

한영 병렬 말뭉치를 이용한 한국어 조사 ‘가’와 ‘는’의 선택 원리 연구

유현경·안예리·양수향*

연세대학교

Hyun-kyung Yoo, Yeri An, and Su-hyang Yang. 2007. The Study on the Principles of Selecting Korean Particle ‘Ka’ and ‘Nun’ Using Korean-English Parallel Corpus. *Language and Information* 11.1, 1-23. This study aims to research into the meaning of Korean particle ‘ka’ and ‘nun’ inductively by examining the correspondences of those particles and English articles on the Korean-English parallel corpus. The correspondences were checked in three ways: semantically, syntactically and pragmatically. This study found that when the semantic or syntactic tier is not salient, the pragmatic tier is activated and particles are selected according to the pragmatic elements such as the amount of information or the change of topic. However, if the meaning of the particles is salient or if there is any syntactic motive, particles are selected in accordance with the semantic or syntactic elements. Former studies which focused on one of those three tiers cannot properly explain such correspondences on the Korean-English parallel corpus. This study shows that semantic, syntactic and pragmatic tiers hierarchically affect the selection of a particle and that the selection process is also related to speaker’s intention. This dimensional analysis of particles is expected to contribute to theoretical studies and applied studies like Korean language education as well. (Yonsei University)

Key words: 가(ka), 는(nun), 조사(particle), 관사(article), 병렬 말뭉치(parallel corpus)

1. 서론

최근 한국어 연구에서 말뭉치 자료가 차지하는 비중이 증가하기 시작하면서 말뭉치 자체에 대한 관심도 높아졌다. 균형적으로 잘 만들어진 말뭉치를 연구 자료로 삼을 때 연구

* 120-749 서울시 서대문구 신촌동 134 연세대학교 국어국문학과, E-mail: yoo@yonsei.ac.kr, cherryelee@hanmail.net, shyang@lex.yonsei.ac.kr

† 이 연구의 참여자는 ‘2단계 BK21 연세대학교 한국 언어·문학·문화 국제인력양성 사업단’의 지원비를 받았다. 논문의 내용과 자료에 대한 유용한 조언을 해 주신 심사위원분들께 진심으로 감사드린다.

의 결과도 신뢰성을 확보할 수 있기 때문이다. 말뭉치 자료는 언어 연구에서뿐 아니라 자연언어처리, 기계 번역 등 다양한 분야에서 이용이 가능하기 때문에 국내외에서 국가 사업으로 말뭉치 구축사업이 진행되고 있다. 국내에서는 현재 21세기 세종계획에서 다양한 유형의 말뭉치가 구축되고 있다¹.

이 연구는 21세기 세종계획에서 구축된 특수 말뭉치 중 하나인 한영 병렬 말뭉치 자료를 기반으로 하여 한국어 조사 ‘가’와 ‘는’의 선택 원리에 대하여 살펴보는 것을 목적으로 한다. 조사 ‘가’와 ‘는’ 각각에 대해서는 많은 논의가 이루어졌지만, ‘언제 ‘가’가 선택되고 언제 ‘는’이 선택되는가’에 대해서는 아직 명확한 설명이 제시되어 있지 않다. 본 연구에서는 기존의 연구에서 한국어의 ‘가’, ‘는’과 관련이 있다고 언급되어 온 인구의 정관사, 부정관사와 ‘가’, ‘는’의 대조 분석을 통해, ‘가’와 ‘는’에 대해 객관적인 접근을 시도하여 그 선택 원리를 밝히고자 한다. 또한 이와 더불어 한영 병렬 말뭉치의 활용 가능성에 대하여 살펴보는 것이 또 하나의 목적이다. 다시 말해서 이 연구는 한영 병렬 말뭉치를 이용하여 한국어 연구를 어떻게 할 것인가와 이 연구의 과정에서 드러난 한영 병렬 말뭉치에 대한 여러 문제점들을 지적하여 이에 대한 해결 방안이 앞으로의 한영 병렬 말뭉치 구축에 반영될 수 있게 하는 것의 두 가지 목표를 가지고 있다.

이 글은 다음과 같이 진행된다. 먼저 조사 ‘가’와 ‘는’에 대한 선행 연구를 살피고 기존의 연구에서 조사 ‘가’와 ‘는’에 대한 의미를 추출한다. 이러한 조사의 의미² 중 어떠한 것이 한영 병렬 말뭉치 분석에 유용한가를 실제 말뭉치 분석을 통하여 검증하고 ‘가’와 ‘는’의 선택 원리를 밝힌다. 이때 한국어를 기본 텍스트로 하여 구축된 영어 대역 말뭉치에서 조사 ‘가’와 ‘는’이 어떠한 양상으로 번역되는가에 중점을 두고 살펴보게 될 것이다. 구체적으로 한국어의 조사와 영어의 관사 대응 양상을 의미, 화용, 통사의 세 층위로 나누어 분석해 볼 것이다. 마지막으로 논의 과정에서 제기된 한영 병렬 말뭉치의 여러 문제점과 전망에 대하여 정리하려 한다.

2. 조사 ‘가’와 ‘는’의 의미와 기능

이 장에서는 선행연구에서 조사 ‘가’와 ‘는’의 의미와 기능을 어떻게 설정했는가를 정리

¹ 21세기 세종계획에서는 기초 자료로서 현대국어 말뭉치가 만들어지고 있으며, 특수 자료로서는 구어 전사 자료 말뭉치, 북한 및 해외 한국어 말뭉치, 역사 자료 말뭉치, 한영 병렬 말뭉치 등이 구축되고 있다. 이 중 기초 자료인 현대국어 말뭉치와 역사 자료 말뭉치는 한국어 연구에 활발히 응용되고 있으며 최근 구어 말뭉치에 대한 관심도 높아지고 있는 실정이다. 그러나 북한 및 해외 한국어 말뭉치와 병렬 말뭉치는 실제로 연구에 활용된 예가 전무하다. 말뭉치 구축 단계에서 그 말뭉치를 실제로 연구에 활용하였을 때 나타나는 문제점들이 충분히 검토되고 반영되어야 한다는 측면에서 특수 말뭉치 자료를 활용한 연구가 보다 활성화되어야 한다.

² 한국어에서 조사 ‘가’는 일반적으로 주격조사이고 조사 ‘는’은 보조사 혹은 주제 표지로 알려져 있다. 본 연구에서도 이러한 전제를 부정하지는 않지만 어떠한 명사구에 조사 ‘가’가 결합되지 않는다고 해서 그 명사구가 문장의 주어가 될 수 없는 것은 아니기 때문에 한국어에서 조사 ‘가’의 출현 여부가 문장의 주어를 결정하지는 않는다고 볼 수 있다. 또한 ‘는’이 선택될 경우 주격을 나타내는 ‘가’는 출현하지 않게 되므로 본 연구의 목적인 ‘가’와 ‘는’의 선택 원리와 주격 조사로서의 기능은 별다른 관련이 없다. 그러므로 본 연구에서는 조사의 기능보다는 의미 쪽에 초점을 맞추어 논의를 진행하게 될 것이다.

하여 한영 병렬 말뭉치에 나타난 조사 ‘가’와 ‘는’의 양상을 살피기 위한 가설 자료로 삼으려 한다. 조사 ‘가’와 ‘는’에 대한 선행 연구에서는 ‘가’는 주격의 표지로 보거나 ‘선택지정’, ‘배타적 대립’ 등의 의미를 갖는 것으로 보고 있으며 ‘는’의 경우는 주제를 나타내거나 ‘대조’의 의미를 갖는 것으로 본다.

2.1 조사 ‘가’의 의미와 기능

먼저 ‘가’에 대한 논의들을 살펴보면 신창순 (1975), 이필영 (1982), 고석주 (2002)에서는 공통적으로 ‘가’가 결합되는 명사구 자리에 올 수 있는 여러 명사구들 가운데 하나를 선택하여 지정하는 ‘선택지정’의 의미를 갖는다고 보았다. 이필영 (1982, 421)에 의하면 ‘(무엇이 푸른가?) 하늘이 푸르다’에서 ‘가’는 다른 명사구들이 아닌 ‘하늘’이 ‘선택지정’ 되었음을 나타낸다고 한다. 신창순 (1975), 이필영 (1982), 고석주 (2002)는 조사 ‘가’를 ‘선택지정’의 의미로 본다는 공통점을 가지고 있지만 다음과 같은 차이를 가지고 있다. 신창순 (1975)에서는 ‘가’의 주격 표지 기능을 인정하고 있으며 신창순 (1975)과 이필영 (1982)에서는 ‘선택지정’과 함께 그보다 약한 ‘지정’의 의미도 있는 것으로 보는 반면, 고석주 (2002)에서는 ‘가’의 격표지 기능을 인정하지 않고 ‘선택지정’의 의미만 가지고 있는 것으로 본다.

한편 최수영 (1984)에서는 ‘가’를 의미와 기능의 측면에서 나누어 살펴보았는데 ‘가’의 주된 기능은 ‘주격’을 나타내는 것이며 경우에 따라 ‘주제’를 표시하기도 하고 언제나 ‘배타적 대립’의 의미를 갖는다고 하였다. 여기서의 ‘배타적 대립’은 선택된 명사구 이외의 다른 명사구들을 배제한다는 의미이므로 앞에서 언급한 ‘선택지정’과 흡사한 것으로 볼 수 있다. 최영환 (1987)에서는 ‘가’가 ‘상대초점’을 나타낸다고 보아 여러 대상 중에서 어느 하나를 선택하고 다른 것들을 배제시킨다고 하였으며 장유진 (1999)에서는 ‘가’가 내포문의 주어를 강조하는 ‘지정서술, 선택지정, 신정보, 초점’의 의미기능을 갖는다고 보았다. 반면 이춘숙 (1999)에서는 기존의 논의에서 언급된 ‘배타적 대립’은 조사 ‘가’의 의미가 아니라 ‘강세’에 의해 명시되는 ‘선택 개념’이라고 보았으며 ‘가’는 ‘임자 자리’ 개념을 명시하는 기능을 갖는 것으로 보았다. ‘가’의 의미와 기능에 대한 앞선 연구를 표로 정리하면 [표 1]과 같다.

2.2 조사 ‘는’의 의미와 기능

신창순 (1975)에서는 ‘는’이 여러 동류(同類) 가운데서 하나를 나타내는 ‘대조’의 의미를 갖는다고 보았으며 채완(1977)에서는 ‘는’은 기본적으로 ‘topic’을 나타내며 ‘topic’을 나타내지 않을 때에는 ‘대조’의 의미를 갖는다고 보았다. ‘topic’은 화자와 청자가 공유하고 있는 한정된 지식으로 일반적으로 문두에 나타나며 대개 ‘NP-는’이 문장의 중간에 위치할 때 ‘대조’의 의미를 갖는다고 하였다. 최수영 (1984)에서는 ‘는’의 기능은 ‘주제’ 표시로 보았고 의미는 ‘대조’ 혹은 ‘대조적 대립’이라고 하였다. 한편 국어에 ‘대조’라는 독립적인 범주가 있다고 주장하는 전영철 (2006)에서는 ‘는’이 화제뿐 아니라

연구논문	조사 '가'의 의미/기능
신창순 (1975)	선택지정, 지정/주격조사
이필영 (1982)	선택지정, 지정
최수영 (1984)	배타적 대립/주격조사, 주제표시
최영환 (1987)	상대초점
장유진 (1999)	지정서술, 선택지정, 신정보, 초점
이춘숙 (1999)	선택 개념/임자 자리
고석주 (2002)	선택지정/주격조사 아님

[표 1] 조사 '가'의 의미와 기능

초점에도 작용할 수 있으며, 화제에 작용할 경우 '대조 화제'의 표지, 초점에 작용할 경우 '대조 초점'의 표지로 쓰인다고 보았다.

대부분의 논의에서 '는'의 의미와 기능을 '대조'와 '주제'로 본 반면 최영환 (1987)이나 이춘숙 (1999)에서는 이와 다른 논지를 전개하고 있다. 최영환 (1987)에서는 '가'가 여러 가지 대상들 중에서 서술어와 관계할 수 있는 것을 지시하고 다른 것은 배제시키는 '상대초점'을 나타내는 반면 '는'은 선택되지 않는 것들이 서술어와 관계할 수 있는지의 여부에 대해서는 무관심한 '절대초점'을 나타낸다고 하였다. 최영환 (1987)에 의하면 '나는 간다'의 의미는 '다른 사람이 가는지는 알 수 없지만 나는 간다'는 의미로, 선택되지 않은 개체들에 대해서는 객관적인 태도를 나타낸다는 것이다. 이춘숙 (1999)에서는 '는'은 그것이 결합되는 명사구가 풀이말에 실제로 이끌리는 임자씨향(즉, 영역)으로서 첫 이끌림임을 규정하는 기능을 갖는다고 보았다. 또한 기존의 논의에서 '는'의 의미로 지적되었던 '주제'는 '는' 자체의 의미가 아니라 '월머리'라는 위치와 그 뒤에 오는 '휴지'로 인한 것이며 '대조'의 의미 역시 '강세'에 의해 주어지는 의미일 뿐 '는' 자체의 의미는 아니라고 하였다. 선행연구에서 논의된 '는'의 의미와 기능을 정리하면 다음의 [표 2]와 같다.

연구논문	조사 '는'의 의미/기능
신창순 (1975)	대조
채완 (1977)	topic, 대조
최수영 (1984)	대조 혹은 대조적 대립/주제 표시
최영환 (1987)	절대 초점
이춘숙 (1999)	영역 [설정] 개념 표시
전영철 (2006)	대조 화제, 대조 초점

[표 2] 조사 '는'의 의미와 기능

2.3 한국어 조사와 영어 관사의 대응

앞의 논의들은 대부분 한국어의 문법 체계 안에서 ‘가’와 ‘는’의 의미와 기능에 대하여 살펴본 연구들인데 비해, 목정수 (2003)은 한국어의 조사를 대조언어학적 관점에서 논의하고 있다. 목정수 (2003)은 ‘가’, ‘를’, ‘도’, ‘는’이 모두 ‘한정사(관사)’라는 하나의 범주를 이룬다고 보았으며 ‘가’는 인구어의 부정관사, ‘는’은 인구어의 정관사에 대응된다고 하였다. 이와 비슷한 논의로 신창순 (1975, 164)에서도 국어의 조사 ‘가’가 영어의 정관사처럼 ‘지정’의 의미를 나타낸다고 하였다.

- (1) a. Je cherchais mon chien dans le parc; en vain, semblait-il ... Tout à coup j'aperçois **un petit bout de nez** qui sortait d'un taillis.
- b. 공원에서 내 개를 찾고 있었는데 허사였지 싶었어. 근데 ... 갑자기 자그마한 개 코끝이 덤불속에서 빠져나오고 있는게 보이데. (목정수, 2003, 163 재인용)

(1)은 불어의 부정관사 ‘un’과 한국어의 조사 ‘가’의 대응 양상을 보여주고 있다. 이들 논의에서 지적된 것처럼 한국어의 조사와 인구어의 관사는 분명 대응되는 쓰임을 보인다. 예컨대 (2a)와 (2b)의 차이는 조사 ‘는’과 ‘가’의 출현 여부로, 이는 영어의 정관사, 부정관사의 출현 여부에 따른 차이와 일치한다.

- (2) a. 다람쥐는 도토리를 먹는다.
- a'. The squirrel eats acorns.
- b. 다람쥐가 도토리를 먹는다.
- b'. A squirrel eats acorns.

한국어의 경우 조사의 교체로 인하여 (2a)와 (2b)의 문장은 의미가 분화되는데 (2a)의 경우는 ‘일반적으로 다람쥐라는 동물은 도토리를 먹는다’는 의미로 해석되는 데 비해 (2b)는 ‘한 개체로서의 다람쥐가 도토리를 먹고 있다’는 사실을 나타낸다. 즉, 눈앞에서 다람쥐가 도토리를 먹고 있는 현장을 기술할 때는 (2a)가 아니라 (2b)로 표현한다는 것이다. 이러한 조사 ‘가’와 ‘는’의 차이는 영어, 불어와 같은 인구어의 정관사와 부정관사의 용법과 흡사하다. 이밖에도 한국어에서 신정보에는 ‘가’, 구정보에는 ‘는’이 결합되듯이 인구어에서도 신정보에 부정관사, 구정보에 정관사가 쓰이는 등 한국어 조사 ‘가’, ‘는’의 쓰임과 인구어의 정관사, 부정관사의 쓰임에는 분명 유사한 측면이 있다. 그러나 아직까지는 ‘가’, ‘는’과 관사의 대응 쌍에 대해 대규모 자료를 토대로 한 면밀한 분석이 이루어지지 않았으므로, ‘가’는 부정관사에, ‘는’은 정관사에 대응된다고 단정 지을 수 없다.³

³ 실제로 말뚝치 자료를 살펴보면 ‘가’와 ‘정관사’가 대응되는 경우도 다수 발견된다. 이에 대해서는 3장에서 자세히 살펴보겠다.

3장에서는 2장에서 살펴본 조사 ‘가’와 ‘는’의 여러 의미와 기능 중에서 어떠한 설명이 가장 타당성이 있는지 검증하기 위하여 한영 병렬 말뭉치에서 조사 ‘가’와 ‘는’이 영어의 관사들과 어떤 대응 양상을 보이는지를 살펴보려 한다.

3. 한영 병렬 말뭉치에 나타난 조사 ‘가’와 ‘는’

3장에서는 한영 병렬 말뭉치를 통해 다음의 가설을 검증해보고자 한다.

(3) 가설: 한국어의 ‘가’는 영어의 부정관사에, ‘는’은 정관사에 대응된다.

이러한 가설의 사실 여부를 실제 쓰인 자료를 통해 가능해 보는 것은 영어의 관사 선택 원리와의 대비를 통해 보다 객관적으로 ‘가’, ‘는’의 선택 원리를 밝혀낼 수 있다는 점에서 검증의 의의가 있다. 가설에 합치되는 경우는 ‘가’, ‘는’과 관사 선택의 공통점을 보여줄 것이며, 합치되지 않는 경우는 인구어와는 다른 한국어만의 독자적인 ‘가’, ‘는’ 선택 원리를 보여주게 될 것이다.

3.1 조사-관사 대응 쌍의 유형

본 연구에서 살펴본 말뭉치 자료는 ‘조세희’의 소설⁴ 「난장이가 쏘아올린 작은 공」과⁵ 그 영어 대역 말뭉치인 「The Dwarf」로, 「The Dwarf」는 ‘Bruce and Ju-Chan Fulton’에 의해 번역되었고 하와이 대학 출판부에서 출판되었다. 30만 어절 규모의⁶ 한영 병렬(KK-KE⁷) 말뭉치인 <난쏘공 말뭉치>는 21세기 세종 계획 특수자료 구축 소분과 인 한영 병렬 말뭉치 팀에서 2000년에 구축되었다.

⁴ 현재 21세기 세종계획 한영 병렬 말뭉치에는 소설, 영어 교과서, 연설문, 성경, 잡지, 매뉴얼 등이 포함되어 있다. 본 연구에서는 조사 ‘가’와 ‘는’의 본래의 성격이 잘 드러나려면 새로운 사건의 도입이 많으면서도 기존 사건에 대한 언급 역시 많이 이루어지는 장르가 적합하다고 보아 텍스트 장르를 소설로 선택하였다.

⁵ 현재 한영 병렬 말뭉치의 소설 장르 중에는 ‘난장이가 쏘아 올린 작은 공(1976)’ 외에도 ‘The Door in the wall(1911), The Desert Islander(1930), 암사지도(1956), 오발탄(1959), Animal Farm(1975), 피비우스의 띠(1975), 칼날(1976)’ 등이 포함되어 있다. 이처럼 1900년대 초반-1970년대 출판물로 자료가 한정된 것은 저작권 해결의 어려움이라는 현실적인 한계 때문이다. 이 중 ‘난장이가 쏘아올린 작은 공’은 가장 후대의 소설에 속한다. ‘난장이가 쏘아올린 작은 공 한영 병렬 말뭉치’는 이하 <난쏘공 말뭉치>로 줄여 부른다.

⁶ <난쏘공 말뭉치>는 약 30만 어절로, 그 규모에 대한 문제가 제기될 수도 있다. 현재 한영 병렬 말뭉치를 다루기 위한 마땅한 검색 도구가 개발되어 있지 않은 실정이므로 한영 병렬 말뭉치를 활용할 때에는 수작업이 필요하다. 적절한 검색 도구가 개발된다면 분석 대상 말뭉치의 규모를 좀 더 확장시킬 수 있을 것이다. 말뭉치의 규모 문제는 특수 말뭉치를 이용한 연구가 일반적으로 가지고 있는 한계점이다. 이미 다양한 검색 기능을 갖춘 검색 도구들이 여럿 개발되어 있는 일반 말뭉치의 경우에는 1,000만 어절 규모 이상을 다루는 것도 어렵지 않지만, 구어 말뭉치나 한영 병렬 말뭉치 같은 특수 목적 말뭉치의 경우에는 검색 도구의 미비로 대규모의 자료를 다루는 데에 아직 한계가 있다. 본 논문과 같이 특수 말뭉치를 이용한 연구들이 활발해져서 도구 개발이 이루어진다면 이러한 한계들은 극복될 수 있을 것이다.

⁷ KK는 한국어 원본 말뭉치를, KE는 한국어 원본을 토대로 한 영어 대역 말뭉치를 나타낸다. 한국어와 영어의 병렬 말뭉치는 흔히 한영 병렬 말뭉치라 불리지만 엄밀히 말하면 한국어 원본과 영어 대역본으로 된 한영 병렬(KK-KE) 말뭉치와 영어 원본과 한국어 대역본으로 된 영한 병렬(EE-EK) 말뭉치가 있는 것이다.

본 연구에서는 이들 말뭉치로부터 다음의 조건을 만족시키는 KK 문장과 KE 문장의 대응 쌍을 추출하여 분석의 대상으로 삼았다.

(4) 문장 대응 쌍 선정의 기준⁸

- a. KK 문장은 반드시 주어 자리에 실현된 ‘가’나 ‘는’을 포함한다.
예) 가. 그가 사나이를 맞아 악수하고 함께 차 안으로 들어갔다.
예) 나. 나는 낙원구에서 내렸다.
- b. KE 문장에 관사가 나타나지 않아도 대응되는 KK 문장에 ‘가’나 ‘는’이 있으면 분석 대상에 포함시킨다.
예) 가. 귀여운 영희의 얼굴은 눈물로 젖었다.
예) 가’. Yǒng-hŭi’s cute face was soaked with tears.
- c. 번역 과정에서 문장의 구조가 바뀐 경우는 분석 대상에서 제외하되, KK 문장의 주어가 그대로 KE 문장의 주어 혹은 다른 성분으로 쓰인 경우는 분석 대상으로 삼았다.
예) 가. 작은 목소리라 나는 알아들을 수 없었다.
예) 가’. They spoke softly and I couldn’t understand them.
- d. 이어진 문장과 안긴 문장에서 주어가 실현되어 있을 경우, 이어진 문장과 안긴 문장도 각각 분석의 대상으로 삼았다.
예) 가. 그는 자라면서 더욱 강해졌지만 우리는 자라면서 반대로 약해졌다.
예) 가’. He had become stronger as he grew, but we were the opposite—we weakened.
예) 나. 영호는 내가 방금 물러선 게시판 앞으로 갔다.
예) 나’. Yǒng-ho was approaching the bulletin board I’d just left.

(4)의 기준들 중에서 a-c은 <난쏘공 말뭉치>의 KK 문장들 중 주어 자리에 ‘가’와 ‘는’이 실현된 모든 문장을 살펴본다는 것으로, 본 연구가 한국어 조사와 영어 관사의 대응을 살피고 있지만 그 목적이 한국어 조사의 양상 파악에 있음을 보여준다. 그리고 d의 예문들과 같이 말뭉치 상에서 한 문장으로 나타나더라도 ‘가’나 ‘는’을 포함한 주어가 둘 이상 나타날 경우, 둘 이상의 주어 각각이 연구 대상이 되므로, 둘 이상의 문장을 각각 독립된 문장처럼 나누어 분석하였다.

(4)와 같은 기준에 따라 <난쏘공 말뭉치>로부터 뽑아낸 문장의 대응 쌍은 총 973 쌍으로, 각 대응 쌍의 가설 일치 여부를 살펴본 결과를 간략히 보이면 [표 3]과 같다.

⁸ 이러한 기준에 의해 주어가 아닌 자리에 ‘가’나 ‘는’이 쓰인 KK 문장들이 배제되어, 보어 자리에 쓰인 ‘가’나 목적어에 결합된 ‘는’ 등이 분석 대상에서 제외될 수 있었다. 또한 KE 문장에는 관사가 나타나지 만 KK 문장의 주어 자리에는 조사가 나타나지 않은 문장 대응 쌍도 배제되었는데, 이는 본 연구의 연구 목적이 영어 관사에 대한 분석보다는 국어 조사 ‘가’, ‘는’에 초점을 두고 있기 때문이다.

가설 일치 여부		수치	
일치	가 - a	17쌍 (1.75%)	8.33%
	는 - the	64쌍 (6.58%)	
불일치	가 - the	88쌍 (9.04%)	91.67%
	는 - a	1쌍 (0.10%)	
	가/는 - 무관사	803쌍 (82.53%)	

[표 3] 조사와 관사의 대응 양상

[표 3]에 나타난 바와 같이, ‘가’와 부정관사, ‘는’과 정관사가 대응되어 가설에 일치 되는 경우는 10%에도 못 미쳤으며, 가설 불일치의 경우가 90% 이상으로 나타났다. 그러나 이러한 수치는 본 연구의 대상 말뭉치에 나타난 조사와 관사의 대응의 대략적인 면모만을 보여줄 뿐 수치 자체가 절대적인 의미를 갖는 것은 아니다. 이제 각 대응 쌍의 유형을 예문을 통해 간단히 살펴보겠다.

(5) 가설 일치 문장 쌍

- a. 큰 공장이 문을 닫으면 수많은 공원들은 갈 곳이 없었다.
- a'. If a large plant closed its doors numerous workers would have no place to go.
- b. 영희가 치는 기타 소리는 영희에게 아주 잘 어울렸다.
- b'. The sound she produced on the guitar was a perfect complement to her.

(5a)와 (5a')는 ‘가’와 ‘a’의 대응을, (5b)와 (5b')은 ‘는’과 ‘the’의 대응을 보여준다. 한편, (6)의 예문들은 조사와 관사가 역으로 대응되는 경우에 해당되는데, (6a)와 (6a')는 ‘가’와 ‘the’의 대응을, (6b)와 (6b')는 ‘는’과 ‘a’의 대응을 보여준다.

(6) 가설 불일치 중 역대응 문장 쌍

- a. 동네 사람들이 골목으로 나와 뭐라고 소리치고 있었다.
- a'. The neighbors had come out into the alley and were shouting about something.
- b. 이런 사람들만 사는 땅은 죽은 땅입니다.
- b'. A land with such people and no others is a dead land.

(5)의 예문에 나타난 것과 같은 가설 일치 문장 쌍의 경우, ‘는’과 ‘the’의 대응은 64쌍, ‘가’와 ‘a’의 대응은 17쌍으로 ‘는’과 정관사의 대응이 ‘가’와 부정관사의 대응보다 더 많이 나타났으며, 가설 불일치 중 (6)과 같이 조사와 관사의 대응이 역으로 나타난 경우, ‘는’과 ‘a’의 대응은 (6b)-(6b')의 단 1쌍인 반면, ‘가’와 ‘the’의 대응은 88쌍으로 나타났다. 이를 통해 ‘가’는 부정관사뿐만 아니라 정관사에 대응되는 경우도 많지만, ‘는’은 거의 정관사에만 대응됨을 알 수 있었다. 이는 ‘는’과 ‘the’의 관계가 ‘가’와 ‘a’의 관계보다 더 긴밀하다는 암시를 주는데 이에 대해서는 3.2에서 상세히 살펴볼 것이다.

(7) 가설 불일치 중 무관사 대응 문장 쌍<1>: 고유명사, 대명사

- a. 나는 명희와의 약속을 지킬 수 없었다.
- a'. I couldn't keep the promise I'd made to Myöng-hüi.
- b. 영호는 내가 방금 물러선 게시판 앞으로 갔다.
- b'. Yöng-ho was approaching the bulletin board I'd just left.
- c. 아버지와 어머니는 내가 공부를 계속하기를 바랐다.
- c'. Mother and Father hoped I'd continue my studies.

한국어 주어 자리에 나타나는 ‘NP-가/는’과 영어의 무관사 명사구의 대응은 특정 명사구 앞에는 관사가 올 수 없다는 영어의 제약으로 인한 것으로, 본 연구 결과에서는 무관사 대응의 대부분이 대명사와 고유명사의 경우였다. (7a)-(7a')는 대명사의 경우이고, (7b)-(7b'), (7c)-(7c')는 고유명사의 경우에 해당된다. 여기서 (7c)-(7c')의 경우는, ‘Father’, ‘Mother’, ‘Elder Brother’가 문장에서의 위치에 관계없이 그 첫 글자가 대문자로 나타나기 때문에 고유명사로 분석한 경우로, 이는 <난쏘공 말뚱치>를 구성하는 텍스트의 특수성으로 볼 수 있다.

(8) 가설 불일치 중 무관사 대응 문장 쌍<2>: 기타

- a. 작은 공장들이 채용할 인원은 한정이 되어 있다.
- a'. The number of employees that small plants could take on was limited.
- b. 그의 승용차는 게시판 앞에 세워져 있었다.
- b'. His car was parked in front of the bulletin board.
- c. 약당은 돈이나 많지.

c'. Bad guys have lots of money and stuff.

d. Sleep overcame me.

d'. 잠이 나를 눌러 왔다.

대명사나 고유명사 외에도 무관사 대응의 유형에는 (8a)-(8a')와 같이 비한정적 복수형 명사 앞에 관사가 오지 않는 경우, (8b)-(8b'), (8c)-(8다')와 같이 명사 앞에 소유격이나 형용사가 와서 관사가 나타나지 않는 경우, (8d)-(8d')와 같이 추상명사 앞에 관사가 쓰이지 않는 경우가 있었다.

이상으로 한영 병렬 말뭉치에 나타난 조사 '가', '는'과 정관사, 부정관사의 대응에 대해 살펴보았다. [표 3]에 나타난 바와 같이 무관사 대응의 경우가 분석 문장 쌍 중 가장 높은 비중을 차지하지만, 이는 영어의 특성으로 인해 관사가 나타나지 않는 것으로, 한국어-영어의 대응 쌍을 통해 조사 '가', '는'의 특성을 살펴보려고 하는 본 연구의 연구 대상으로는 적절하지 않다. 따라서 3.2에서는 무관사 대응의 유형을 제외한 나머지 유형들에 대해 자세히 살펴보려고 한다.

3.2 조사-관사 대응 관계 분석

3.2에서는 [표 3]에 나타난 대응 쌍의 각 유형들을 의미적, 화용적, 통사적 측면에서 분석해 보았다.⁹ 의미적 측면에서는 2장에서 이미 살펴본 바와 같이 신창순 (1975), 이필영 (1982), 고석주 (2002) 등에서 '가'의 의미로 제시한 '선택지정', 신창순 (1975), 채완 (1977), 최수영 (1984) 등에서 '는'의 의미로 제시한 '대조'의 두 가지 측면을 살펴보았다. 또한 화용적 측면에서는 유동석 (1984)의 '통보기능량'의 관점에서의 논의, 류구상 (1986)의 신정보와 구정보에 대한 논의를 참고하여 정보량에 따른 조사와 관사의 선택에 대해 살펴보았으며, 이와 더불어 주제¹⁰ 전환의 표지로서의 조사 선택에 대해서도 살펴보았다. 이인영 (1996, 84)에 언급된 바와 같이 주제가 바뀔 경우에는 주제의 전환으로 인한 청자의 부담을 줄이고자 화자가 특정한 방법을 통해 주제가 바뀌었음을 나타내게 되는데, 본 연구에서는 조사의 선택에도 이와 관련되는 면이 있다고 보았다. 통사적 측면에서는 채완 (1977, 106), 최수영 (1984, 243) 등의 논의에서 언급된 내포문

⁹ 의명의 심사위원회에서는 관계절이나 수식어구의 존재 여부에 대한 분석이 이루어져야 한다고 점을 지적하였는데, 한확성 (1995, 143-144)에서 언급된 바와 같이 최상급 형용사의 수식을 받는 명사나 'She is the mother of my best friend'와 같이 of에 의해 수식을 받는 명사는 지시 내용이 유일한 것으로 결정되어 the를 사용할 수 있다. 그러나 한국어에서는 수식어구로 인해 조사 선택이 달라지는 현상이 발견되지 않으므로 이를 분석 기준에 넣지 않았다.

¹⁰ '주제'는 '화제'라는 용어로도 쓰이고 있지만, 본 연구에서는 '주제'와 '화제'를 구별하여 설명한 이인영 (1996)의 논의를 따라 '화제'가 아닌 '주제'를 선택하였다. 이인영 (1996)은 '주제'와 '화제'는 모두 이야기의 중심이 되는 대상이지만 '주제'가 청자와 화자에게 친숙한 것인 반면 '화제'는 새로운 대상이라고 하였다. 따라서 기존의 이야기에서 새로운 대상에 대한 이야기로 넘어가는 것은 '화제의 전환'이라고 해야 하는 반면, 한참 전에 언급되었던 대상이 다시 이야기의 중심으로 떠오를 때에는 '주제의 전환'이 이루어졌다고 봐야 한다(이인영, 1996, 84). 본 연구의 분석 자료에 나타난 '전환'은 한참 전에 언급되었던 대상, 즉 구정보인 대상에 대한 이야기로 다시 돌아가는 것이었으므로 이를 '주제의 전환'으로 부른 것이다.

주어 제약을 살펴보았다. 내포문 주어 제약이란 내포문의 주어는 그것이 대조의 대상이 아닌 경우에는 언제나 ‘가’로 나타난다는 것이다.

(9) 의미 층위¹¹

- ‘가’의 의미: 선택지정
- ‘는’의 의미: 대조

(10) 화용 층위

- 정보량: 신정보, 구정보
- 주제의 전환

(11) 통사 층위

- 내포문 주어 제약

(9)–(11)의 분석틀은 한국어의 조사 ‘가’, ‘는’에 대한 앞선 연구들을 토대로 한 것이므로 실제 분석을 진행하기에 앞서 영어의 관사 선택에 대한 앞선 연구들을 통해 정관사와 부정관사의 선택 원리를 밝히기 위한 분석틀 역시 마련해 두어야 할 것이다. 영어의 관사에 대한 앞선 연구들을 살펴보면, 한학성 (1995)에서는 관사 다음의 명사가 가리키는 지시 내용이 문맥 또는 기타 이유로 결정되어 있을 때는 정관사를 쓰고 그렇지 않을 때는 부정관사를 쓴다고 하였으며,¹² ‘the’가 사용되는 경우로 ‘문맥상 지시 내용이 결정되는 경우(앞서 언급된 대상, 혹은 그와 밀접한 관련을 갖는 대상)’, ‘상황에 의해 지시 내용이 결정되는 경우(눈에 보이는 사물, 주어진 상황에서 당연히 기대할 수 있는 대상)’, ‘유일한 사물을 지시하는 경우’, ‘수식에 의해 지시 내용이 결정되는 경우(유일성 나타내는 형용사 수식, 최상급 형용사 수식, of에 의한 수식)’로 나누어 설명하였다. 이는 모두 ‘지시 내용이 이미 결정되어 있는 경우’라는 점에서 공통적이고, 본 연구의 ‘정보량이 높은 경우 구정보의 경우’와 상통한다. 또한 박혜숙 (2001, 392)에서는 관사의 화용 영역에 대해 설명하면서 관사는 화자와 저자의 독립적인 결정으로 만들어지는 것이 아니라 실제 진행되는 의사소통에서 대화자 간의 공유지식의 반영이라고 하였다. 이렇게 볼 때 영어의 관사 선택은 화용 층위에서 이루어지기 때문에 (9)–(11)의 분석틀 외의 다른 기준을 추가하지 않는다.

¹¹ 본문에서 언급한 의미 층위의 ‘선택지정’과 ‘대조’는 문장 차원의 현상이 아니라 화용 층위에 속하는 개념으로 볼 수도 있다. 그러나 ‘딸기는 좋은데 포도는 싫다’의 ‘는’이 갖는 대조의 의미를 화용적 의미로 본다면 ‘좋다’, ‘싫다’의 의미 역시 화용적 의미로 보아야 한다. 즉, 딸기와 포도의 무엇이 좋고 싫은 것인지(색깔, 맛, 모양...)는 앞뒤의 문맥이나 반화의 상황에 대한 정보가 없이는 알 수 없다. 이는 ‘이, 그, 저’와 같은 관형사나 ‘이것, 저것, 그것’과 같은 대명사의 경우도 마찬가지이다. 이들 관형사나 대명사의 의미는 문맥이나 상황적 맥락에 따라 달라질 수밖에 없다. 의미를 어휘 의미에 한정하는 어휘미론에서 본다면 이러한 의미는 의미 층위에 속한다고 볼 수 없겠지만, 의미를 보다 넓은 개념으로 보는 형식미론의 관점에서는 ‘선택지정’, ‘대조’와 같은 조사의 의미 역시도 의미라고 볼 수 있다. 그러므로 본 연구에서는 의미 층위는 화용적 의미까지를 포함하는 것으로 본다.

¹² 한학성 (1995)에서는 관사와 관련한 핵심 사항으로 첫 번째, 수의 문제와 관련한 명사의 가산성, 두 번째, 지시 내용의 결정 여부 문제를 들었는데, 이 중 첫 번째 문제는 부정관사와 무관사에 해당되는 경우이며 두 번째가 정관사와 부정관사에 해당되는 것이다.

분석의 결과를 먼저 간단히 요약하면, 영어에서 정관사와 부정관사의 선택에 있어서는 ‘정보량’이 절대적인 기준으로 작용한 반면, 한국어에서는 ‘정보량’ 외에도 여러 가지 요인에 따라 조사의 선택이 결정됨을 알 수 있었다. [표 3]에서 가설 일치를 보이는 ‘가-a’ 대응의 17쌍과 ‘는-the’ 대응의 64쌍은 거의 대부분이 (9)-(11)의 항목들 중 ‘정보량’에만 해당되었으며 몇몇 예에서 일반적인 사실을 나타낼 때 KK 문장에서 ‘는’이, KE 문장에서 ‘the’가 쓰인 경우가 있었다.¹³ 한편 가설에 어긋나며 역대응으로 나타나는 ‘가-the’ 대응의 88쌍은 구정보이지만 주제 전환, 선택지정, 내포문 주어 제약으로 인해 ‘는’이 아닌 ‘가’가 선택된 경우이다. 또한 ‘는’과 ‘a’ 대응의 1쌍은 신정보이지만 대조의 대상이므로 ‘가’가 아닌 ‘는’이 선택된 경우이다. 가설 일치의 경우는 모두 정보량으로 설명이 되므로 더 이상 논의하지 않고, 3.2.1-3.2.4에서는 조사와 관사가 가설과 달리 역대응된 경우를 중점적으로 살펴보겠다.

13

- a. 싸움은 언제나 옳은 것과 옳지 않은 것이 부딪쳐 일어나는 거야.
 a'. **The fight** is always a conflict between what's right and what's wrong.
 b. 세상은 공부를 한 자와 못 한 자로 너무나 엄격하게 나뉘어져 있었다.
 b'. **The world** was divided too arbitrarily between those who had studied and those who hadn't.

3.2.1 가-the: 주제 전환¹⁴.

(12) KK-KE 문장 대응 쌍¹⁵

KK	<p><p id=1.1.p393><s id=1.1.p393.s1>「팔지 말고 기다려요.」</s></p> <p id=1.1.p394><s id=1.1.p394.s1>승용차 안에서 ⊖<u>한 사나이가</u> 말했다.</s></p> <p id=1.1.p449><s id=1.1.p449.s2>그날 밤 ⊖<u>승용차 안의 사나이가</u> 우리 동네의 나머지 입주권을 모두 사 버렸다.</s><s id=1.1.p449.s3>그는 다른 투기업자들이 이십 이만 원에 사는 것을 이십 오만 원씩 주고 모두 사 버렸다.</s><s id=1.1.p449.s4>그날 밤에도 영희는 팬지꽃 앞에 앉아 기타를 쳤다.</s><s id=1.1.p449.s5>영희는 팬지꽃 두 송이를 따 하나는 기타에 꽂고 하나는 머리에 꽂았다.</s><s id=1.1.p449.s6>그리고, 꼼짝도 하지 않고 기타만 쳤다.</s><s id=1.1.p449.s7>⊖<u>사나이가</u> 아버지에게 담배를 권했다.</s></p></p>
KE	<p><p id=1.1.p285><s id=1.1.p285.s1>"Wait before you sell," said ⊕<u>a man in a car.</u></s> <p id=1.1.p317><s id=1.1.p317.s1>That night ⊕<u>the man in the car</u> bought up the occupancy rights of all the neighbors who still had them.</s><s id=1.1.p317.s2>He bought them all up at two hundred fifty thousand apiece;</s><s id=1.1.p317.s3>other brokers had paid two hundred twenty thousand.</s><s id=1.1.p317.s4>Again that night Yǒng-hüi sat in front of the pansies playing her guitar.</s><sid=1.1.p317.s5>She picked two of the pansies and stuck one in the guitar and the other in her hair.</s><s id=1.1.p317.s6>She didn't budge, merely played the guitar.</s><s id=1.1.p317.s7>⊖<u>The man</u> offered Father a cigarette.</s></p></p>

(12)에서 ‘⊖승용차 안의 사나이’는 ⊖의 사나이와 동일 인물로 정보량의 관점에서 보면 구정보에 해당된다. 이때 KE 문장에서는 ‘the’가 쓰였으나 KK 문장에서는 구정보를 나타내는 ‘는’ 대신 ‘가’가 쓰였다는 점이 주의를 끈다. 이 점은 ⊖과 ⊕에서도 마찬가지이다. KE 문장에서는 구정보이므로 ‘the’가 쓰였으나 KK 문장에서는 구정보임에도 불구하고 ‘가’가 쓰인 것은, 한참 전에 언급되었던 대상이 다시 주제로 떠오르는 과정에서 주제의 전환의 표지로서 ‘가’가 사용된 것이다. <p id=1.1.p449><s

¹⁴ 익명의 심사위원께서 주제 전환을 야기하는 중간 개입 담화의 길이나 구조적 특성에 대한 계량적인 혹은 기술적인 설명이 필요하다고 지적하였으나 이러한 내용은 본 논문 안에서 보충되기보다는 별도의 논문에서 다루어져야 한다고 보아 본 연구에서는 단순히 주제가 전환되었는가의 여부만을 살펴보았다.

¹⁵ <p id=1.1.p449><s id=1.1.p449.s2>는 각각 문단 id와 문장 id를 나타낸다. 즉, 문단 id <p id=1.1.p449>는 이 텍스트 내에서 449번째 문단임을 나타내며, 문장 id <s id=1.1.p449.s2>는 449번째 문단의 2번째 문장임을 나타낸다.

id=1.1.p449.s2> 이후를 살펴보면, 처음엔 사나이에 대한 이야기를 하다가 <s id=1.1.p449.s4>~<s id=1.1.p449.s6>에서 ‘영희’로 주제가 전환되었으므로 <s id=1.1.p449.s7>에서 다시 주제를 바꿔 ‘사나이’에 대한 이야기로 다시 돌아오기 위해서는 독자들의 주위를 환기시킬 필요가 있고, 이러한 이유로 ‘가’가 쓰인 것이다.¹⁶

(13) KK-KE 문장 대응 쌍

<p>KK</p>	<p><s id=1.1.p452.s4>㉠용달차가 좁은 골목을 뚫고 들어와 명희네 집을 실었다.</s><s id=1.1.p452.s5>명희 어머니가 치마를 올려 눈물을 닦았다.</s></p> <p id=1.1.p453><s id=1.1.p453.s1>「에유, 정이란 게 뭔지!」</s></p> <p id=1.1.p454><s id=1.1.p454.s1>명희 어머니가 말했다.</s></p> <p id=1.1.p455><s id=1.1.p455.s1>「정이란 게 이렇게 더러운 게라우.」</s></p> <p id=1.1.p456><s id=1.1.p456.s1>그 말이 우리의 눈에 고춧가루를 뿌렸다.</s><s id=1.1.p456.s2>㉡용달차가 집 앞을 지나갔다.</s><s id=1.1.p456.s3></p>
<p>KE</p>	<p><s id=1.1.p320.s4>㉠A moving truck threaded its way into the narrow alley and loaded Myöng-hŭi's family's belongings.</s><s id=1.1.p320.s5>Myöng-hŭi's mother wiped away her tears with the hem of her skirt.</s></p> <p id=1.1.p321><s id=1.1.p321.s1>"Isn't it funny!" she said with a great sigh.</s><s id=1.1.p321.s2>"The thing that makes us close makes a time like this so difficult."</s></p> <p id=1.1.p322><s id=1.1.p322.s1>These words were pepper in our eyes.</s><s id=1.1.p322.s2>㉡The moving truck went past our house.</s></p>

(13)의 예문들에서 ‘용달차’와 ‘moving truck’은 ㉠과 ㉡에서 처음 나타난다. ㉠과 ㉡는 이미 언급되었던 구정보로, KE 문장에서는 ‘the’가 쓰인 반면, KK 문장에서는 구 정보임에도 불구하고 ‘가’가 쓰였다. 이 경우에도 <s id=1.1.p452.s4>에서 용달차 이야기가 나온 이후에 <p id=1.1.p453><s id=1.1.p453.s1>에서 <p id=1.1.p456><s id=1.1.p456.s1>까지 다른 이야기가 진행된다. 그리고 다시 <s id=1.1.p456.s2>에서 용달차에 대한 이야기로 돌아온다. 이때의 ‘가’는 (12)의 ㉠, ㉡과 마찬가지로 주제 전환에 따라 독자의 주위를 환기하는 기능을 한다.

¹⁶ (12)와 (13)에서 영어의 경우를 보면 한번 등장한 뒤 다시 언급되는 명사들이 앞의 명사와 몇 문장 혹은 몇 단락의 거리를 두고 나타날 때 구정보를 표시하기 위해 ‘the’가 쓰였음을 알 수 있다. 한학성 (1995, 135)에서는 앞에서 언급된 명사를 반복할 때 나타나는 the는 반복이 바로 연이어 나타나는 경우보다 상당한 거리를 두고 반복되는 경우에 주로 사용되며, 바로 연이은 반복의 경우에는 this가 쓰이거나 대명사로 나타나는 것이 보다 일반적이라고 하였다.

이상의 논의를 통해 영어에서는 정보량에 따라 관사가 선택되는 경향이 짙은 반면, 한국어에서는 조사의 선택에 정보량 외에도 주제 전환이라는 다른 요인이 개입함을 알 수 있었다.

3.2.2 가-the: 선택지정.

(14) KK-KE 문장 대응 쌍

<p>KK</p>	<p><s id=1.1.p270.s5>내가 방송통신고교의 강의를 받기 위해 라디오를 사러 갈 때 영희가 따라왔었다.</s><s id=1.1.p270.s6>㉠<u>쓸 만한 라디오가</u> 있었다.</s><s id=1.1.p270.s7> 그런데, 영희가 먼지 속에 놓인 기타를 들어 통겨 보는 것이었다.</s><s id=1.1.p270.s8> 영희는 고개를 약간 숙이고 기타를 쳤다.</s><s id=1.1.p270.s9> 긴 머리에 반쯤 가려진 옆얼굴이 아주 예뻐다.</s><s id=1.1.p270.s10> 영희가 치는 기타 소리는 영희에게 아주 잘 어울렸다.</s><s id=1.1.p270.s11> 나는 ㉡<u>먼저 골랐던 라디오를</u> 살 수 없었다.</s><s id=1.1.p270.s12> 좀 더 싼 것으로 바꾸면서 영희가 든 기타를 가리켰다.</s><s id=1.1.p270.s13> ㉢<u>그 라디오가</u> 고장이 나고 기타는 줄이 하나 끊어졌다.</s></p>
<p>KE</p>	<p><s id=1.1.p194.s5>That’s where I had gone to buy a radio for my high school correspondence course lessons, and Yǒng-hŭi had tagged along.</s><s id=1.1.p194.s6>I had found ㉣<u>a radio in usable condition</u>.</s><s id=1.1.p194.s7>But Yǒng-hŏi had picked up a dust-covered guitar lying in the dust and tried it out.</s><s id=1.1.p194.s8>She bent over the instrument and strummed it.</s><s id=1.1.p194.s9>Her profile, half covered by her long hair, was so pretty.</s><s id=1.1.p194.s10>The sound she produced on the guitar was a perfect complement to her.</s><s id=1.1.p194.s11>I wouldn’t be able to afford ㉤<u>the radio</u> now.</s><s id=1.1.p194.s12>So I found a cheaper one and gestured toward the guitar Yǒng-hŭi held.</s><s id=1.1.p194.s13>㉥<u>The radio</u> had broken down and one of the guitar strings had snapped.</s></p>

(14)의 ㉠과 ㉡의 ‘라디오(radio)’는 ㉠-㉢과 ㉣-㉥에서 이미 언급되었으므로 구경 보이다. 또한 ㉠과 ㉡의 바로 앞 문장인 KK 문장 <s id=1.1.p270.s12>와 KE 문장 <s id=1.1.p194.s12>도 ‘라디오(radio)’에 대한 이야기이므로 주제가 전환된 것도 아니다. 그럼에도 불구하고 ㉠에서 ‘는’이 아닌 ‘가’가 쓰인 이유는, ㉠에서 ‘라디오’가 다른 사물이 아니라 바로 그 라디오로 선택지정되었기 때문이다.

(15) KK-KE 문장 대응 쌍

<p>KK</p>	<p><s id=1.1.p586.s6>안쪽에서 ㉠<u>사무장이</u> 일어서며 나를 손짓해 불렀다.</s><s id=1.1.p586.s7>㉡<u>그는</u> <행복 제1동장> 위에 직인을 찍었다.</s><s id=1.1.p586.s8>그것을 내주기 전에 나를 창가로 데리고 갔다.</s><s id=1.1.p586.s9>㉢<u>사무장은</u> 큰길 건너 포도밭 아랫동네를 가리켰다.</s></p> <p id=1.1.p587><s id=1.1.p587.s1>「위에서 세번째 집야.」</s></p> <p id=1.1.p588><s id=1.1.p588.s1>㉣<u>그가</u> 말했다.</s></p></p>
<p>KE</p>	<p><s id=1.1.p419.s7>㉤<u>The section chief</u> rose and beckoned me.</s><s id=1.1.p419.s8>㉥<u>He</u> stamped his official seal above the words "Chief, Felicity Precinct 1."</s><s id=1.1.p419.s9>Before handing it to me, he took me to the window.</s><s id=1.1.p419.s10>㉦<u>He</u> indicated a neighborhood below a grape patch, across the main street.</s></p> <p id=1.1.p420><s id=1.1.p420.s1>"Third house from the upper end," ㉧<u>he</u> said.</s></p>

선택지정으로 인해 구정보임에도 ‘가’가 결합되는 현상은 (15)의 ㉣에도 해당된다. 대사가 나오고 그 대사를 말한 사람이 누구인지를 밝히는 문장 ‘그가 말했다’에서 조사 ‘가’는 다른 등장인물이 아닌 대사를 말한 바로 그 사람을 특별히 지정하는 것이다. ㉤의 ‘그’는 ㉠-㉡에 이미 언급된 ‘사무장’에 해당되지만, ‘위에서 세 번째 집야’라고 말한 사람이 누구인지를 선택적으로 지정하는 문장에서는 선택지정의 의미를 갖는 ‘가’가 쓰이는 것이다. 그러나 KE 문장에서는 이런 경우에 ‘the’ 대신 ‘a’가 쓰인 경우가 없었다.

3.2.2의 논의를 통해 영어와 달리 한국어에서는 구정보일지라도 주어가 선택지정의 의미를 가지면 ‘는’이 아닌 ‘가’가 쓰임을 알 수 있었다.

3.2.3 가-the: 내포문 주어 제약.

(16) KK-KE 문장 대응 쌍

KK	<p><s id=1.1.p451.s2>㉠<u>사나이를 따라온 나이 든 사람이</u> 검은 가방을 열어 돈을 보여 주었다.</s></p> <p>.....</p> <p><s id=1.1.p505.s6>처음 일이 떠올랐다.</s><s id=1.1.p505.s7>그는 [㉡<u>나이 든 사람이</u> 매매 계약서를] 쓰는 동안 내 옆에서 있었다.</s></p>
KE	<p><p id=1.1.p319><s id=1.1.p319.s1>㉡<u>An older man who had accompanied the man in the car</u> opened a black briefcase and displayed the money.</s></p> <p>.....</p> <p><s id=1.1.p358.s6>I thought back to the beginning.</s><s id=1.1.p358.s7>He had stood next to me while [㉢<u>the older man</u> wrote out the sales contract].</s></p>

(17) KK-KE 문장 대응 쌍

KK	<p><p id=1.1.p4><s id=1.1.p4.s1>「그게 뭐냐?」</s></p></p> <p><p id=1.1.p5><s id=1.1.p5.s1>「㉠<u>철거 계고장</u>예요.</s></p></p> <p><p id=1.1.p6><s id=1.1.p6.s1>「기어코 왔구나!」</s></p></p> <p><s id=1.1.p505.s9> [㉡<u>철거 계고장이</u> 나온] 날 내가 동사무소 앞으로 달려갔을 때부터 그는 나를 보았다.</s></p>
KE	<p><p id=1.1.p4><s id=1.1.p4.s1>"What is it?" she asked.</s></p></p> <p><p id=1.1.p5><s id=1.1.p5.s1>"㉡<u>A condemnation notice.</u>"</s></p></p> <p><p id=1.1.p6><s id=1.1.p6.s1>"So, finally," Mother said.</s></p> <p>.....</p> <p><s id=1.1.p358.s9>He had noticed me running up to the precinct office the day ㉢[<u>the condemnation notice</u> arrived].</s></p>

(16)과 (17)에서 각각의 ㉠은 구정보인 명사에 ‘가’가 결합된 경우로, (16)과 (17)의 ㉡은 []로 묶인 내포문의 주어이다. 채완 (1977, 106)에서 언급된 대로, 내포문의 주어는 그것이 특별히 대조의 의미를 갖지 않는 한 언제나 ‘가’와 결합한다. 따라서 해당 명사가 구정보임에도 ‘가’가 결합되는 요인의 한 유형으로 내포문 주어 제약을 들 수 있을 것이다.

3.2.4 는-a: 대조.

(18) KK-KE 문장 대응 쌍

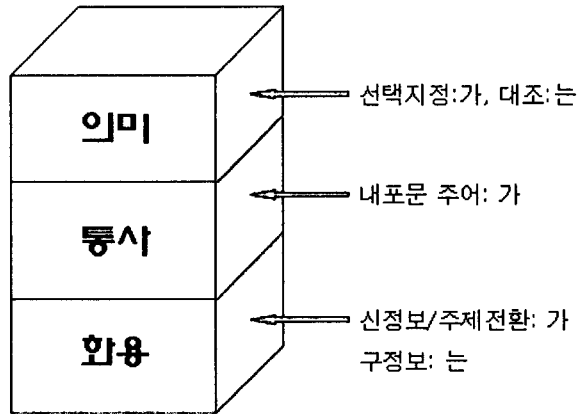
<p>KK</p>	<p><p id=1.1.p274><s id=1.1.p274.s1>「사람들은 사랑이 없는 욕망만 갖고 있습니다.</s><s id=1.1.p274.s2>그래서 단 한 사람도 남을 위해 눈물을 흘릴 줄 모릅니다.</s><s id=1.1.p274.s3>⊙<u>이런 사람들만 사는 땅은</u> 죽은 땅입니다.</s></p></p> <p>..</p> <p><s id=1.1.p282.s4>이제 이 죽은 땅을 떠나야 됩니다.</s></p></p> <p><p id=1.1.p283><s id=1.1.p283.s1>「떠나다니?</s><s id=1.1.p283.s2>어디로?</s></p></p> <p><p id=1.1.p284><s id=1.1.p284.s1>「달나라로!</s></p></p>
<p>KE</p>	<p><p id=1.1.p196><s id=1.1.p196.s1>"Because the only thing people have are loveless desires." said Chi-söp.</s><s id=1.1.p196.s2>"There is not a soul who knows what it is to shed tears for others.</s><s id=1.1.p196.s3>ⓐ<u>A land with such people and no others</u> is a dead land.</s></p></p> <p>..</p> <p><s id=1.1.p204.s4>Now you must leave this dead land.</s></p></p> <p><p id=1.1.p205><s id=1.1.p205.s1>"Leave?</s><s id=1.1.p205.s2>Where to?</s></p></p> <p><p id=1.1.p206><s id=1.1.p206.s1>"To the moon.</s></p></p>

(18)의 ⊙은 3.2.1-3.2.3에서 살펴본 것과 반대로, 신정보인데도 ‘는’이 사용된 예에 해당된다. (18)에 나타난 예문들은 난장이 가족의 이웃집 가정교사인 ‘지섭’이 난장 이에게 하는 말로, 사회주의 사상을 가진 지섭은 현재의 세상을 비판하고 혁명을 통해 더 나은 세상을 만들어야 한다고 주장한다. 여기서 ‘이런 사람들만 사는 땅’은 ‘더 나은 땅’을 의미하는 ‘달나라’와 대조되는 대상이다. 이렇게 신정보일지라도 대조의 의미를 담고자 할 때에는 ‘가’ 대신 ‘는’이 선택된다. (18)ⓐ은 ‘는’과 ‘a’의 대응을 보여주는 유일한 예로, <난쏘공 말뭉치>에서 대조의 의미를 나타내는 ‘는’은 거의 구정보를 나타내는 경우와 겹쳐 쓰였다.

3.3 한영 병렬 말뭉치를 통해 본 ‘가’와 ‘는’의 선택 원리

이상의 논의를 종합해 보면, 조사 ‘가’와 ‘는’의 선택에 있어서 의미, 통사, 화용의 세 층위가 일정한 위계를 가지고 작용함을 알 수 있었다. 3.2의 분석 결과를 요약해 보면, 의미 층위나 통사 층위에서의 해당 사항이 없는 경우에는 화용 층위의 정보량이나 주제 전환 여부에 따라 ‘가’나 ‘는’이 선택된다. 그러나 화용 층위와 더불어 의미 층위나 통사 층위에도 해당 사항이 있는 경우, 조사 선택은 의미나 통사 층위에 따라 이루어진다. 이렇게 볼 때 ‘가’와 ‘는’의 선택에 있어서 화용 층위는 의미 층위나 통사 층위보다 하위의

층위임을 알 수 있다. 또한 통사 층위와 의미 층위의 관계에서는 통사 층위가 의미 층위보다 낮은 층위라고 할 수 있다. 내포문 주어일 경우라도 대조의 의미가 있을 때에는 통사 층위의 내포문 주어 제약과 관계없이 대조의 의미를 나타내기 위해 ‘는’이 선택되기 때문이다. 이렇게 볼 때, ‘가’와 ‘는’의 선택에 있어서 의미 층위가 가장 상위의 층위에 해당되고, 그 다음이 통사 층위, 가장 낮은 층위가 화용 층위라는 결론을 내릴 수 있다. 이를 그림으로 나타내면 [그림 1]과 같다.



[그림 1] ‘가’와 ‘는’의 선택 층위

‘가’와 ‘는’의 선택은 [그림 1]에 나타난 각 층위 간의 위계에 따라 가장 상위인 의미 층위로부터 가장 하위인 화용 층위로 내려가며 각 층위 별 해당 사항이 있는가의 여부에 따라 이루어진다. 둘 이상의 층위에 해당 사항이 있을 경우에는 보다 상위 층위의 정보에 따라 조사가 선택되는 것이다.

[그림 1]에 정리된 각각의 층위별 조사 ‘가’와 ‘는’에 대해서는 이미 앞선 연구들에서 수차례 논의된 바 있다. 그러나 기존의 논의들에서 지적된 ‘가’와 ‘는’에 대한 기술들은 각각의 층위 중 일부 층위만을 인정하거나 혹은 층위를 뒤섞어 살펴보았기 때문에 일부의 ‘가’와 ‘는’에 대해서는 옳지만 조사 ‘가’와 ‘는’의 모든 용법을 설명할 수 없다는 점이 문제가 되어 왔다. 본 연구에서는 조사 ‘가’와 ‘는’의 선택에는 화용 층위, 의미 층위, 통사 층위가 모두 작용하되 그 층위간의 위계가 있다는 결론을 내렸다.

한국어의 틀 내에서 설명력 있는 해결책을 찾기 어려운 경우, 다른 언어와의 대조 연구를 통해 한국어의 특성을 파악해보는 것은 문제 해결의 실마리를 찾게 해 줄 수 있다. 이러한 연구를 위해 유용한 자료가 본 연구에서 활용한 한영 병렬 말뭉치와 같은 이중 언어 혹은 다중 언어 말뭉치이다. 본 연구에서 한영 병렬 말뭉치를 통해 조사와 관사의 대응 관계를 살펴본 결과, 영어의 관사 선택에는 화용 층위 중 ‘정보량’이 강력한 기

준으로 작용한 반면, 한국어의 조사 선택에는 ‘정보량’ 외에도 여러 가지 화용, 의미, 통사 층위가 영향을 미친다는 점을 알 수 있었다. 관사의 선택과 조사의 선택을 통해 볼 때, 영어는 그것이 처음 언급되는 대상이나 다시 언급되는 대상이나 하는 ‘사실 여부’가 결정적인 반면, 한국어에서는 ‘청자의 주의를 환기시키고자 하는가’, ‘해당 명사구를 선택적으로 지정하려고 하는가’, ‘다른 대상과 특별히 대조하여 말하고자 하는가’ 하는 ‘화자의 의도’가 강하게 작용한다는 점을 알 수 있었다.

3.4 영한 병렬(EE-EK) 말뭉치를 통한 검증

병렬 말뭉치를 토대로 연구를 진행할 때에는 번역의 방향이 고려되어야 한다. 한국어 텍스트를 원본으로 하는 한영 병렬(KK-KE) 말뭉치인지, 영어 텍스트를 원본으로 하는 영한 병렬(EE-EK) 말뭉치인지가 변수가 되어 연구 결과가 달라질 수도 있기 때문이다. 본 연구는 한국어 원본 텍스트와 그 영어 대역 텍스트로 이루어진 한영 병렬(KK-KE) 말뭉치를 토대로 이루어졌다. 따라서 3.4에서는 영어로 작성된 ‘조지 오웰(G. Orwell)’의 『Animal Farm』과 그에 대한 한국어 대역본인 『동물농장』으로 이루어진 영한 병렬(EE-EK) <Animal Farm 말뭉치>를 토대로 본 연구결과의 신뢰성을 검증해보고자 한다.¹⁷

이를 위해 <Animal Farm 말뭉치>에서 100쌍의 EE-EK 문장들을 뽑았다. 분석 대상 문장 쌍의 추출 기준은 (4)와 같다.

<난쏘공 말뭉치>와 <Animal Farm 말뭉치>에서 추출한 대응 쌍들의 가설 일치 여부 및 대응 유형에 따른 수치를 비교하여 제시하면 [표 4]와 같다.

가설 일치 여부		KK-KE(난쏘공)	EE-EK(동물농장)
일치	1	가 - a	1.75%
	2	는 - the	6.58%
불일치	3	가 - the	9.04%
	4	는 - a	0.10%
	5	가/는 - 무관사	82.53%
			72%

[표 4] KK-KE 말뭉치와 EE-EK 말뭉치의 분석 결과 비교

한영 병렬(KK-KE)과 영한 병렬(EE-EK) 말뭉치 모두에서 가장 큰 비중을 차지한 것은 무관사 대응의 경우이며, ‘는-a’ 대응이 가장 작은 비중을 차지했다는 점은 공통적이다. 그러나 <난쏘공 말뭉치>의 경우 그 비중이 5 > 3 > 2 > 1 > 4인 반면, <Animal Farm 말뭉치>의 경우 5 > 2 > 3 > 1 > 4로 나타나, 2와 3의 순서가 뒤바뀐 차이를 보인다. ‘는-the’ 대응의 비율이 한영 병렬(KK-KE)의 경우가 6.58%, 영한

¹⁷ <Animal Farm 말뭉치>는 21세기 세종 계획 특수자료 구축 소분과인 한영 병렬 말뭉치 팀에서 1999년에 구축된 말뭉치이다.

병렬(EE-EK)의 경우 14%로 2배가 넘는 수치의 차이를 보이는 것은 <Animal Farm 말뭉치>에서 구정보에 해당되는 주어의 출현이 잦았기 때문이며 이는 조사 선택 원리의 문제라기보다는 개별 텍스트의 성격과 관련된다.¹⁸ 이렇게 개별 텍스트의 특성으로 인한 차이 외에 번역의 방향에 따른 차이는 발견되지 않았다.

4. 결론 및 남은 문제

지금까지 한영 병렬 말뭉치에 나타난 한국어 조사 ‘가’, ‘는’과 영어 관사의 대응 양상을 살펴봄으로써 기존의 연구에서 제시한 조사 ‘가’와 ‘는’의 의미를 귀납적으로 검토해 보았다. 한국어에서 조사는 그 의미가 부각되지 않을 때는 정보의 양이나 주제의 전환과 같은 화용 층위에서 선택이 활성화되고 조사의 의미가 부각되거나 통사적인 동기가 있을 때는 화용 층위보다 의미, 통사 층위가 활성화됨을 알 수 있었다. 내포문의 주어와 같은 통사 층위에서의 조사 선택은 조사 ‘가’에만 해당되고 조사 ‘는’은 무관하다는 사실은 조사 ‘가’는 통사적인 기능을 가지고 있으나 조사 ‘는’은 화용적, 의미적인 기능만을 지닌다는 것을 보여 준다. 조사의 의미나 화용 혹은 통사적 기능의 어느 한쪽만을 부각시킨 기존의 논의들은 한영 병렬 말뭉치와 같은 자료를 설명할 수 없으며 한국어에서의 조사 선택은 의미, 화용, 통사의 세 가지 층위가 화자의 의도와 맞물려 상호 작용하고 있음을 알 수 있었다. 이러한 조사의 입체적인 의미와 기능에 대한 분석 결과는 한국어의 이론적 연구에도 기여할 수 있지만 한국어교육과 같은 응용 분야에 더욱 효과적으로 적용할 수 있다.

다음에서는 본 연구 과정에서 드러난 한영 병렬 말뭉치의 문제점을 몇 가지 지적하고자 한다. 이러한 점들이 개선된다면 언어 연구에서 한영 병렬 말뭉치의 활용 가능성이 보다 높아질 것으로 기대된다.

첫째로 한영 병렬 말뭉치의 주석에 관한 문제를 제기할 수 있다. 한영 병렬 주석 말뭉치를 위한 태그셋은 어절 분석을 위한 태그셋으로 품사 주석(part of speech annotation: POS annotation)을 위주로 한다. 그러나 품사 단위의 주석만으로는 이 말뭉치가 언어 연구를 위해 활용되는 데 있어 상당한 제약이 따를 것으로 보인다. 한 예로, 본 연구에서 주어에 나타난 ‘가’나 ‘는’ 혹은 ‘a’나 ‘the’를 찾기 위해 주석 말뭉치를 활용하려 하였으나, ‘가/JKS’, ‘는/JX’로 검색을 하면 ‘는’의 경우는 주어의 위치가 아닌 경우도 모두 검색이 되고, 영어 관사 ‘a/DT’, ‘the/DT’의 경우도 마찬가지였다.¹⁹ 즉, 문장 성분이나 문장 구조 차원에서의 주석이 이루어져 있지 않기 때문에 주석 말뭉치를 활용하지 못하고 원시 말뭉치를 처음부터 끝까지 읽어가며 수작업으로 분석 대상 문장을 찾아내야 했다. 한영 병렬 주석 말뭉치가 주석 말뭉치로서 언어 연구에 제대로 활용되

¹⁸ <난쏘공 말뭉치>나 <Animal Farm 말뭉치>의 경우 하나의 작품만으로 구성된 말뭉치이므로 해당 텍스트의 특수성이라는 제약에서 자유로울 수 없었다. 이러한 문제점을 해결하기 위해서는 적정 크기의 다양한 표본들로 구성된 대규모 한영 병렬 말뭉치를 토대로 연구를 진행해야 할 것이다.

¹⁹ JKS는 주격조사, JX는 보조사, DT는 관사를 나타내는 태그이다.

기 위해서는 품사 주석을 넘어서서 문장 구조, 더 나아가다면 담화 구조까지 반영한 입체적인 주석 말뭉치가 구축되어야 할 것이다.

둘째로 들 수 있는 문제는 한영 병렬 말뭉치 검색 프로그램에 관한 것이다. 현재 한영 병렬 말뭉치를 위한 검색 프로그램은 hepman 한 종류밖에 없다. hepman으로는 문장 단위 대응 쌍을 검색할 수 있으며 어절의 빈도를 추출할 수 있다. 이는 번역을 위한 자료로서는 쓰임새가 있겠지만, 언어 연구 프로그램으로서는 한계가 있다. 한국어와 영어의 주제부 실현 양상을 연구한다든지, 영어의 접속사와 한국어의 어미의 대응 관계를 연구한다든지 할 때에는 문장 단위가 아닌 문장 성분 혹은 형태소 단위의 검색이 가능해야 한다.

셋째, 텍스트의 특수성과 말뭉치의 저작권 해결이 시급한 문제점이다. 이번 <난소공 말뭉치>의 분석은, <난소공 말뭉치>의 문체적 특성이 뚜렷하여 다른 무관사 및 대명사 대응이 많은 경우에 속하는 텍스트일 수 있다. 그러나 특히 소설의 경우 까다로운 저작권의 문제로 인하여 말뭉치 구축에 많은 제약이 따른다. 따라서 한영 병렬 말뭉치의 구축에 필요한 저작권 문제가 해결되어야 다양한 장르의 텍스트, 그리고 최근의 언어 사용 모습을 반영한 텍스트가 활용될 수 있다. 이것이 선결되어야 말뭉치의 균형성과 대표성의 문제를 해결할 수 있고 이는 곧 말뭉치의 언어 연구 활용도를 높이는 것으로 이어질 수 있을 것이다.

마지막으로 범용성과 특수성의 문제가 있다. 말뭉치의 구축은 범용성을 목적으로 하기 쉽다. 그러나 범용성을 염두에 두고 구축된 말뭉치는 여러 연구에 활용되는 데에 제한점이 많이 따른다. 즉 어느 연구에도 적합하지 않거나 추가로 많은 작업을 필요로 할 수 있다. 이번 연구에서도 형태 분석이 이루어진 말뭉치임에도 불구하고 연구 목적에 부합하지 않아 원시 말뭉치에서 수작업으로 분석 대상을 찾는 작업부터 시작되었다. 이러한 문제를 해결하기 위한 방안으로 범용 말뭉치 외에 특수 목적 말뭉치를 구축할 필요가 있다. 의미 주석이 이루어진 말뭉치, 화용 정보 주석이 이루어진 말뭉치 등 직접적으로 관련 언어 정보를 추출할 수 있는 말뭉치가 구축된다면, 특정 목적에 맞는 언어 자료를 선택할 수 있을 뿐 아니라 목적에 맞는 주석 작업으로 말뭉치 구축이 곧 즉각적인 활용으로 이어질 수 있을 것이다. 이러한 특수 목적 말뭉치는 범용 말뭉치보다 작은 규모로 구축될 것이나 더 정확한 언어 정보를 제공하고 말뭉치의 활용도를 높인다는 점에서 그 구축의 정당성을 확인할 수 있다.

<참고문헌>

- 고석주. 2002. 조사 '가'의 의미. 국어학 40, 221-247. 국어학회.
- 류구상. 1986. 주격조사에 대하여. 柳穆相(편) 저 國語學新研究 : 若泉 金敏洙 教授 華甲 紀念에서. 塔出版社, 서울, 333-347쪽.
- 목정수. 2003. 한국어 조사의 분류체계와 유형론. 월인.

- 박혜숙. 2001. 문법의식향상을 통한 영어관사 습득. *영어교육* 56.2, 383-402. 한국영어교육학회.
- 신창순. 1975. 국어의 「주어문제」 연구. *문법연구* 2, 131-170. 문법연구회.
- 유동석. 1984. 양태 조사의 통보기능에 대한 연구. 석사학위 논문, 서울대학교.
- 이인영. 1996. ‘주제’와 ‘화제’—기존 주제 개념에 대한 재고—. *어학연구* 32.1, 61-69. 서울대학교 어학연구소.
- 이춘숙. 1999. 토씨 {가}와 {는}의 의미기능—자리개념과 영역개념에서 본—. *한글* 243, 177-209. 한글학회.
- 이필영. 1982. 조사 ‘이/가’의 의미분석. *관악어문연구* 7, 417-431. 서울대국문과.
- 장유진. 1999. 국어 격조사의 기능 연구—의미 기능을 중심으로—. 석사학위 논문, 성신여자대학교.
- 전영철. 2006. 대조 화제와 대조 초점의 표지 ‘는’. *한글* 274, 171-200. 한글학회.
- 채완. 1977. 조사 ‘-는’의 의미. *국어학* 4, 93-113. 국어학회.
- 최수영. 1984. 주제화와 주격조사: ‘-는’과 ‘-가’를 중심으로. *어학연구* 20.3, 233-250. 서울대학교 어학연구소.
- 최영환. 1987. {가}와 {는}의 의미 비교. 석사학위 논문, 서울대학교.
- 한학성. 1995. 영어 관사의 문법: 국내 관사 교육의 문제점 및 그 개선을 위한 관사 교육 기본 모형의 모색. *언어연구* 13, 109-146. 경희대학교 언어연구소.

접수 일자: 2007년 3월 19일

게재 결정: 2007년 5월 19일