

인 기준을 형태(소) 결합 관계를 고려한 형태·통사론적 표지로 설정해야 한다고 주장하는 바이며 그 대상은 특정 어휘나 조사, 연결어미 등 표층에서 그 결합 관계를 명시적으로 확인할 수 있는 것이어야 한다.

연구 대상과 관련 연구

본 연구에서는 21세기 세종계획 국어 기초자료 구축 분과에서 구축한 형태분석 말뭉치(이하 세종 말뭉치)와 연구보고서를 대상으로 검토하였다. 또한 자동 형태 분석에서의 품사 중의성 문제들을 검토하기 위한 도구로 고려대 자연어처리 연구실 이도길 연구원의 통계 확률 모델 기반 태거(KOMA)의 결과¹⁾를 활용하였다.

형태 분석을 위한 기준 설정에 대해 논구된 자료는 그리 많지 않으며 그 중 대표적인 것은 임흥빈·송철의 (1998)이다. 그들은 그전까지 개별적으로 진행되어 오던 품사 분류 체계 설정 방법론과 원칙들을 종합하고, 이를 기반으로 한국어 문법과 규범에 맞는 표준안을 세우고자 하는 데서부터 논의를 출발하였다. 나아가 다양하고 실제적인 용례와 함께 정밀한 분석이 가능한 체계적인 대·중·소분류의 품사 표지를 설정하였다.

다만 형태 분석을 위한 품사 표지 체계 연구를 목적으로 하면서도 통사적인 조건은 기본 원칙 등에서만 간접적으로 밝혀져 있고, 의미론적인 세분류가 다수 반영되면서 기준을 정의하는 경계의 모호성, 형태적 일관성 부족 등이 문제로 남게 되었다.

결국 생성이 아닌 분석의 입장에서 요구되는 것은 의미가 아닌 표층구조에 나타난 근거들이라는 점에서 기존의 문법적인 접근에 의한 형태 분석 관련 연구는 실제적으로 적용하는 데 부족함이 있었다. 따라서 형태 분석은 표층구조에서 확인 가능한 근거들을 기반으로 기준을 세우고 분석이 진행되어야 한다. 이는 여러 형태들 간의 결합 관계를 기반으로 각각의 조건이 설명될 수밖에 없음을 의미하며 이렇게 분석되어야만 자연어처리에서 요구되는 명료성과 일관성을 획득할 수 있다.

또한 본 연구에서 분석 대상으로 다루고 있는 세종 말뭉치는 어절을 기본 단위로 한 분석 방식을 택하고 있기 때문에 단어 내에서의 형태적 결합 관계뿐만 아니라 통사적 결합 관계도 확인할 수 있다. 이는 김영길의 (2003)과 같은 선행 연구들을 통해서도 알 수 있듯이 굳이 세종 말뭉치만이 아니라도 어절 단위를 기반으로 이루어지는 자연어처리라면 형태뿐만 아니라 통사적

인 결합 관계도 확인할 수 있다.

이와 같이 본고에서 형태소들의 연결 혹은 통합 관계를 형태소 결합 관계 및 결합 규칙으로 보고 있는 것은 기존의 연구들과 별다른 차이점이 없다. 그러나 어절 내의 결합 관계뿐만 아니라 어절 경계를 넘어선 형태 결합 관계가 어떠한 문법적인 타당성을 획득할 수 있는가에 대한 연구는 부족했던 것이 사실이다. 이는 본고에서 다루고자 하는 품사 중의성과 관련된 예뿐만 아니라 전반적인 형태 분석의 기준과 근거를 요구하는 것에도 닿아 있는 것이다.

하지만 현재에도 이러한 표층형을 고려한 형태 분석 방법론은 보완을 요구하고 있으며, 오랜 시간의 노정에도 불구하고 여전히 형태 분석 관련 연구는 정밀한 기준을 세우는 일과 이에 대한 문법적인 타당성을 확보하는 일이 과제로 남아있다.

형태 분석과 형태론적 중의성

형태 분석에서 발생하는 중의성은 한국어의 교착적 특징이 문법 현상에 반영되면서 비롯된 형태소 분리 중의성과, 품사 표지를 부착하면서 초래되는 품사 중의성으로 크게 나누어 볼 수 있다. 이 중 본고에서 해결하고자 하는 것은 품사 중의성이 내포하고 있는 문제점들이다.²⁾

강승식(2002 : 268~273)에서는 품사 중의성 문제를 '형태론적 중의성'의 일부분으로 설명하는 한편, 광범위한 중의성 현상이 품사 체계의 계층적 차이에 의해 발생하기 때문에 별도로 '품사 체계에 의한 중의성'을 구분해서 이들을 3가지 유형으로 정밀화하고 있다. 그는 대분류 체계 간에서 '1차 품사 중의성', 중분류 체계 간에는 '2차 품사 중의성', 소분류 체계 간에는 '3차 품사 중의성'이 발생하는 것으로 보고 각각을 정의하였다.

본고에서는 각 체계별로 품사 중의성을 보이는 대표적인 3가지 유형을 살펴보고, 문법적인 타당성과 자연어처리에서 요구되는 일관성, 명료성 등을 종합적으로 고려한 품사 표지 태깅(tagging) 방안을 제안하고자 한다.

1. 명사-부사 통용어

품사통용어에 대한 정의는 '하나의 동일한 형태가 아무런 형태적 변화 없이 서로 다른 범주로 나뉘어 품사

1) KOMA 태거와 관련된 연구는 현재 이도길(2005, 고려대학교 박사학위논문)으로 진행 중임

2) 형태소 분리 중의성 문제는 황화상·최정혜(2003)에서 자세히 다루고 있다. 이 논문에서는 형태론적 중의성이 형태소를 분절하는 과정과 해석하는 과정에서 발생하는 것으로 파악하고, 각 과정에 따라 중의성 유형을 체계화 하였다.

가 규정된 어사(語辭)’이며, 이는 강승식(2002)에서 제안한 ‘1차 품사 중의성’에 해당하는 경우이다. 본고에서의 논의 대상은 그 중 교차 범주적 특징이 가장 활발하게 나타나는 명사-부사 간의 통용어에 중심을 두고자 한다. 실제로 이들은 술어와의 관련성에서 보면 논항(argument)과 부가어(adjunction)라는 큰 차이를 보임에도 불구하고, 형태적으로 동일하며 문법적 기능과 구조 또한 매우 유사하다.

때문에 기존의 문법적 연구는, 최호철(2000)에서처럼 통합적인 관점에 따라 이들을 명사 범주로 보면서 다만 여러 가지 부사적인 기능을 갖는다고 보는 주장과, 그와 달리 남기십·고영근(1985)에서와 같이 각각을 명사-부사에 해당하는 별개의 단어로 파악하고 그 단어가 속하는 품사 체계의 다의성을 인정하는 주장으로 나뉜다.

Table 1. 명사-부사 통용어 중 시간 부사의 품사 분류

통용어	사전	표준 국어 사전	동아 새 국어 사전	연세 한국어 사전	우리말 큰사전
내일	「명」 「부」	「명」 「부」	「명」 「부」	「명」	「명」
매년 (每年)	「명」 「부」	「명」 「부」	「명」 「부」	「부」	「명」 「부」
매일 (每日)	「명」 「부」	「명」 「부」	「명」 「부」	「부」	「명」 「부」
어제	「명」 「부」	「명」 「부」	「명」 「부」	「명」	「명」
언제	「대」 「부」	「대」 「부」	「대」 「부」	「대」	「부」 「명」
엊그저께	「명」 「부」	「명」 「부」	「명」 「부」	「명」	「명」
오늘	「명」 「부」	「명」 「부」	「명」 「부」	「명」	「명」
일시 (一時)	「명」 「부」	「명」 「부」	「명」 「부」	「명」	「명」
처음	「명」	「명」	「명」	「명」 1 「부」 2	「명」 「부」
한참	「명」	「명」	「명」	「명」 1 「부」 2	「명」
현재 (現在)	「명」 「부」	「명」 「부」	「명」 「부」	「명」	「명」

위의 표에서 보면, ‘오늘’, ‘내일’, ‘일시’, ‘현재’ 등 시간 부사에 대한 정의가 해당 사전이 통합론과 통용론 중 어떤 방법론상에서 있느냐에 따라 크게 두 가지로

나뉘고 있다.³⁾

자연어처리상에서 두 방법론은 서로 상보적인 장단점을 갖게 된다. 예를 들어, 통합론에 따라 ‘오늘’의 품사를 ‘명사’로만 한정할 경우 모호성이 사라지므로 형태 분석 단계에서의 중의성 문제가 사라지게 되고 자동 처리가 단순해진다. 그러나 부사적 기능에 관한 정보가 표시되지 않으므로, 상위 단계의 구문 분석 등을 수행하는 데 있어서 논항과 부가어, 관형어 중 어떤 기능을 하고 있는지 파악하는 데 어려움을 겪게 된다.

통용론의 경우, 명사와 부사가 구별되는 정보가 주어지므로 기능을 파악하기에는 유리하나, 표층형이 동일하거나 유사한 경우가 많아 통계나 규칙에 따른 표지 부착이 쉽지 않다. 그렇더라도 자연어처리에서는 보다 많은 정보를 표시해줄 수 있는 방법론이 요구되므로, 통용론이 통합론에 비해 유리할 것이다. 그러면 통용론에서 문제가 될 수 있는 예를 살펴보자.

- (1) i. 철수는 [오늘 공부]를 포기했다
- ii. 철수는 [오늘 [공부를 포기했다]]

(1, i)의 ‘오늘’은 관형어로 기능하므로 명사로 볼 수 있으나 (1, ii)에서는 동일한 표층구조임에도 불구하고 ‘오늘’이 부사어로 기능한다. 만약 직관에 의해 문장 이상의 문맥을 고려할 경우에는 문장 의미로 이들을 구별해낼 수 있겠지만, 자연어처리 단계에서는 표층구조가 동일하므로 이를 자동으로 처리하는 것은 불가능하다.

현재 세종 말뭉치에서는 이러한 문제에 대해서 다음과 같은 원칙에 의해서 분석할 것을 기준으로 삼고 있다.

- (2) 일반부사는 종종 일반명사와 동일형태를 띠고 있어 구분이 애매한 경우가 있다. 이들은 뒤에 조사가 결합하느냐의 여부와, 문맥에서 후행 명사를 수식하느냐의 여부에 따라 부사와 명사로 분석될 수 있다.

세종 말뭉치는 수작업을 전제로 한 방식이기 때문에 (2)의 방식이 가능할 수 있지만, 자동 처리를 위한 방법론에서 문맥을 고려한 후 품사 표지를 부착하는 것은 거의 불가능하다.

따라서 본고에서는 (2)의 일반적인 해결 방식에 덧붙여 자동 처리를 고려한 ‘절충적 통용론’을 주장하고자 한다. 즉, 명사-부사 통용어가 관형어나 관형격, 관형형 어미 뒤에 오거나 격조사와 결합된 경우에는 명사 표지를

3) 시간 부사의 품사 체계에 대해, 표준·동아 사전은 주로 통용론을, 연세·우리말 사전은 통합론을 따르는 경향이 확인된다.

부여하고, 보조사가 결합하거나⁴⁾ 조사 없이 단독으로 표층형을 이룰 경우에는 ‘통용어’ 표지를 부여하는 것이다.

- (3) i. 철수는 자신의 오늘/NNG+을/JKO 소중히 여긴다 1)
- ii. 철수는 오늘/MAG 포기한 것을 후회했다
- iii. 철수는 오늘/NMG 공부를 포기했다

(3, i)은 ‘오늘’에 선행하는 관형격조사와 후행하는 목적격조사를 근거로 ‘오늘’이 명사(NNG)로 표지 부착된 예이다. (3, ii)는 ‘오늘’이 후행하는 술어를 수식하므로 부사(MAG)로 판명된 것이다.⁵⁾ (3, iii)은 선행의 결합 관계를 통해 판정하는 것이 불가능하므로 통용어 표지(NMG)를 부여한 것이다. 결합 관계를 통해 판정할 수 없는 경우라도 이들이 문장 내에서 독립해서 기능할 때에는, 대부분 부사에 해당할 확률이 높다. 하지만 일부 부라도 문제될 여지가 있으므로 격조사 없이 독립 어절로 나타나는 경우, 일부 후행 명사와의 결합 관계를 고려한 후 무표적인 표지를 부착하게 되는 것이다.⁶⁾

결국 효율성과 일관성을 중시하는 자료로 활용하려면 통용어 표지를 부사(MAG)로 변환하면 되고, 문법적인 면을 고려한 자료로 쓰일 때는 통사적 기능을 고려한 판단이 보다 용이하도록 하는 것이다. 이러한 방식은 자연어처리를 위한 임시적인 절차이지만 실제 문법적으로도 ‘상당어(相當語)’의 개념에도 해당할 수 있다. 이는 자연어처리에서 일괄적인 처리로 인해 발생하는 현상의 왜곡을 줄여주고 문맥으로 판단해야 하는 부분을 표시해 준다는 점에서 대안이 될 수 있을 것이다.

2. 보조용언 구성

보조용언⁷⁾은 용언 대분류 내에서 중의성을 보이는 경우이며 ‘2차 품사 중의성’에 해당한다. 동일한 형태로

4) 보조사는 ‘오늘도 영화를 만났다’, ‘일을 열심히 했다’처럼 부사에 부가되는 것이 자연스럽게 때문에 명사/부사를 구분하는 근거로 삼기에 부족한 면이 있다.
 5) 이와 같은 결합 관계를 조건으로 포함시킨다면 통용어 표지를 부여하는 빈도는 더 줄어들게 된다. 이외에도 통용어 뒤에 고유명사가 오는 경우나 관형사가 통용어에 후행하는 경우 등도 규칙화할 수 있는 결합 관계를 이룬다.
 6) 이는 대용량의 자연어처리에서 자동화와 일관성을 고려한 것이므로 일부 문법 현상의 왜곡이 있을 수 있다. 이에 대한 통계적 유의미성은 후고로 연구할 것을 기약한다.
 7) 세종 말뭉치에서는 보조용언을 보조동사와 보조형용사로 하위분류하지 않고 있다. 임홍빈·송철의(1998)에서는 이들을 구분할 것을 제안하였으나, 보조용언이 가지고 있는 상적 속성 등으로 인해 이들을 문법적으로 구분하는 것이 어렵고, 얻을 수 있는 효과도 적다고 보았기 때문이다. 그리고 보조용언(VX) 표지 또한 명사-부사 통용어와 마찬가지로 보조동사와 보조형용사 범주의 상위에 있는 무표적이고 중립적인 표지로 볼 수 있다.

인해 본용언-보조용언 간 중의성을 내포하는 단어는 ‘하다, 보다, 주다, 두다, 놓다, 말다’ 등 일반적으로 쓰임이 빈번한 것들이며 빈도(token) 또한 어느 한 부류가 많이 쓰인다고 규정하기 어려울 정도로 다양한 양상을 보인다.

한편, 보조용언은 음운, 통사, 의미론적 의존성을 띠는데 이는 대부분 선행 용언이나 그에 후행하는 연결어미와 관련해서 논의된다. 즉, 통사론적인 의존성-문장 내에서 특정한 요소와 서로 공기할 것을 요구한다는 점 -이 주로 보조용언과 관련된 논의의 초점을 이루게 된다. 따라서 형태 분석 단계에서의 보조용언 처리 방법 역시 통사적인 결합 관계를 중심으로 규칙을 확인할 수 있다.

지금까지 보조용언과 연결어미와의 공기 관계는 대체로 다음에서 유형화 한 것과 같이 비교적 일관된 결합 양상을 보여 왔다.⁸⁾

- (4) i. -고 계시다/있다 v. -ㄴ가/는가/나 보다
- ii. -게 만들다/하다 vi. -는가/나/르가 싶다
- iii. -고야 말다 vii. -어야 하다
- iv. -기는/도/만 하다 viii. -지 말다/못하다/않다
- 남기심 · 고영근(1985)⁹⁾

이처럼 보조용언이 어미와 이루는 결합 관계는 다양하게 연구되어 온 반면, 보조용언 구성과 중의적이면서도 대부분 규칙적으로 본용언 연결 구성에 해당하는 경우에 대한 결합 관계 연구는 많지 않았다. 아래의 표는 본용언과 중의성을 보이는 보조용언뿐만 아니라, 보조용언과 중의성을 보이는 본용언도 결합관계를 가질 수 있음을 보여준다.

Table 2. 본용언 + 본용언 구성

본용언 연결 구성	예문
1. -게 보다/두다	좋게 보다
2. -러 가다/오다	탐사하러 가다
3. -고 가다/오다	들고 가다
4. -면 주다	가면 주다
5. -려고 오다	가려고 오다
6. -라고/자고/다고 하다 ¹⁰⁾	가라고 하면
7. -든지 하다	젓는다든지 하는
8. -지 하다	살아보지 하고

8) 이밖에도 선행행 보조 용언 간의 공기 관계를 통해서도 결합 관계를 규칙화할 수 있으며, 이는 이미 여러 규칙 기반 태거에 적용되고 있다.
 9) 반면, 가장 많은 빈도수를 보이고, 결합 관계도 활발한 연결어미 ‘-어/아-’는 문맥이나 직관에 의해서만 본용언+본용언, 합성용언, 본용언+보조용언 구성 여부를 판가름할 수 있다.

위의 표에 나타난 후행 용언들은 모두 본용언-보조용언간의 중의성을 내포하는 것들로, 선후행어의 품사 표지만을 고려해서 분석하는 자동화 도구에서는 중의성으로 인해 혼란을 일으킬 확률이 높다. 하지만 어휘와 품사 정보를 함께 고려해보면 위와 같은 결합 관계가 추출되는데, 이들을 분석 정보로 활용한다면 본용언-보조용언간의 중의성 문제 해결에 도움이 될 수 있을 것이다.

세종 말뭉치에서는 본용언-보조용언, 본용언-본용언 결합 구성의 판단을 형태 분석 작업자의 직관에 따라서 품사 표지를 부착하였다. 이는 의미적으로 구분하기 힘든 부분이므로 작업자별 직관의 차이에 따라 일관성의 문제가 발생할 수 있다. 그러므로 규칙적으로 공기하는 유형을 파악한 후, 그 나머지에 대해서는 품사통용어 처리와 마찬가지로 무표적인 표지를 부여하는 방법을 생각해 볼 수 있다.

3. 고유명사

고유명사(proper noun)¹¹⁾의 중의성은 일반명사와 고유명사 간에 경계를 나눌 수 있는 형태·통사론적인 기준이 없다는 점이다. 이들은 같은 체언 범주에 해당하며 '고유(固有)'라는 명칭에서 드러나듯 의미론적인 기준에 의해서 나누어진 것이기 때문에 형태론적인 기준만을 가지고 이들을 별도의 표지로 분리한다는 것은 쉽지 않은 일이다.

때문에 강승식(2002 : 272)에서는 고유명사-일반명사간의 중의성을 '3차 품사 중의성'으로 정의하고 소분류내에서도 의미론적인 기준을 요구하므로 이는 형태 분석의 수준을 넘어서는 것으로 보았다. 또한 고유명사는 상당수가 미등록어이기 때문에 형태 분석 성공률이 낮을 뿐만 아니라 일정 정도의 사전이나 데이터베이스가 만들어지기 전에는 그 유형을 파악하기가 쉽지 않다. 그럼에도 불구하고 고유명은 어절 또는 그 이상의 단위를 유효적으로 구분하게 해주는 중요한 실마리(keyword) 역할을 할 수 있기 때문에 이들을 형태 분석단계에서 구별해주는 것은 중요한 과제이다.

고유명사는 '특정한 개체'를 구별하기 위해 쓰이는데, 이를 기준으로 고유명사를 일반명사 등으로부터 달

리 구분해내는 것은 쉽지 않은 일이다. 각각의 고유명이 가진 어휘 개념에 따라 '특정성'이 달리 규정될 수 있기 때문이다.

따라서 고유 명사의 형태 분석을 위해서는 의미 범주에 따른 기준을 구체적으로 요구하게 되는데, 기존 연구에서는 대체로 인명, 지명, 구조물명, 기관명, 단체명, 작품명, 상품명, 도서명, 언어명 등으로 종합하였다.

그러나 그 중 기관명이나 단체명, 상품명, 작품명 등은 '특정한 개체'라는 고유명사의 정의를 고려할 때 문제가 있으므로 이들 중 일부만을 고유명사로 분석하고 나머지는 일반명사로 분석하는 것이 보다 합리적인 해결책이 될 것이다. 그럼 각각의 경우를 살펴보도록 하자.

3.1 특정 기관명이나 단체명

대표적인 기관인 정부 기관이나 특정 회사 등은 고유명사로 처리하는데 문제가 없지만 다음의 (5, ii)와 같은 경우는 중의적이다.

- (5) i. 대한축구협회, 전국은행협회,
한국전자통신연구소
ii. 국문과축구협회, 입주자대표자협회,
생활체육연구소

(5, ii)의 '국문과축구협회'는 전국의 수많은 대학을 대상으로 범위가 확대될 수 있으므로 이를 특정한 개체로 보기는 어렵다. 더욱이 이를 형태 분석했을 때에도 고유명에 해당하는 형태를 확인할 수 없으므로 (5, ii)의 예들은 (5, i)과 달리 구성성분에 따라 분리 분석해야 한다.

그러므로 본고에서는 이러한 문제를 해결하기 위해 특정한 고유명이 포함된 기관·단체명만 고유명사가 될 수 있는 것으로 한정할 것을 주장하고자 한다. 즉, 인명, 지명 등이 포함되거나 '한국, 전국, 국제, 대한' 등 특정 기관이나 단체명에서 높은 빈도로 사용되는 어휘들이 포함되는 경우만 고유명사로 분석할 수 있도록 범위를 제한하는 것이다.

3.2 특정 회사의 상품명, 작품명

차준경(2003 : 276)에서는 상품명은 특정 개체를 지칭할 뿐만 아니라 대량 생산 체제에 따른 현실을 고려해야 하므로 일반명사와 중의성을 보인다고 보았다. 또한 상품명으로 계열 관계의 부류를 지시하는 경우가 빈번히 일어나는데 이는 고유명사의 의미가 확장되는 경우로 볼 수 있다.

고유명사의 의미 확장 현상은 상품명, 작품명 등에도

10) 본용언 '하다'의 연결 규칙은 조진현(2001 : 102)에서 형태소 단위 문맥 규칙으로 제안한 바 있다.

11) 고유명사의 의미와 그 기능에 대한 연구는 이경호(1998), 차준경(2003)에 자세히 소개되어 있다. 그 중에서도 차준경(2003)은 고유명사의 의미 기능에 대한 이론적인 접근과 함께 술어-논항 간의 격률 정보와 의미 정보를 이용한 의미적 중의성 해소 방안을 제안하였다.

일어날 수 있다. 여러 부류의 고유명사들 중에서도 상품명은 의미 확장이 가장 빈번하고, 대량으로 생산된다. 이러한 점들을 고려하면 상품명은 다른 고유명사들이 가지는 '특정 개체'와는 다른 해석이 이루어져야 하며 이들은 결국 형태별로 분리해서 분석해야 할 것이다.

작품명 또한 도서 등은 대량 생산된다는 점을 고려한다면 이들 또한 상품명과 동계에서 처리하는 것이 합당할 것이다. 이러한 처리 방법은 실제 직관과 다를 수 있겠지만, 의미를 기준으로 삼는 고유명사 범주의 특성상 의미적 모호성이 나타나는 것은 불가피하며 이를 최소화하기 위한 형태적 조건을 설정하는 것이라는 점에서 의의가 있을 것이다.

결 론

이상에서 살펴본 바와 같이 본고에서는 강승식(2002)에 따라 품사 중의성을 3단계로 나누고, 각각의 경우에 있어서의 분석 기준과 방안을 모색해 보았다. 그러나 실제로 이러한 방법론이 다양한 예외적인 언어 현상을 포괄할 수 있는 대안이 될 수 있도록 하기 위해서는 더욱 다양한 품사 중의성 문제를 실증적으로 연구하는 것이 필요할 것이다.

본고에서는 우선 품사 중의성 문제를 해결하기 위한 품사 체계 및 범주 상의 대안을 제시하였는데, 나아가 형태론적 중의성 유형들을 자동 처리할 수 있도록 자연어 처리에 적합한 경계 설정을 통해 그 결합 관계 등을 규칙화하는 연구가 이어져야 할 것이다.

형태론적 중의성 전반에 대한 국어학적 연구는 황화상·최정혜(2003), 황화상(2004)에서 진행한바 있다. 이 연구에서는 형태론적 중의성을 분리와 해석의 관점으로 구분해서 형태소 분리 중의성 중 일부는 품사 체계에 의해서 초래되는 것으로 보고, 이를 다양한 어절 유형에서 확인하는 동시에 각 유형별 중의성 해소 방안을 모색하였다. 그러나 실제 말뭉치에서 형태론적 중의성을 일으키는 예 중 숫자나 기호, 영자 등이 결합된 어절에서 초래되는 중의성은 다루지 않고 있는데 이 또한 높은 빈도를 보이는 것들이 많이 있으므로 이에 대한 추후적인 연구도 필요할 것이다.

앞으로 형태 분석 기준 설정에 대한 실증적인 연구가 다양하게 이루어지기 위해서는 단어 형성과 형태 분석의 구분, 형태 변화의 통시성과 공시성 등 국어학의 오랜 쟁점으로부터 풀어가야 할 것이다. 이를 통해 형태

분석 경계 및 기준뿐만 아니라 기준 적용의 우선순위까지 획정할 수 있어야 하며, 이는 또한 한국어 전반의 형태 분석을 위한 일관된 기준 설정 연구를 필요로 하는 것이다.

참고 문헌

- [1] 강승식(2002) : 「한국어 형태소 분석과 정보 검색」, 홍릉과학출판사
- [2] 김영길, 양성일, 홍문표, 박상규(2003) : “형태소 어휘 문맥에 기반한 태깅 오류 정정”, 제15회 한글 및 한국어 정보처리 학술대회 논문집, 63-68
- [3] 김의수(2004) : 국어의 격과 의미역 연구, 고려대학교 박사학위논문
- [4] 김홍규·강범모(2000) : 「한국어 형태소 및 어휘 사용 빈도의 분석」, 고려대학교 민족문화연구원
- [5] 김홍규·강범모 외(1998-2003) : 「21세기 세종계획 국어 기초자료 구축」 - 연구보고서, 문화관광부, 국립국어연구원
- [6] 남기심·고영근(1985) : 「표준국어문법론」, 탑출판사
- [7] 이경호(1998) : 국어 고유명의 의미 기능 연구, 고려대학교 석사학위논문
- [8] 이상주, 류원호, 김진동, 임해창(1998) : “품사 태깅을 위한 어휘 규칙의 자동 획득”, 제10회 한글 및 한국어 정보처리 학술대회 논문집, 20-27
- [9] 임흥빈·송철의(1998) : 「한국어 정보처리를 위한 어절 분석 표지의 표준화 연구」, 21세기 세종계획 기초언어자료기반 분과 세부연구과제, 1998년 최종연구보고서
- [10] 조진현(2001) : 형태소 분석 말뭉치 구축을 위한 태깅 방안의 연구와 개발 - 형태소 기분석 사전과 관계형 데이터베이스 관리 시스템을 이용한 반자동 방법으로-, 고려대학교 석사학위논문
- [11] 차준경(2003) : “고유 명사의 의미와 중의성 해소”, 한국어학 20, 265-286.
- [12] 최호철(2000), “국어의 형태론과 어휘론”, 국어학 35, 313-365.
- [13] 황화상·최정혜(2003) : “한국어 어절의 형태론적 중의성 연구”, 한국어학 20, 287-311.
- [14] 황화상(2004) : 「한국어 전산 형태론」, 월인.