어 *Zusammenbildung*가 있다. 공성어는 파생의 구조를 분석할 때 형태적인 근거에 의하여 분석을 할 경우, 파생의 기초가 되는 것이 형태소가 아니라 어군이 되는 경우가 보통이다. 이 경우 파생어의 의미는 어군의 의미에서 파악되어야 하며, 어군이 관용어일 경우에는 파생어의 의미와 직접구성분요소소의 의미사이에서는 유연성을 찾아보기 어렵다. 이러한 경우 파생의 기초가 되는 것은 형태적인 기준이 아니라 의미적인 기준이 되어야하므로, 이러한 파생어 역시 본 논문의 분석대상에서 제외한다.

3.1. 독일어의 파생 유형

독일어 파생어의 어휘구조를 나타내기 위해서 먼저 파생의 유형을 구분하여야 한다. 독일어의 파생어는 크게 접두사조어 *Präfixbildung*와 접미사

조어 Suffixbildung로 구분할 수 있다. 그런데, 접두사조어 Präfixbildung를 조어의 어떠한 유형으로 분석해야 할 것인가에 대해서는 여러 이견들이 있다. 플라이셔 Fleischer(1975)는 접두사조어를 접미사조어와 동일한 차원에서 파생어에 귀속시키지 않고, 합성어, 파생어 이외의 조어의 또 다른 유형으로 구분하고 있다. 반면에 에르벤 Erben(1975)은 접두사조어와 접미사조어를 파생어의 2가지 유형으로 구분하고 있다¹¹⁾. 또한 분리동사의 경우에는 파생어가 아니라 합성어라고 주장하는 학자(Henzen 1965, Zifonun 1973)도 있다¹²⁾.

이 논문에서는 파생어를 어떻게 분류하는 것이 중요한 것이 아니라 파생어의 구조를 어떻게 보여 줄 수 있는가가 문제이기 때문에 접두사 조어와 접미사 조어를 어떠한 구별을 두지 않고, 파생어로 간주한다. 또한 독일어 파생어에 대한 구조를 어휘통사론적인 입장에서 설명하기 위해서 엑스바 X-bar이론에서 사용되는 개념을 도입한다. 이는 구구조규칙의 체계와 어휘구조규칙의 체계가 형식적으로 유사하며, 어휘구조에서도 구구조에서와 마찬가지로 ‘머리(Kopf)’의 개념을 이용할 수 있기 때문이다. 파생어 어휘 구조에서 머리의 개념은 범주전환의 기능을 지니고 있는 구성성분으로 이해할 수 있다(Olsen, 1986: 41).

이러한 이유로 우리는 독일어의 파생어에 대해 다음과 같은 규칙을 정의할 수 있다.

$N^0 \rightarrow V N^{-1}$: Begehung, Arbeiter
$N^0 \rightarrow A N^{-1}$: Fähigkeit, Dummheit
$A^0 \rightarrow V A^{-1}$: begehbar, arbeitsam

11) 진성복(2000)은 파생어를 접두어 파생어, 접미어 파생어, 유사 접두어 Präfixoide, 유사 접미어 Suffixoide로 구분하여 설명하고 있다.

12) 분리동사의 분리 전철은 대부분이 전치사로서 독립적으로 나타나기도 한다. 즉 분리전철이 자립형태소로 사용되기도 하고 또한 접두사로서 동사와 결합하여 동작양태 Aktionsart를 바꾸어 주기 때문이다. 두덴문법에서는 (1995:449) 분리전철을 반접두사 Halbpräfix로 규정하고 분리동사를 접두사조어에서 합성어로 넘어가는 중간영역에 설정하고 있다.

$A^0 \rightarrow N A^{-1}$: blumig, künstlich, verdächtig
$A^0 \rightarrow ADV A^{-1}$: heutig, gestrig
$V \rightarrow V^{-1} N$: versalz-, entgift-
$V \rightarrow V^{-1} A$: verbesser-, erfrisch-

위의 예에서 알 수 있듯이, 파생의 과정에서 독일어의 접미사는 품사를 전환시키는 기능을 수행하고, 접두사는 기초어의 동작양태를 수정하는 의미적인 역할을 담당하고 있다. 이러한 규칙은 일반화하여 나타내면 다음과 같다.

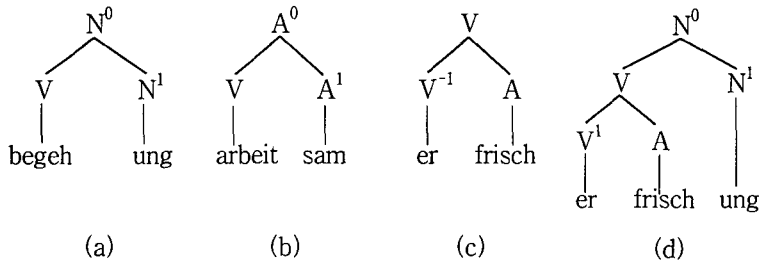
$$X^0 \rightarrow Y X^{-1}$$

$$X \rightarrow X^{-1}Y$$

여기서 X^0 은 X 의 범주에 속하는 자립형태소를 나타내고, X^{-1} 은 어휘통사론적인 머리자질의 기능을 지니고 있으며, X 라는 범주에 속하도록 파생어의 범주를 규정하는 종속형태소를 나타낸다. 그리고 Y 는 임의의 품사범주를 나타내고 있으며, 굴절하지 않는 단어의 어근을 대표한다. 여기서 동사에 대하여는 $V0$ 라는 범주가 없이 V 만이 설정되어 있는 것도 V 가 파생어의 형태가 아니라 동사의 어근으로만 존재하는 것을 나타내기 위함이다.

' $V \rightarrow V^{-1} N$ '과 ' $V \rightarrow V^{-1} A$ '은 접두사 조어에 대한 규칙을 나타내는데, 여기서 V^{-1} 은 명사나 형용사의 앞에 나타나서 'N'이나 'A'와 결합하여 동사의 어간을 형성한다.

이러한 조어규칙을 바탕으로 어휘구조의 수형도를 <도표 6>과 같이 나타낼 수 있다.



<도표 6> 독일어 조어구조의 수형도

3.2. QPATR를 이용한 어휘구조 분석

QPATR를 이용하여 어휘구조를 분석하기 위해서 우리는 먼저 어휘구조 규칙을 일반화하여 QPATR의 문법형식으로 나타내야 한다. 우선 <도표 6>의 (a),(b)와 같은 형태의 어휘구조 분석을 위해 우리는 (1)이라는 규칙이 필요하다. 그리고 (c)와 (d)의 어휘구조분석을 위해서는 (1)의 규칙과 함께 (2)의 규칙이 필요하다.

- (1) WORT --> Wurzel Suffix
- (2) Wurzel --> Präfix Wurzel
- (3) WORT --> Präfix Wort

그밖에 ‘*er-ziehen*’과 같이 동사에 접두사가 붙어 그대로 동사로 머무르는 이중파생어의 어휘구조를 분석하기 위해서는 (3)의 규칙이 도입되어야 한다. 규칙 (3)은 순수접두어 뿐만 아니라, 접두사가 결합하여 파생된 동사에 또 다른 접두사가 결합하는 이중접두어의 파생어(예: *auf-er-steh-en*, *aus-ver-kauf-en*, *ver-aus-gabe-n*, *be-an-spruch-en*)를 분석하는데도 필요하다.

이제 이러한 규칙과 함께 각 어휘구조의 구성성분들간에 머리자질에 대한 제약을 설정할 수 있다.

구성성분이 지니고 있는 머리자질은 위계형태의 어휘구조에서 지배를 하고 있는 상위마디에 계승되어 보다 복잡한 형태의 어휘자질을 구성하는데 기여하기 때문에, 조어의 구성성분의 어느 곳에 머리를 설정하는가가 중요하다.

조어 구조에서 ‘머리’의 위치설정에 대한 논란은 크게 두 가지로 나뉜다. 하나는 오른쪽 머리 규칙을 주장하는 비대칭적인 견해이고 다른 하나는 오른쪽 성분뿐만 아니라 왼쪽 성분도 머리가 될 수 있다는 대칭적인 견해이다. 윌리엄스 Williams(1981)는 조어 구조에서의 머리를 항상 오른쪽에 나타나는 성분으로 보고 이를 오른쪽머리규칙이라고 부른다. 반면에 리버 Lieber(1981)의 경우에는 합성어와 파생어에서 각각 달리 머리의 위치를 설명하고 있다. 즉 합성어에서는 오른쪽 성분이 머리가 되고, 파생어의 경우에는 오른쪽 성분인 접미사도 머리가 될 수 있지만, 왼쪽 성분인 접두사도

머리가 될 수 있다고 설명하고 있다¹³⁾. 이 논문에서는 리버와 같이 접두사도 접미사와 마찬가지로 머리가 될 수 있다는 대칭적인 입장을 취하고 있는데, 그 이유는 이러한 입장을 근거로 각 구성성분의 통합이 가능한 제한 조건을 쉽게 설정할 수 있기 때문이다.

이제 (1)에서 (3)까지의 규칙을 미명세정보와 함께 확장하여 정의하면 다음과 같다.

```

1 # wort(Wort) ---> wurzel(WZ), suffix(SF) ::
    Wort/kopf/kat *= SF/kopf/derivat/kat,
    Wort/kopf/basis/kat *= WZ/kopf/kat,
    Wort/kopf/basis/basis/kat *= WZ/kopf/basis/kat.

2 # wurzel(WZ) ---> praefix(PF), wurzel(WZ1) ::
    WZ/kopf/kat *= PF/kopf/derivat/kat,
    WZ/kopf/basis/kat *= WZ1/kopf/kat,
    (WZ1/kopf/kat *= adj; WZ1/kopf/kat *=n).

3 # wort(Wort) ---> praefix(PF), wort(Wort1) ::
    Wort/kopf/kat *= Wort1/kopf/basis/kat,
    Wort/kopf/basis/kat *= Wort1/kopf/basis/kat,
    Wort/kopf/basis/kat *= PF/kopf/derivat/kat.

4 # wort(Wort) ---> praefix(PF), wurzel(WZ) ::
    Wort/kopf/kat *= WZ/kopf/kat,
    Wort/kopf/basis/kat *= WZ/kopf/kat,
    Wort/kopf/basis/kat *= PF/kopf/derivat/kat.

```

% 형태소 목록

```

salz    lex  wurzel(F) ::
        F/kopf/kat *= n.
zieh    lex  wurzel(F) ::
        F/kopf/kat *= v.
treu    lex  wurzel(F) ::
        F/kopf/kat *= adj.
er       lex  praefix(F) ::
        F/kopf/derivat/kat *= v.
un       lex  praefix(F) ::
        F/kopf/derivat/kat *= adj.
ver      lex  praefix(F) ::
        F/kopf/derivat/kat *= v.
en       lex  suffix(F) ::

```

13) 오예옥(2000) 34쪽에서 재인용

		F/kopf/derivat/kat *= v.
ung	lex	suffix(F) ::
		F/kopf/derivat/kat *= n.
ig	lex	suffix(F) ::
		F/kopf/derivat/kat *= adj.

<도표 7> 구구조규칙과 형태소목록

<도표 7>의 형태소목록에서 자유형태소는 'F/kopf/kat *= X'라는 형식의 머리자질을 포함하고 있는데 이것은 자유형태소가 속하는 문법범주가 X라는 것을 의미한다. 그리고 접사의 경우에는 X라는 문법범주의 단어를 파생하기 위해 사용되는 구속형태소를 'F/kopf/derivat/kat *= X'로 정의하였다.

규칙 1의 미명세정보가 의미하는 것은 파생어(Wort)의 머리자질속에 있는 문법범주는 접미사(SF)의 머리자질에 있는 문법정보와 통합이 되어야 하며(Wort/kopf/kat *= SF/kopf/derivat/kat), 파생의 기초가 되는 기본어 Basiswort의 문법범주는 어근(WZ)의 문법범주와 통합(Wort/kopf/basis/kat *= WZ/kopf/kat)되어야 한다는 것을 의미한다. 또한 'Er-frisch-ung'이나 'ver-salz-en'과 같이 파생이 다단계에 걸쳐 일어나는 경우, 파생어 기본어의 기본어의 문법범주는 어근의 기본어의 문법범주와 통합되어야 함을 (Wort/kopf/basis/basis/kat *= WZ/kopf/basis/kat) 나타내고 있다.

2의 규칙은 1의 규칙과 함께 접두사조어의 분석에 사용되는 규칙이다. 파생어(WZ)의 머리자질의 문법범주는 접두사 머리자질의 문법범주와 통합 가능하여야 하며(WZ/kopf/kat *= PF/kopf/derivat/kat), 파생어 기본어의 문법범주는 기본어(WZ1) 머리자질의 문법범주와 같아야 한다(WZ/kopf/basis/kat *= WZ1/kopf/kat). 이때 기본어 머리자질의 문법범주는 형용사(adj)이거나 명사(n)이어야 한다(WZ1/kopf/kat *= adj; WZ1/kopf/kat *=n).

3의 규칙은 'er-zieh-en'이나 'auf-er-steh-en'과 같이 기본어에 접두사가 붙어 파생어가 만들어지는 경우 품사의 전환은 일어나지 않고 단지 동작양태의 변화만이 일어나는 이중접두어의 분석을 위한 규칙이다. 이 때, 파생어(Wort) 머리자질의 문법범주는 기초어(Wort1) 기본어 머리자질의 문법범주의 값과 같으며(Wort/kopf/kat *= Wort1/kopf/basis/kat), 파생어 기본어의 문법범주는 파생어 기본어의 기본어의 문법범주와 같고(Wort/kopf/

basis/kat *= Wort1/kopf/basis/kat), 접두사 머리자질의 품사범주와도 같아야 함을 나타낸다(Wort/kopf/basis/kat *= PF/kopf/derivat/kat).

한편 1부터 3의 규칙을 가지고는 'un-treu'나 'Ab-gas'와 같은 순수접두어를 분석할 수 없다. 이를 위하여 4의 규칙이 정의되어야 한다. 순수접두어에서는 접두사가 품사의 전환에 아무런 영향을 미치지 않으며, 기초어는 그 자체가 파생어가 아닌 순수한 자유형태소이다. 따라서 파생어(Wort)의 품사는 자유형태소(WZ)의 머리자질이 가지고 있는 품사범주와 같아야 하며(Wort/kopf/kat *= WZ/kopf/kat), 파생어 기본어 품사범주도 당연히 자유형태소의 문법범주와 같다(Wort/kopf/basis/kat *= WZ/kopf/kat). 그리고 여기서 주목할 만한 내용은 순수접두어의 자유형태소 머리자질의 품사범주와 접두사 머리자질의 품사범주가 통합 가능한지를 검사하는 일(Wort/kopf/basis/kat *= PF/kopf/derivat/kat)이다. 이는 순수접두어의 접두사는 품사전환의 기능을 가지고 있지 않기 때문에 머리자질에 문법범주가 정의되어 있지 않다. 따라서 접두사의 머리자질의 문법범주와 기본어가 되는 자유형태소의 머리자질을 통합하면, 자유형태소의 머리자질에 있는 문법범주가 그대로 유지된다. 그러나 명사나 형용사와 결합하여 동사를 파생시키는 'ver-'와 같은 접두사는 머리자질에 문법범주가 정의되어 있어서, 접두사와 기본어의 머리자질을 통합할 경우, 속성-값 사이에 모순이 생기면 통합이 실패한다. 이러한 정의를 통하여 'ver-un-treu'와 같은 단어는 파생될 수 없으며, 'ver-un-treu-en'이나 'Ver-un-treu-ung'은 파생이 가능하여 그 구조를 분석할 수 있도록 한다¹⁴⁾.

다음은 이러한 규칙을 바탕으로 입력되는 단어의 어휘구조분석이 어떠한 과정을 거쳐 이루어지는지 살펴보도록 한다.

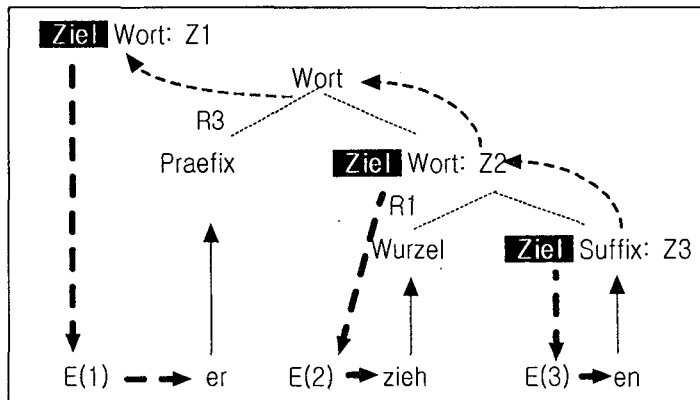
QPATR의 구문분석 Parsing방법은 상향식 Bottom-up구문분석방법과 하향식 Top-down구문분석방법을 결합한 left-corner 구문분석방법을 채택하고 있다. left-corner 구문분석방법을 대략적으로 설명하면 다음과 같다¹⁵⁾.

14) 여기서 통합을 정의하는 방법이나 정의에 사용된 속성과 값은 정해진 것이 아니라, 문법 개발자에 의해 임의로 정해질 수 있다.

15) left-corner 구문분석방식에 대한 자세한 설명은 최병진(2000) 참조

먼저 left-corner 구문분석과정에서는 증명해야 할 목표가 설정이 되면, 입력데이터가 순서대로 하나씩 입력된다. 그리고 읽혀진 입력데이터에 의하여 문법규칙 $N_0 \rightarrow N_1...N_k$ 에서 설정된 목표를 충족시키기 위하여 적용할 규칙이 있는 지를 검토한다. 여기서 주목할 것은 바로 적용규칙을 찾는 방법이다. 적용규칙은 입력된 데이터가 규칙 오른쪽의 왼쪽 구석(N_{1j})에 자리 잡고 있는 목표범주(N_0)에 대한 규칙을 찾아서 적용된다. 그리고 이 N_{1j} 을 지배하는 범주 N_0 가 또다시 규칙 오른쪽 부분의 왼쪽 구석에 위치하는 규칙을 찾아서 적용을 한다. 이와 같이 규칙 오른쪽의 왼쪽 구석에 나타나는 성분이 규칙을 적용해 나가는 열쇠역할을 하기 때문에 Left-corner라는 이름이 붙여졌다. 그런데, N_0 에 의해 지배를 받는 N_{1j} 범주가 다른 구성성분을 자매마디로 가지고 있는 경우, 이 자매 마디는 또다시 증명되어야 할 하나의 목표로 설정이 된다. 그러면 또 다른 목표가 설정이 되었으므로, 목표를 증명하기 위하여 데이터를 입력받는다.

한편 어휘부에 대한 규칙은 일반적으로 'A ---> B'의 형태로 이루어져 있다. 즉 어휘부에 등록된 최소의 입력단위 B는 자매마디를 가지고 있지 않으며, 품사범주 A에 의해서만 지배를 받는다. QPATR에서는 이러한 어휘부에 대한 규칙을 마크로 정의를 이용하여 'B lex A ::'라는 형태로 정의하고 있다.



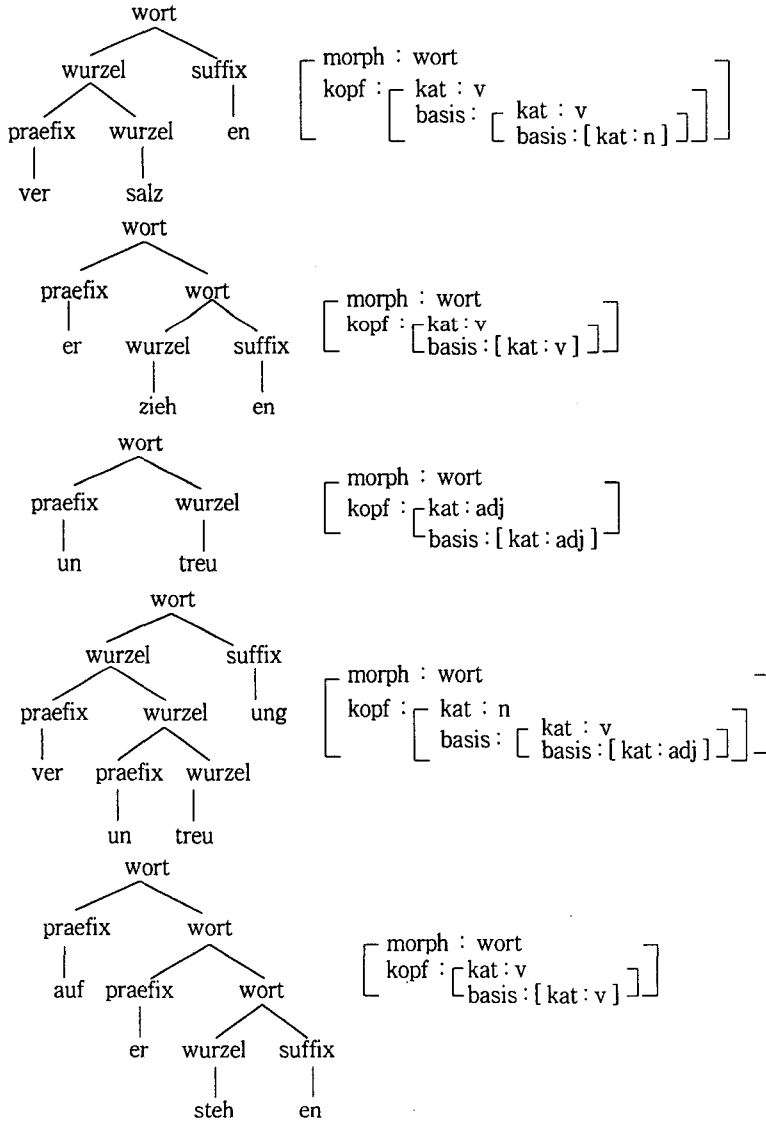
<도표 8> 'er-zieh-en'의 left-corner 분석과정

위의 그림은 지금까지의 설명을 바탕으로, 실제로 'er-zieh-en'이라는 어휘구조의 분석이 어떻게 이루어지는지를 보여주고 있다. 먼저 'er-zieh-en'이 독일어의 올바른 조어구조를 지닌 단어인지를 검사하기 위하여 Wort라는 가정 하에 목표(Z1)로 설정된다. Wort라는 목표가 설정이 되면 입력데이터의 첫 번째 요소(여기서는 형태소)가 읽혀진다(E(1)). 읽혀진 형태소는 'er'이며 'er'는 형태소 목록에서 Praefix라는 것이 확인된다. 형태소 'er'는 Praefix라는 범주로 대체되고, 목표 Wort를 충족시키기 위해 Praefix가 문법규칙 오른 쪽의 왼쪽 구석에 나타나는 규칙을 찾는다. 여기서 규칙 3 '3 # wort ---> praefix , wort'(R3)이 적용되면서 자매마디 Wort가 새로운 목표(Z2)로 설정된다. 목표설정이 되었으므로 다시 새로운 형태소를 입력받는다. 이번에는 'zieh'라는 형태소가 입력되고(E(2)) 형태소 목록에 의해 'Wurzel'로 대체된다. 그리고 목표 Wort를 충족시키기 위해 Wurzel이 규칙 오른쪽의 왼쪽 구석에 나오는 규칙 '1 # wort ---> wurzel , suffix'(R1)을 찾는다. 그리고 Wurzel은 또 'Suffix'라는 자매마디를 가지므로, Suffix는 새로운 목표(Z3)가 된다. 목표 설정에 의해 형태소 'en'이 입력되고(E(3)) 형태소 목록에서 'Suffix'로 대체된다. 이 때 대체된 'Suffix'가 목표 'Z3'를 충족시키는지는 미명세정보의 통합에 의해 확인된다. 목표 'Z3'가 충족되면, 목표 'Z2'가 규칙 'R1'에 의해서 충족되는지, 그리고 목표 'Z1'이 규칙 'R3'에 의해서 충족되는지가 검토된다. 이 모든 과정에서 목표를 충족시키기 위한 조건들이 규칙의 미명세정보에 선언된 통합에 의해서 이루어진다.

이러한 과정을 거쳐 분석된 독일어 어휘의 분석결과는 <도표 9>와 같다.

■ constituent structure:

■ feature structure:



<도표 9> 어휘구조 분석결과: 수형도 및 자질구조

4. 맺는 말

지금까지 통합기반문법에 대한 기본개념을 살펴보고, 문장분석기로 사용할 수 있는 QPATR를 이용하여 독일어 어휘구조를 분석하였다. 독일어 조어의 생산성은 매우 높아서 구문분석기를 이용하여 어휘구조를 분석한다는 것이 어떻게 보면 무리한 시도라고 생각할 수도 있다. 반면에 우리는 종종 독일어 어휘 구조와 조어 생성에 대한 규칙을 찾아내어 일반화하곤 한다.

이 논문에서는 QPATR라는 언어분석도구의 사용과 전산언어학에 대한 이해를 돕기 위하여, 독일어 파생어 구조에 대한 예를 가지고 규칙과 미명세정보를 정의하는 과정에 초점을 맞추었다. 그리고 조어의 생성이나 어휘구조의 분석을 컴퓨터로 실행하여 봄으로써, 조어 규칙이나 유형에 관한 이론을 확인하였다.

이 논문에서 제안한 방법을 통하여 독일어 파생어의 어휘구조를 완전하게 파악하고 분석할 수 있다는 말할 수 없으나, 대부분의 독일어 파생어의 구조를 파악하고 분석하는 것은 가능하다고 본다.

그러나 이 논문에서는 QPATR를 이용하여 어휘구조를 분석할 때 형태론적인 요소만을 고려하였기 때문에 **Kochung*, **Backung* 또는 **Erwachsung*과 같은 단어도 사실상 독일어에 존재하지 않는 단어이지만, 그 어휘구조를 분석할 수 있다는 문제점을 가지고 있다. 이러한 문제점을 해결하기 위해서는 조어가 이루어질 때 기본어와 결합할 수 있는 접미사의 상관관계에 대한 의미적론 내지는 통사론적 요소들을 고려하여 통합을 위한 제약조건을 설정하여야 할 것이다. 이와 같이 특정 접미사나 접두사가 기본어와 결합할 수 있는 상관관계에 관한 내용, 유사한 의미를 지니면서 서로 경쟁관계에 있는 접미사나 접두사가 일정한 기본어와 결합할 수 있는 가능성에 대한 내용, 또는 접미사나 접두사의 의미와 관련된 전체 파생어의 의미표상에 대한 내용 등은 앞으로의 연구과제로 매우 흥미롭다고 생각된다. 이러한 내용들은 언어자료를 토대로 규칙이나 상관관계를 찾아내고 이를 프로그래밍함으로써 독일어 어휘구조의 형태와 의미 분석의 가능성을 더 확장해 줄 것이다.

참고문헌

- 권영을(1999): 파생어 의미의 유연화와 관용화. 독일언어문학 11집, 1-20.
- 오예옥(2000): 어휘 통사론. 한신문화사.
- 이민행(1999): 독일어 어휘부에 대한 연구 - 전산언어학적 접근 -. 독일문학 69집, 308-331.
- 조건(1999): 현대 도이치어 파생명사의 조어의미성분. 독일문학 71집, 445-473.
- 조자경(1995): 의미선택 제약의 기술. 독일문학 57집, 367-386.
- 조자경(2000): 독일어 형용사 어미변화의 새로운 이해. 독일문학 76 집, 269-289.
- 진성복(2000): 독일어 어휘론. 성균관대학교 출판부
- 최병진(2000): 한국어 문장분석과 어휘정보의 연결에 관한 연구. 언어와 정보 4권 2호, 55-68.
- Erben, J. (1975). Einführung in die deutsche Wortbildungslehre. Berlin
- Fliedl, G. (1999): Natürlichkeitstheoretische Morphosyntax. Tübingen.
- Fleischer, W. (1975): Wortbildung der deutschen Gegenwartssprache. Tübingen.
- Henzen, W. (1965): Deutsche Wortbildung. Tübingen.
- Kilbury, J. / Naerger, P. / Renz, I. (1992). New Lexical Entries for Unknown Words. Theorie des Lexikons: Arbeiten des Sonderforschungsbereichs 282, no. 29.
- Koelzer, A. (1992): Eine Semantische Repräsentation für die SIMLEX-Grammatik. Theorie des Lexikons: Arbeiten des Sonderforschungsbereichs 282, no. 29.
- Olsen, S. (1986): Wortbildung im Deutschen. Eine Einführung in die Theorie der Wortstruktur. Stuttgart.
- Shieber, S.M. (1986). An Introduction to Unification-Based Approaches to Grammar. CSLI Lecture Notes Nr. 4. Stanford Univ.
- Zifonun, G. (1973): Zur Theorie der Wortbildung am Beispiel deutscher Präfixverben. München.

Zusammenfassung

Eine Wortsyntaktische Analyse der deutschen Ableitung durch QPATR

Choi, Byung-Jin(Mokpo Nat'l Uni.)

In diesem Aufsatz wird versucht, die Wortbildungsstruktur der deutschen Ableitungen zu analysieren und in QPATR zu implementieren. Dafür werden die deutschen Ableitungen in morphosyntaktischer Hinsicht betrachtet. Die Wortbildungsstruktur der Ableitungen wird durch Phrasenstrukturregeln mit den detaillierten Unterspezifikationen beschrieben, die durch Unifikation definiert sind.

Es wird davon ausgegangen, daß Sätze durch Phrasenstrukturregeln, die die syntaktische Kombinierbarkeit der unmittelbaren Konstituenten definieren, zu analysieren sind. In Wörtern sind auch die Regeln für die Erzeugung und Analyse von abgeleiteten Wörtern zu finden. Die Wortbildungsregeln und die Phrasenstrukturregeln von Sätzen haben Ähnlichkeit von der Form her. Außerdem kann man die morphosyntaktischen Phänomene der Kongruenz oder die Selektionsrestriktion nicht nur auf der Satzsyntax, sondern auch auf der Wortsyntax betrachten. D.h. wenn man durch einen Parser mit Hilfe der Phrasenstrukturregeln einen Satz analysieren kann, kann man auch ein Wort aufgrund der Wortbildungsregeln parsen.

Im ersten Abschnitt werden die grundlegenden Begriffe der "unifikationsbasierten Grammatiken", nämlich die Begriffe 'Merkmalstruktur', 'Unifikation', und 'Subsumption' eingeführt. Im zweiten Abschnitt wird der QPATR-Parser vorgestellt, der aufgrund des 'PATR-II'-Formalismus in Düsseldorf entwickelt worden ist. Im dritten Abschnitt werden die deutschen abgeleiteten Wörter durch QPATR analysiert. Bei der Analyse der Wörter wird gezeigt, wie die

Left-corner-parsing-Strategie funktioniert, und wie die spezifizierten Informationen durch Unifikation als 'Constraints' die zu kombinierenden unmittelbaren Konstituenten beschränken.

Zum Schluß werden die Ergebnisse der Analyse durch Baumstruktur und Merkmalstruktur repräsentiert, die durch QPATR automatisch erzeugt werden.