

언어 변화와 언어 처리

- '-는게/는데'의 문법화와 자동 태깅 시스템 -

The Language Change and Language Processing

최 운 호*
(Woon-Ho Choi)

요약 본 논문에서는 현대 한국어에서 나타나는 언어 변화 현상에 대한 설명과 그러한 언어 현상이 언어 처리 시스템에 미칠 수 있는 영향을 연구한다. 현대 한국어에서는 [관형형 어미 + 의존명사 + (조사)]와 같은 통사론적 구성이 형태론적 구성으로 변화되는 과정이 나타나고 있으며 몇몇 형태에서는 문자 언어 생활에서도 두드러지게 나타나고 있다. 이러한 예로 통사론적 구성 [관형형 어미 + 의존명사 '데' (+조사)]이 '-는데'로, [관형형 어미 + 의존명사 '것' + 조사] 구성이 '-는게'로 나타나고 있으며, 음성 언어 생활에서는 더욱 두드러지고 있어서 다른 어미와 구별하기 어렵다. 이와 같은 유형의 형태는 다른 접속문 어미나 내포문 어미처럼 복합문 구성에 관여하는 것으로 파악할 수 있는데, 다른 어미와는 달리 이 형태 자체에 문법적인 격기능이 융합되어 있다. 따라서, 이러한 형태에 대한 분석 방법은 언어 처리 시스템의 구성에 영향을 미칠 수 있으며, 자동 태깅 시스템, 통사 분석 시스템 등에는 특히 그러하다. 그러므로, 언어 처리 시스템의 설계에 이러한 언어 변화 현상이 반영될 필요가 있다.

주제어 언어변화, 형태소 분석, 자동 태깅, 문법화

Abstract This paper aims to research the language changes in modern Korean and its effect to the language processing systems. In modern Korean, the syntactic constructions like [Adnominal Ending + Bound Noun (+ Postposition)] are changing into the morphological constructions, and some of these constructions are reflected in the written language. For example, the syntactic construction [Adnominal Ending + '-de(Bound Noun)' (+ Postposition)] co-exists with the mixed form '-neunde', and [Adnominal Ending + 'geot'(Bound Noun) + '-i'(Postposition)] does with '-neunge'. These constructions are used frequently in the spoken language. As like other verbal endings, these forms also participate in the construction of the complex sentence, and these forms have its own case function fused into themselves. So, the analytic approach to these forms can make great effect on the automatic morphological analysis systems, automatic tagging systems, and the syntactic analysis systems. So, in the design phase of a language processing systems, the language change phenomena like these must be taken into consideration.

Keywords language change, morphological analysis, automatic tagging, grammaticalization

1. 들어가는 말

한국어 전산 처리를 위한 첫번째 단계는 형태소 분석이다. 이 분석 결과를 바탕으로 자동 태깅, 통사 분석 등의 연산을 수행할 수 있으며, 이를 응용해서

기계 번역, 정보 검색, 말뭉치 검색 등의 응용 프로그램을 개발할 수 있다.

자동 형태소 분석과 관련한 기존의 연구 및 시스템이 빠른 처리 속도와 정확한 분석 결과를 보여주고 있다. 본 논문에서는 자동 형태소 분석 시스템의 문법 형태소 목록 선정 및 문법 형태소 분석 방법과 관

* (주)언어과학 부설 언어공학연구소

런된 몇 가지 문제를 제기하고자 한다.

언어는 끊임없이 변화하고 있는데 현대 한국어에도 이러한 변화 양상이 많이 나타나며, 이러한 변화된 특성은 텍스트 자료 및 음성 발화에서 관찰할 수 있다. 이러한 변화 양상의 대표적인 예로 ‘-는게’, ‘-는데’ 등의 유사 어미 형식이 있으며, 이러한 유사 어미 형식의 형태-통사적인 특성과 자연언어처리의 관계를 밝히는 것이 본 논문의 목적이다.

2. 언어 변화와 언어 분석 시스템

한 시대의 언어는 끊임없이 변화하고 있는 과정에 놓여 있다. 현대 한국어에서도 단순히 언어 사용상의 실수 또는 오용이라고 보기에는 어려운 많은 언어 현상들이 나타난다. 이러한 현상에 대한 접근 방법은 여러 가지 양상으로 나타난다.

(1) 카메라를 향해 시중 울기만 하는게 아니라 눈을 삼킬 줄 아는 연기를 하고 싶다고 한다. (스포츠조선, 97. 11 7.)

(1)에서 “-는게”는 어미와 유사한 형태로 사용된다. 이와 같이 [관형형 어미 + 의존명사 + 조사]의 복합

| | |
|-------|---|
| 카메라를 | (N "카메라")<N:20> + (j "를")<1> |
| 향해 | (V "향하")<IT:24> + (e "어")<9> |
| 시중 | (Z "시중")<A:10> (N "시중")<N:20> |
| 울기만 | (N "울기")<N:20> + (j "만") (V "울")<I:23> + (e "기") + (j "만") |
| 하는게 | (V "하")<KTV:20> + (e "는게") |
| | |

(그림 1) HAM 4.0a의 분석 예

형은 현대 한국어에서 다양하게 사용되고 있으며, 이러한 유형의 형태에 대한 분석 방법은 시스템을 어떻게 설계하느냐에 따라서 다른 접근 방식을 보일 수 있다. HAM 4.0a¹⁾에서는 (1)에 대해 (그림 1)과 같은 분석 결과를 보여준다.

강승식(1996)에서는 한국어 어절²⁾의 구조를 52개로 분류해 놓았다. 이 분류는 한글 맞춤법의 띄어쓰기 규정을 기반으로, 한 어절은 어휘형태소로 시작되며 이러한 어휘형태소에 결합하는 문법형태소의 종류에 따라서 결합 가능한 유형을 분류해 놓은 것이다. 이 유형 분류에 기초해서 단순화된 어절 유형이 제시되어 있는데 <표 1>과 같다.

이 분류는 ‘어미+조사’ 결합형을 하나의 어미로 설정하고 분석하는 방식을 취했기 때문에 어미에 교착되는 조사에 대해서는 분석을 하지 않는다. 따라서, 조사 또는 어말어미의 결합형을 사전에 수록한다고 가정하였고, 어말어미 뒤에 교착되는 조사에 대해서도 어말어미의 일부로 가정한다.

(1)의 “하는게”와 같은 어절의 분석 방식은 시스템마다 다른 접근 방식을 보일 수 있다. HAM의 경우에는 [어미 + 조사]의 결합형을 복합어미로 간주하기 때문에 “이루었으면서도”와 같은 어절은 “이루+었+으면서도”로 분석이 되며, 이후 통사 분석 단계에서 조사 ‘-도’에 대한 정보가 필요한 경우에는 어미-조사 결합형에 대한 재분석이 필요하다. 이와 유사한 차원에서 HAM은 ‘-는게’ 유형도 복합 어미로 간주하여 처리하였다. ‘-는게’ 유형의 분석에 대한 접근 방식은 시스템마다 다른 입장을 보일 수 있는데, 그 유형을 (2)와 같이 분류해 볼 수 있다.

- (2) ‘-는게’에 대한 형태소 분석 시스템 접근 방법
- ㄱ. 하(용언) + 는게(어미)
 - ㄴ. 하(용언) + 는(어미) + 게(?)
 - ㄷ. 하(용언) + 는(어미) + 거(의존명사) + |/이(조사)
 - ㄹ. 하(용언) + 는(어미) + 것(의존명사) + 이(조사)

1) HAM은 한성대학교 강승식 선생이 개발한 Hangul Analysis Module이며, 본 논문에서 사용한 것은 HAM 4.0a이다.

1) HAM은 한성대학교 강승식 선생이 개발한 Hangul Analysis Module이며, 본 논문에서 사용한 것은 HAM 4.0a이다.
2) 강승식(1996)에서는 이를 표준어 단어 유형 분류라는 말로 표현했다. 여기에서 ‘단어’라는 용어는 형태론적 단어(Morphological Word)와 혼용해서 사용하였기 때문에 맞춤법 상의 띄어쓰기 단위인 ‘어절’이라는 용어로 다시 수정해서 표현한다.

(표 1) 강승식(1996)의 단순화된 단어 유형 분류

| 유형 표지 | 유형 기호 | 유형 예 |
|-------|-----------|---------------------------------|
| N1. | PTN_N | 체인 |
| N2. | PTN_NJ | 체인 + 조사 |
| N3. | PTN_NSM | 체인 + 용언화접미사 + 어미 |
| N4. | PTN_NSMJ | 체인 + 용언화접미사 + 'ㄴ/기' + 조사 |
| N5. | PTN_NSMXM | 체인 + 용언화접미사 + '아/어' + 보조용언 + 어미 |
| N6. | PTN_NJCM | 체인 + '에서/부터/에서부터' + '이' + 어미 |
| V1. | PTN_VM | 용언 + 어미 |
| V2. | PTN_VMJ | 용언 + 'ㄴ/기' + 조사 |
| V3. | PTN_VMCM | 용언 + 'ㄴ/기' + '이' + 어미 |
| V4. | PTN_VMXM | 용언 + '아/어' + 보조용언 + 어미 |
| V5. | PTN_VMXMJ | 용언 + '아/어' + 보조용언 + 'ㄴ/기' + 조사 |
| A1. | PTN_AID | 단일어 |
| A2. | PTN_ADVJ | 부사 + 조사 |
| X1. | PTN_NVM | 체인 + 동사 + 어미 |

(2.ㄱ)의 접근 방법은 '-는게'를 복합형 어미로 인정하는 경우로 통사 분석 단계에서 필요한 경우 재분석을 해야 한다. (2.ㄴ)~(2.ㄷ)은 (표 1)과는 다른 어절 구조를 가질 경우의 분석이다.

(2.ㄱ)과 같은 분석의 문제점은 '-는게'에 대해 다른 어미와의 통사적 차이를 밝혀 주지 않는다는 것이다³⁾. (2.ㄴ)~(2.ㄷ)은 언어 사용자의 언어 의식이 반영되지 않는다는 것이다. 이 접근 방법은 이러한 형태를 사용자의 오류로 보는 것이며 따라서 맞춤법 교정의 대상으로 보는 것이다. 그렇지만 일상적인 발화에서는 '-는게' 유형 형태의 사용이 일반화되어 있는 언어 현실에서 언어 변화 현상이 반영되지 않는 경우, 음성 언어 이해를 위한 형태-동사 분석 시스템에서 이러한 유형의 유사 어미 형태에 대한 해석 방식에 적절한 전략을 수립하지 못할 수도 있다.

(1)과 같은 예에 대해 언어 분석 시스템은 2가지 서로 다른 태도를 취할 수 있다. 하나는 (1)의 예를 현재 진행 중인 언어 변화가 반영된 자료로 파악하고 이에 대한 분석 전략을 수립하는 것이고, 다른 하나는 언어 정책적인 입장에서 교정 대상으로 파악하고

이를 교정하는 시스템을 개발하는 것이다.

(1)에서 보인 것과 같은 유형의 언어 자료를 교정 대상으로 파악하는 경우로 최재혁 외(1998)를 들 수 있는데, [관형형 어미 + 의존 명사 '데']가 어미 유형으로 사용되는 경우를 교정의 대상으로 파악했다. 최운호(1999)에서는 이를 언어 변화 양상이 진행 중인 현상으로 파악하고 분석 시스템에 반영되어야 한다는 입장을 취하고 있다. 최운호(1999)에서는 이러한 현상을 현대 한국어에서 문법화가 진행 중인 것으로 파악하고 이러한 유사 어미 형태의 형태-통사 특성을 제시했다.

(3) '-ㄴ데/-ㄴ 데'의 비교 (최재혁 외, 1998: 44)

- ㄱ. 목적은 지식을 얻는데 있다.
- ㄴ. 그를 설득하는 데에 며칠이 걸렸다.
- ㄷ. 꼼꼼히 생각해 봤는데, 그 일은...
- ㄹ. 오늘은 휴일인데 회사에 왜 나왔니?
- ㅁ. 눈이 펑펑 쏟아지는데 그가 왔다.
- ㅂ. 형편이 어려운데 계속 공부한다.
- ㅅ. 그것을 이해하는데 도움이 되지 않는다.

최재혁 외(1998)에서 정의 및 제시된 것은 현대 한국어에서 의존명사 '데'가 관형형 어미와 결합하여 붙여쓰는 오류가 많이 나타나기 때문에 이에 대해 자동 교정을 해야 한다는 것이고, 일부 예에 대해서 자동

3) 권재일(1985)에서는 어미를 내포문 구성에 관여하는 어미, 접속문 구성에 관여하는 어미로 분류하였으며, 내포문 구성에 관여하는 어미는 다시 명사구 내포문 구성에 관여하는 어미, 동사구 내포문 구성에 관여하는 어미로 분류하였다. '-는게'와 같은 유형은 내포문 구성에 관여하지만 권재일(1985)의 분류에는 포함되지 않는 유형의 어미로 분류된다.

교정 방법이 있다는 것이다. 그러나 이 글에서는 관형형 어미와 의존명사가 결합한 '-는데'는 단순히 띄어쓰기 오류라고 보지 않고 현대 한국어에서 활발히 진행되고 있는 언어 변화 양상의 하나로 파악하며 이러한 구성을 문법화 현상의 일부로 간주하고자 한다.

3. '-는게'의 문법화

현대 한국어에서 구어체와 문어체는 문법 형태소들의 사용 양상에서 차이를 보인다. 청자높임법 중에서 서술법을 표현하는 어미의 경우 일반적으로 문어체에서는 [+격식]을 나타내는 '-다'의 사용이 일반적인 반면, 구어체에서는 [-격식]의 '-어/지'의 사용이 일반적이다. 그리고, 문어체에는 구어체에서 나타나는 언어 형태의 변화가 반영되는 속도가 느리다. 그럼에도 현대 한국어의 문어체에는 의존명사 '것'과 주격 조사 '이'가 융합된 형태인 '게'의 사용이 일반화되어 있다. 따라서, 구어체에서 그 사용이 일반화된 형태소가 문어체로 확산되었다고 할 수 있다.

전통적으로 문법화는 '한 형태가 어휘적인 것에서 문법적인 것으로, 그리고 덜 문법적인 것에서 더 문법적인 것으로 발전하는 변화'를 말한다(Hopper & Traugott 1993). 70년대 이전까지의 문법화 연구가 주로 역사적인 시각에서의 연구들이었다면, 70년대 들어서서 공시적인 언어 현상을 이해하기 위해서도 문법화는 중요한 개념이라는 인식이 싹트게 되었는데, Givón은 "오늘의 형태는 어제의 통사"라는 주장을 하게 된 것이다(이기갑, 1993:3).

따라서, 현대 한국어라는 공시적인 언어 현실에는 과거의 통사론적 구성에서 변화된 형태론적 단위가 존재하거나 혹은 통사론적 구성과 이러한 구성으로부터 변화가 진행 중인 형태론적 구성이 공존하고 있다고 볼 수 있으며, 이러한 진행 중인 변화에 대한 자료의 분석은 역사언어학의 기본 정신인 경험주의⁴⁾를 충실히 지향한다고 할 수 있다.

3.1. '것'의 형태 변화

현재 우리말에서는 '것'의 준말인 '거'의 사용이 일반화되어 있다. '거'라는 형태에 대해서 한글학회 「우리말 큰사전」에서는 다음과 같이 설명하고 있다.

(4) '거' 형태의 「우리말 큰사전」 설명
거 (매이) '것'의 준말 : 세상 일이란 다 그렇고 그런 것야.

이렇게 줄어든 형태 '거'는 (5)에서처럼 그 사용이 일반화되고 있다.

(5) 내가 돈만 보고 출연해 왔다면 지금보다 재산은 훨씬 많을 거예요. (조선일보, 97. 11. 7)
위와 같이 '것'의 준말 '거' 뿐만 아니라, 이러한 '거' 형태에 문법 형태소가 결합된 '게' 형태도 사용되는데, 「우리말 큰사전」에서는 '게'를 (6)과 같이 설명하고 있다.

(6) '게' 형태의 「우리말 큰사전」 설명
게¹³ '것'의 준말. : 그게 뭐 게냐? 참 좋은 게로구나.

(6)에서 보인 형태 '게'는 의존명사 '것'이 지정사 '이'와 결합한 뒤, 음운탈락 및 축약의 과정을 거쳐서 나타난 형태인데, 권영환(1996)에서는 이러한 형태 '게'가 관형형 어미 '-르'와 함께 선어말어미로 되어가는 과도기에 있는 것이라고 주장하고 있다.

현대 한국어에서는 '것'과 지정사 '이'의 결합형에서 변화한 '게' 형태 뿐만 아니라, '것'과 주격조사 '이'가 결합한 형태론적 단어에서 변화한 형태 '게'도 그 사용이 일반화되고 있다.

(7) 비디오 대여점 5천여개가 사라진 게 영화산업 위기상황으로 비친 것뿐이다. (조선일보, 97. 11. 7)

이러한 형태 '게'는 '것'의 준말 '거'에 주격 조사가 결합한 형태라고 설명할 수는 없다. 현대 우리말에서 주격 조사는 '이/가'가 음운론적 환경에 따라 교체형으로 나타나는데, 조사 '이'는 선행하는 형태의 발음절이 폐음절인 경우에만 결합하며 따라서 '거'라는 형태에 조사가 결합해야 한다면 '가'가 결합해야 한다. 그러므로 '게'라는 형태는 의존명사 '것'과 조사 '이'가 결합한 형태론적 단어 '것이' 자체가 탈락 및 축약이라는 과정을 동시에 거치면서 형성된 것이라고 설명할 수밖에 없다.

(8) '게'의 형성
것이 > 게

4) Paul(1920: 1)에서는 다음과 같이 언어학에서 "역사적"이란 곧 "경험적"이라는 것을 의미한다고 하였다. "Am allerwenigsten darf man diesem allgemeinen Teile der Sprachwissenschaft den *historischen* als den *empirischen* gegenüberstellen."

(8)에서처럼 '것이'에서 형성된 '게'는 형태-통사론적으로 '것이'와 똑같은 특성을 지니고 있다. 우선, 현대 한국어에서 격조사 '이/가', '을/를'의 뒤에는 보조사가 결합할 수 없다. 따라서 *'것이-도/는/만'이 불가능한 형태론적 구성인 것처럼 *'게-도/는/만'도 역시 불가능한 형태론적 구성이다. 통사적으로 '것이'에서 격조사 '-이'는 [[...]내포문 것]명사구과 같은 통사 구성에서 명사구에 주격이라는 문법 기능을 할당한다. 반면, '게'라는 형태에는 이러한 격기능이 하나의 형태에 융합되어 있다. 이러한 특성은 '게'가 관형형 어미와 결합하여 어말 어미를 구성하면서 이제까지의 우리말 어말 어미와는 다른 통사적 특성을 보이게 된다.

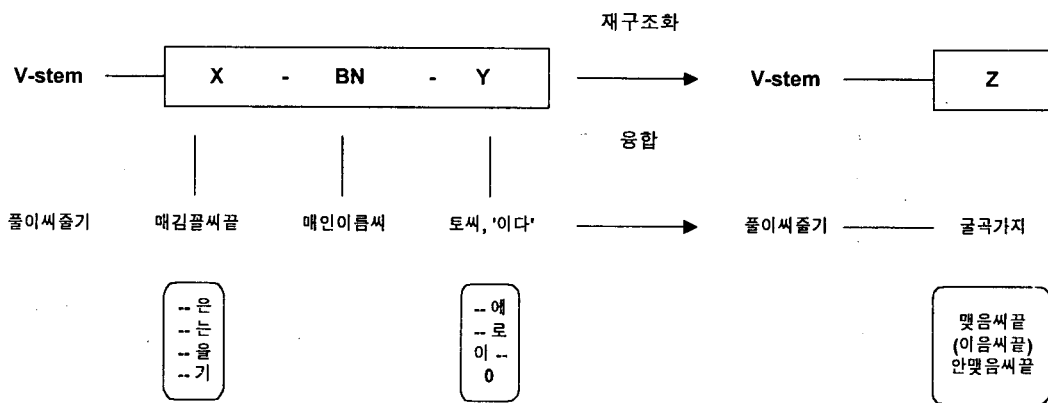
3.2. '-는게'의 형성

요즘 우리의 언어 사용에는 단순히 언어 사용상의 실수, 또는 잘못된 언어 습관이라고만 보기에는 일반적인 현상들이 있다. 이러한 현상 중에는 의존 구문 구성, 특히 의존 명사 구문과 관련된 것이 있는데, 어떠한 형태의 일부가 의존명사인지 어미의 일부인지를 언어 사용자가 결정하기 어려워져 생기는 문제가 그러하다.

권영환(1996: 14)에서는 의존명사의 통사적 구성이 어미라는 형태적 구성으로 되는 과정을 (9)와 같이 형식화하고 있다.

(10) (-은/는/을)지라, (-은/는/을)지, -을지라도, ... -을것이다

그런데, 권영환(1996)에서 제시된 의존명사 구성에서 '-을지라도'와 같은 경우는 이미 어미로 인정되는 구성⁵⁾이다. 그렇지만 '-을것이다'와 같은 의존명사 구성은 형태론적 구성으로, 즉 어미로 기능한다고 보기에는 설명의 무리가 있다. 우리말에서 어미의 형태론적 분포는 용언의 활용에 참여하는 것이다. 용언 어간은 선어말어미, 어말어미와 결합하여 형태론적 단어를 구성하는데, 용언의 형태론적 단어에서 시제 어미는 어말어미의 앞에만 나타난다. 그런데, '-을것이다'와 같은 구성은 '어간-았/었-을 것-이-었/았-다'와 같은 구성을 이룰 수 있다. 이러한 형태소 배열상의 제약으로 인해 '-을것이다'와 같은 구성은 '-을지라도'와 같은 구성과는 달리 분명한 통사적 구성이며, 따라서 이러한 구성을 '-을지라도'와 같은 형태와 동일한 형태론적 구성을 이룰 수 있다고 보기는 어렵다. 그러므로 (9)와 같이 통사론적 구성이 형태론적 구성으로 문법화되어서 어미가 형성되었다고 보기 위해서는 그렇게 생성된 어미의 형태론적 분포가 다른 어미와 동일해야 하며, 그러한 구성이 형태론적 단어의 구성에 참여할 때 다른 어미와 동일한 형태론적 기능을 지녀야 한다.



(9) 매인이름씨 구성의 씨끝되기(권영환, 1996:14)

(9)와 같은 형식화에 기초하여서 권영환(1996)에서는 어미로 기능하는 의존명사 구성에 해당될 수 있는 것들을 제시해 놓았는데 그 중 일부를 인용하면 (10)과 같다.

반면, '것이'가 탈락 및 축약의 과정을 거쳐서 형성된 '게'라는 형태는 관형형 어미와 결합하여 어미로 문법화되어 가고 있다고 볼 수 있다. (11)에서 예로

5) 한글학회 「우리말 큰사전」에 '-르지라도/을지라도'는 어미로 등재되어 있다.

제시된 '-는게'의 사용 빈도는 늘어 가고 있으며, 단순히 맞춤법의 오류라고 보기만은 어렵다.

(11) ㄱ. 내 그림 속에서 우스꽝스런 몸짓으로 꼬물대는 사람의 형상 또한 롤러스케이트를 타고 이 삭막한 세상 속을 미끄러져 가는 채플린의 모습과 닮아 있는**게** 아닌가 싶기도 하다. (조선일보, 97. 11. 7)

ㄴ. 카메라를 향해 시종 울기만 하는**게** 아니라 눈물을 삼킬 줄 아는 연기를 하고 싶다고 한다. (스포츠조선, 97. 11. 7)

ㄷ. 조연 시절엔 주연 연기에 맞추다보니 촬영 전에 열심히 연구**한게** 소용없어지는 경우가 많았다. (스포츠조선, 97. 11. 7)

ㄹ. "그나마 한 두군데라도 불러주는 곳이 있어야 보이콧을 하든 말든 선택의 여지가 있을**게** 아닙니까. ..." (스포츠조선, 97. 11. 7)

ㅁ. 다양한 색상의 하려함보다 두 가지 정도의 색을 적절하게 사용하는**게** 안정감을 준다. (스포츠조선, 97. 11. 7)

ㅂ. 북한이 경제상황 악화를 더 이상 견디지 못하고 갑자기 무너지는**게** 아니냐는 것이 국내외의 공공연한 전망이었다. (중앙일보, 97. 11. 22)

ㅅ. 한국민들에게는 지난 날 극복의 경험이 있는**게** 사실이지만 '대통령 선거전을 들여다 보면 희망을 걸기가 어렵다'**는게** 이 사람의 걱정이다. (중앙일보, 97. 11. 22.)

'-는게' 형태가 의존명사 구문이라는 통사론적 구성에서 어미라는 형태론적 구성으로 문법화되었다고 보는 것은 우리말 의존구문의 통사적 성격에 대한 논의에서도 알 수 있다. 권재일(1987, 1994)에서는 현대 한국어 의존명사 구문의 인식 양상은 통사적 구성으로 인식되는 경우와 형태적 구성으로 인식되는 경우가 있으며, 한국어 문법사에서 수많은 활용어미나 조사들의 발달과정에서 알 수 있듯이 의존명사 구문은 형태적 구성으로 변화되고 있음을 예측할 수 있다고 하였다. 그리고 의존명사 구문이 형태적 구성으로 인식된다고 하더라도 그 본질은 통사적 구성임을 지적하였다(권재일, 1994: 245). 따라서, '-을것이다'와 같은 구성은 비록 형태적 구성으로 인식된다고 하더라도 서로 다른 형태론적 단어가 참여하여 구성되는 통사적 구성이다⁶⁾.

반면, '-는게' 형태는 이미 다른 활용어미와 동일한 형태론적 특성을 지니고 있다. 활용어미가 용언의 활용에 참여하는 것처럼 '-는게' 형태는 용언의 어간과 결합하여 형태론적 단어를 이룬다. 그리고 '-는게'와 어간 사이에는 주체 존재 선어말 어미 '-시-'가 결합할 수 있다. 그러나, 시제 선어말 어미 '-았/었-'의 결합 가능성은 다른 관형형 어미와 동일한 양상을 보인다. 권재일(1994)에서 의존명사 구문이 형태론적 구성으로 인식된다고 하더라도 그 본질을 통사적 구성으로 본 것은 의존명사 구문이 서로 다른 형태론적 단어가 결합하여 이루어진 구성이기 때문이다. (8)에서 제시된 예에서 보인 '-는게'가 단순히 맞춤법의 오류라고만 보기는 어렵다. 우리말 띄어 쓰기는 하나의 형태론적 단어를 하나의 어절로 표현하는 것을 기본 원칙으로 하는데, '-는게'는 언어 사용자가 이것을 통사론적 구성이 아닌 형태론적 구성으로 인식하고 있다는 것을 반영한다고 보는 편이 타당하다. 권재일(1992: 392)에서는 의존명사 구문의 변화 과정을 '통사론적 구성 > 통사론적 구성, 형태론적 구성 > 형태론적 구성'으로 설정하였는데, [관형형 어미 + '게']의 구성은 현대 한국어에서 통사론적 구성과 형태론적 구성이 공존한다고 볼 수 있다. 즉, 언어 사용자가 이것을 통사론적 구성으로 인식하는 경우에는 서로 다른 형태론적 단어의 연쇄인 '-는 게'로 표현하고, 형태론적 구성으로 인식하는 경우에는 활용어미처럼 '-는 게'로 표현한다고 볼 수 있다.

3.3. '-는게'의 통사론적 특성

'-는게'라는 형태에서 '게'는 의존명사 '것'과 주격조사 '이'가 결합한 '것'이라는 형태가 변화한 것이다. 따라서, '-는게' 형태에는 '게' 형태가 지니고 있던 격기능이 그대로 남아 있다. 이러한 특성은 이전의 우리말 어미와는 다른 통사적 특성을 보이게 된다.

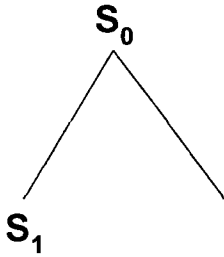
우리말 복합문의 유형은 접속문과 내포문으로 나누어 볼 수 있다. 권재일(1985: 19)에서는 우리말의 복합문 구성 방식을 상위문과 하위문의 관계를 바탕으로 형식화하고 복합문의 유형을 접속문과 내포문으로 분류하였는데 그 유형은 (12)와 같다. 그리고 통사적으로 복합문의 구성에 관여하는 활용어미에는 문법적인 격기능이 없다.

적 구성으로 인식됨 (특히 '것', '지').

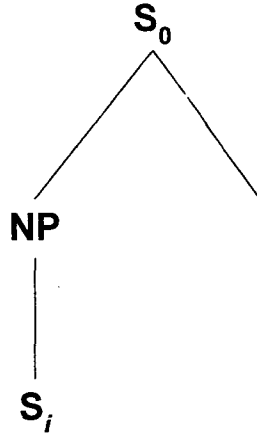
6) 권재일(1994: 244-245)에서는 형태적 구성으로 인식되는 의존명사 구문을 다음과 같이 설명하고 있다.
1. 서로 이웃하고 있는 관형화 어미와 의존명사인 내포문 명사 사이에 긴밀한 통합 관계가 이루어져 그 결과 형태

2. 서로 이웃하고 있는 의존명사인 내포문 명사와 상위문 서술어(하다일 경우에) 사이에 긴밀한 통합 관계가 이루어져 그 결과 형태적 구성으로 인식됨 (특히 '등, 듯, 만, 뻔, 법, 양, 척, 체').

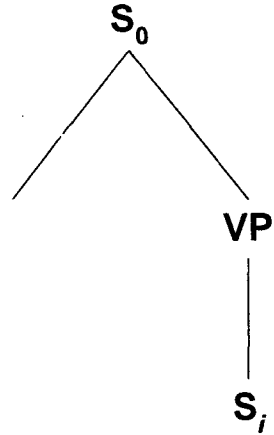
ㄱ. 접속문



ㄴ. 명사구 내포문

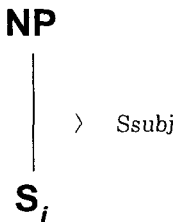


ㄷ. 동사구 내포문



(12) 복합문의 유형(권재일, 1985: 19)

(12)와 같은 유형 분류에 근거하면 우리말의 활용어미는 통사적으로 동사구 내포문의 구성에 관여하는 동사구 내포 어미, 명사구 내포문의 구성에 관여하는 명사구 내포 어미, 그리고 접속 어미로 분류할 수 있다. 내포문을 하위문으로 포함하고 있는 명사구가 상위문 안에서 주격, 목적격 등의 문법적인 격기능을 나타내는 방법은 명사구에 격조사가 결합하는 방식으로 실현되었다. 그런데 '-는게'에는 이미 주격조사가 하나의 형태로 융합되어 있으며, 주격조사가 표현하던 문법적인 격기능도 그대로 그대로 유지되고 있다. '-는게'는 통사적으로는 상위문이 하위문을 직접 관할하는 접속문과 같은 통사 구성을 이루기 때문에 복합



문 구성에서 통사적으로는 접속 어미와 같은 유형이라고 볼 수 있다. 그러나, '-는게'가 결합하여 구성된 하위문은 상위문 안에서 주격이라는 문법 기능을 수행한다. 따라서, '-는게'가 결합한 하위문은 (12.ㄴ)에서 NP와 NP가 관할하는 Si가 주격 기능을 지닌

Ssubj로 변화한 것으로 볼 수 있으며, (13)과 같다.

(13) '-는게'의 통사적 특성 형성 과정
 [(Si) NP] > [Ssubj]

3.4. '-는게'의 어미 설정

'-는게'의 문법화 과정과 통사론적 특성에 대해서 살펴 보았다. 우리말의 문법 변화 과정에서 의존명사구문은 활용어미나 조사로 변화한 경우가 있는데, '-는게'의 문법화 과정에 대한 논의를 통해 이러한 변화가 현대 한국어에서도 진행중임을 알 수 있다.

'-는게'는 [관형형어미 + 의존명사 '것' + 주격조사 '이']라는 통사적 구성이 형태적 구성으로 변화한 것이다. 우선, '것이'라는 형태론적 단어가 탈락과 축약이라는 음운현상을 거치면서 '게'라는 형태가 형성되고, 이렇게 생성된 '게'와 관형형 어미 사이에 긴밀한 통합 관계가 이루어져 형태론적 구성으로 인식되면서, 이제는 형태론적 구성과 통사론적 구성이 공존하고 있는 형태이다.

'-는게' 형태가 나타내는 통사론적 특성은 복합문의 구성에서 상위문에 의해 직접 관할되는 접속문과 동일한 구조를 구성하지만, '-는게'가 결합해서 구성된 하위문이 주격이라는 격기능을 한다는 것이다. 이러한 특성은 '-는게'가 명사구와 명사구에 의해 지배되는 하위문이 동일한 층위로 배열되는 구조 조정을 거치면서 형성되었기 때문에 나타나는 것이다.

3.5. '-는데'의 어미 설정

관형형 어미와 의존명사가 결합한 '-는데'도 '-는게'와 유사한 방법으로 설명이 가능하다.

(14) '-는데'

- ㄱ. 목적은 지식을 얻는데 있다.
- ㄴ. 그를 설득하는데 며칠이 걸렸다.
- ㄷ. 그것을 이해하는데 도움이 되지 않는다.

(14)에서처럼 '-는데'도 '-는게'와 같은 유형의 양상을 보이면 일반적으로 사용된다. (14)는 현대 한국어에서 (15)와 공존한다.

(15) 관형형 어미 + 의존 명사 '데'

- ㄱ. 목적은 지식을 얻는 데 있다.
- ㄴ. 그를 설득하는 데 며칠이 걸렸다.
- ㄷ. 그것을 이해하는 데 도움이 되지 않는다.

(14), (15)는 (16)에서처럼 조사 '에'가 동일한 모음의 반복으로 축약되면서 생성된 것으로 볼 수 있다.

(16) [의존명사구문] + '에'

- ㄱ. 목적은 지식을 얻는 데에 있다.
- ㄴ. 그를 설득하는 데에 며칠이 걸렸다.
- ㄷ. 그것을 이해하는 데에 도움이 되지 않는다.

위의 (14), (15), (16)에서 '-는데' 구문, [관형형 어미] + 의존명사, [의존명사구문] + '에'는 (17)과 같은 명사구 + '에' 구문과 교체 가능하며, 동일한 통사적 환경에서 나타난다.

(17) [명사구] + '에'

- ㄱ. 목적은 경제 발전에 있다.
- ㄴ. 그 작업에 며칠이 걸렸다.
- ㄷ. 이해에 도움이 되지 않는다.

'-는데'도 '-는게'처럼 현대 한국어에서 문법화가 진행 중인 것으로 보이며, '-는데'는 특히 기존의 접속어미 '-는데'와 동일한 형태로 통합되어 가고 있다.

3.3.에서 '-는게'가 문법화되면서 '-는게' 형태에는 주격 조사가 융합되어 있음을 보였다. 따라서, '-는게' 유형에는 다른 문법 형태소가 결합할 여지가 없다. 반면, '-는데'는 의존 명사 '데'에 융합되어 있다고 볼 수 있는 조사 '에'가 의존 명사 '데'와의 음성/음운적 유사성 때문에 그 흔적을 보기 어렵다. 아래 (18)에

서 보듯이 '-는데'에는 조사가 결합할 수 있다⁷⁾.

(18) 형태소 분석이나 구문 분석, 의미 분석 등 각 기술의 난이도에 따라 차이가 있지만, 자연언어 처리 시스템의 성능을 향상시키는데는 한계가 있다.

(18)과 같이 '-는데'의 사용 빈도와 범위가 확장되어 가고 있다는 것은 현대 한국어에서 하나의 형태론적 단위로 인식되어 가고 있음을 보여 준다고 할 수 있다.

4. 결론

현대 한국어에서 나타나는 언어 변화의 양상을 '-는데/는게' 유형의 유사 어미 형태를 중심으로 살펴 보았다.

'-는데/는게' 유형의 형태를 현대 한국어에서 나타나는 언어 변화 현상으로 보았으며, 이를 문법화라는 관점에서 형태-통사적 특성을 설명하였다. 또한 이러한 언어 변화 현상은 한국어 복합문 구성 방법에 새로운 통사 구조를 생성해 내고 있다고 볼 수 있다.

이러한 유형의 언어 변화 현상은 언어 처리 시스템에 영향을 미칠 수밖에 없다.

우선, 형태소 분석 시스템의 경우에는 이러한 유형의 형태를 어미 목록으로 설정해야만 한다. 단순 정보 검색을 위한 형태소 분석 시스템에서는 이러한 유형의 형태가 어미 목록에 추가되는 것이 전체 형태소 분석 시스템의 구조에는 아무런 영향을 미치지 않을 것이다. 어절 구조 자체가 달라지는 것은 아니기 때문이다. 그리고 규범 문법의 입장에서는 최재혁 외(1998)에서 보였듯이 이러한 유형의 형태가 교정의 대상이 될 수 있다. 그렇지만 음성 언어 이해를 위한 음성 언어 형태 분석 시스템의 경우에는 이와 유사한 언어 현상에 대한 연구가 필요하다. 텍스트 자료와는 달리 음성 언어 환경에서는 다양한 유형의 언어 변화 양상이 이미 반영되어 있기 때문이다.

형태-통사 태깅 시스템의 경우에는 기존 접속문 어미나 내포문 어미와는 다른 유형의 통사적 특성을 보이는 이러한 어미에 대해 적절한 태깅 값을 설정할 필요가 있다. 또한 통사 분석 시스템의 경우 복합문 구성에 기존의 통사 구조와는 다른 유형의 통사 구조를 가정할 필요가 있는데, 특히 음성 언어의 경우에는 이러한 유형의 어미 목록과 그 형태-통사 특성에 대한 책략이 필요하리라 본다.

7) (18)은 강승식(1998)의 본문 중 일부를 인용한 것이다.

참고 문헌

- 강승식(1993). "음절 정보와 복수어 단위 정보를 이용한 한국어 형태소 분석", 서울대학교 컴퓨터공학과 박사학위논문.
- 강승식(1996). "한국어 형태소 분석을 위한 단어 유형 분류와 자료구조", 96년도 한글 및 한국어 정보처리 학술대회 논문집, pp. 241-245
- 권영환(1993). '도움풀이씨의 문법화', 부산한글 12, 한글학회 부산지회.
- 권영환(1996). "매인이름씨 구성의 씨끝되기에 대하여", 우리말연구 6, 우리말연구회.
- 권재일(1985). 「국어의 복합문 구성 연구」, 집문당.
- 권재일(1992). 「한국어 통사론」, 민음사.
- 이기갑(1993). "한국어의 문법화", 언어와 문화 8, 목포대학교 어학연구소.
- 최운호(1999). "한국어 전사처리를 위한 형태-통사 태깅 시스템 연구", 서울대학교 언어학과 석사학위 논문.
- 최재혁 외(1998). "한글 문서에서 형태적 중의 오류의 교정", 제10회 한글 및 한국어 정보처리 학술대회 논문집, pp. 41-48.
- Hopper, P. & E. Traugott(1993), *Grammaticalisation*, Cambridge University Press.
- Paul, H.(1920). *Prinzipien der Sprachgeschichte*, Max Niemeyer Verlag.