
신어의 정착과 낱말 용법의 변화

이기황 · 연세대학교 BK21 한국 언어문학문화
국제인력양성 사업단 박사 후 연구원

1. 머리말

우리는 종종 언어는 시간의 흐름에 따라 끊임없이 변하고 있다는 사실을 잊고 지낸다. 우리가 잘 의식하지 못해도 언어의 변화는 언어의 여러 국면에서 진행된다. 우선은 말소리가 변한다. 최근에 와서 많은 사람들(주로 젊은이들)이 된소리를 많이 사용하는 것, 또 복모음을 단순화하여 단모음으로 발음하는 것은 말소리 변화의 한 예이다. 이러한 발음의 변화는 대체로 지엽적이고 단편적으로 일어나지만 결국에는 우리말의 음운 체계를 변화시킬 가능성도 있다. 언어 변화에 있어서 어휘의 변화도 예외가 아니다. 어휘의 변화는 대규모로, 그리고 매우 빠른 속도로 이루어진다. 새로운 지식과 정보가 감당하기 힘들 정도로 창출되는 현대 사회에서 새로운 개념을 표현하기 위한 신어가 대량으로 생성되는 것은 매우 당연한 일이다.¹⁾

1) 물론 신어가 새로운 개념을 표현하기 위해서만 만들어지는 것은 아니다.

새로운 개념, 제도, 사물 등이 만들어지거나 외국에서 유입될 때 다양한 경로와 방법으로 신어들이 만들어져 사용되는데, 이들이 언중들의 지지를 받아 성공적으로 정착하기까지는 의사소통에 있어서 상당한 혼란과 장애가 발생할 수도 있다(남기심, 1983). 의사소통의 혼란과 장애를 최소화하기 위해서는 적절한 신어를 만들어 내야 하며, 만들어진 신어를 꾸준히 정리하고, 신어의 정확한 의미와 용법을 언중들에게 효율적으로 전달해야 할 것이다. 이와 같은 노력의 일환으로 국립국어원에서는 매년 신어 조사 작업을 수행하고 있다.²⁾ 특히, 2005년 신어 조사인 김한샘(2005)에서는 1995년에 조사된 신어들과 2002년에서 2004년까지 조사된 신어들이 2005년에 어떻게 사용되고 있는가를 보이고 있다. 이는 매우 긍정적인 발전으로 국어 어휘의 생성과 소멸에 대한 의미 있는 자료를 제공하고 있다.

이 글에서는 김한샘(2005)에서 제시한 자료를 바탕으로 신어가 정착해 가는 과정을 그 용법 및 함께 어울려 쓰이는 낱말들의 변화를 통해 관찰해 보고자 한다. 이러한 연구는 대규모의 정밀한 작업이 요구되지만 여러 가지 제약으로 인해 이 글에서는 비교적 소규모로 이루어진 관찰 결과만을 보인다. 또한 신어의 정착 과정을 보다 정밀하게 추적하기 위한 몇 가지 방법도 제안한다.

2. 신어의 개념과 용법의 변화

먼저 이 글의 논의를 위해 신어의 개념을 잠시 살펴보자. 남기심(1983)에서는 새말을 “이미 있었거나, 새로 생겨난 개념이나 사물을 표현하기 위해 지어낸 말, 그리고 이미 있던 말이라도 새 뜻이 주어진 것을 통틀어 일컬으며, 다른 언어로부터 사물과 함께 차용되는 외래어도 포함된다”라고 정의하였다. 우리는 보통 신어라 하면 ‘지어낸 말’, 즉 새로 만들어진 형태를 가진 단어를 떠올린다. 그런데 앞의 신어의 정의에서 밝히고 있는

2) 국립국어원의 신어 조사는 1994년에 시작되어 그 이듬해까지 이어지고 잠시 중단되었다가 2000년부터 다시 시작되었다.

것처럼 ‘이미 있던 말’에 새 뜻이 주어지는 경우도 있다. 전혀 새롭게 만들어진 신어를 발견하는 것은 새 뜻이 주어진 신어보다 찾기가 상대적으로 쉽다. 기존 사전의 표제어 목록 등과 비교하는 것을 포함한 기계적인 작업에 의한 신어의 인식이 가능하기 때문이다. 그러나 이미 있던 말에 새 뜻이 주어진 것을 찾아내기 위해서는 모든 단어의 의미 변화를 추적해야 할 것이므로 이런 유형의 신어를 골라내는 일은 소극적이며 수동적으로 이루어질 수밖에 없을 것이다. 모든 경우에 적용되지는 않을 것이지만, 새 뜻의 주어짐을 짐작해 볼 수 있는 실마리로 새로운 용법의 등장을 들 수 있다. 다음의 예를 보자.³⁾

(1) 가. 나는 요즘 이상한 꿈을 자주 꾀다.

나. 김 선생은 학창 시절 꿈이 남달리 큰 학생이었다.

다. 여러분, 큰 꿈을 가지십시오.

(1가)에서처럼 명사 ‘꿈’이 본래의 뜻으로 사용될 때에 주로 함께 쓰이는 동사는 ‘꾀다’이다. 그런데 본래의 ‘꿈’과는 거의 관련 없이 ‘희망’의 뜻으로 쓰일 때에는 (1나)와 (1다)에서와 같이 다른 서술어(‘크다’, ‘가지다’)와 함께 쓰일 수 있다. 여기서 알 수 있는 사실은 낱말의 의미 변화는 종종 용법, 혹은 함께 쓰이는 낱말들의 변화를 불러일으킨다는 것이다. 또한 신어는 새로운 개념이나 사물을 표현하기 위해 만들어지는 말이므로, 그 개념의 사물을 둘러싼 사회 상황이 변화하면 신어와 함께 쓰이는 낱말들도 함께 변화할 것이다.

3. 새말의 정착과 용법의 변화 관찰

앞서 언급한 바와 같이 김한샘(2005)에서는 이전에 조사된 신어가 2005년에 어떻게 사용되고 있는가를 보였다. 조사에 따르면 2002년에 조사된 408개의 신어 가운데 187개의 신어가 2005년도 신어 조사 대상 말

3) 이 예는 남기십(1983)에서 제시된 것을 완전한 문장의 형태로 고친 것이다.

뭉치에 출현하였다.4) 이들 2005년에도 사용되는 것으로 보고된 2002년 신어들 가운데 ‘기러기 아빠’, ‘오노스럽다’, 그리고 신어 조사에 포함되지는 않았으나 최근 주로 구어 표현에서 새 뜻을 갖게 된 낱말인 ‘꽃하다’의 용례를 분석한다.5)

3.1. ‘기러기 아빠’의 용례 분석

2000년에서 2005년 사이의 ‘기러기 아빠’의 사용 빈도 관련 통계는 다음과 같다.

통 계	2000	2001	2002	2003	2004	2005	2006
기 사 수	15	12	20	69	80	123	72
빈 도 수	23	13	60	215	160	318	142
기사당 평균 빈도	1.53	1.08	3.0	3.12	2.0	2.59	1.97

이 표에 따르면 신어 ‘기러기 아빠’의 사용 빈도가 계속적으로 증가하고 있다. 사용 빈도뿐만 아니라 이 낱말이 사용된 기사의 수도 계속적으로 증가하고 있다. 흥미로운 것은 기사당 평균 빈도도 대체로 증가 추세에 있다는 것이다. 특히 2002년에 들어서면서 이 낱말의 사용에 급격한 변화가 있었음을 볼 수 있다. 이제 실제 용례를 살펴보자.

(2) 가. 이미지는 19살에 “열아홉 순정”으로 가요계에 데뷔해 “섬 마을 선생님”, “동백 아가씨”, “기러기 아빠” 등의 히트곡을 남기며 “엘레지의 여왕”으로 군림해왔다.[한국경제 2000. 05. 10]

나. 혼자 남은 가장을 ‘기러기 아빠’라고 부른다[동아일보 2001. 03. 13]

4) 실제로 2005년도 용례를 보인 2002년 신어들은 151개이다.

5) 모든 용례는 한국언론재단의 기사 검색 사이트인 <http://www.kinds.or.kr>에서 수집하였다. 검색 기간은 2000년 1월 1일부터 2006년 11월 25일까지로 설정하였다. 2002년 신어는 2002년 이전에 만들어진 것들도 포함하기 때문이다. 정밀한 용례 여과를 하지 않았으므로 추출된 용례에 잡음이 섞여 있는 경우가 있으나 이 글의 논의와 관련하여서는 염려할 수준은 아닌 것으로 판단된다. 검색에 이용한 파이썬 프로그램을 부록에 제시한다.

다. 소득이 높아지는 것에 걸맞게 교육의 질이 개선되지 않으면 기러기
아빠들은 더욱 늘어날 것이다.[동아일보 2002. 01. 12]

라. 그러면 기러기 아빠들은 그렇게도 품에 안고 싶었던 자식을 잃는
견디기 어려운 고통을 겪는다.[한국일보 2002. 09. 24]

(2가)는 검색된 ‘기러기 아빠’의 최초 용례이다. 그런데 이 용례에 나타
난 ‘기러기 아빠’는 “자녀를 외국에서 공부시키기 위해 아내와 자녀를 외
국에 보내 놓고 홀로 국내에 남아 생활하는 남편을 비유적으로 이르는
말”이라고 2002년 신어 조사(박용찬, 2002)에 보고된 신어가 아니라 노래
의 제목이다.⁶⁾ ‘기러기 아빠’가 앞의 정의의 뜻으로 사용된 첫 번째 예는
(2나)이다. 이 용례의 특징은 신어가 처음으로 사용될 때, 그 신어의 등장
을 나타내 주는 기재로 자주 사용되는 ‘부르다’라는 동사가 쓰인 것이다.
‘부르다’와 같이 신어를 도입할 때에 사용되는 표현으로는 ‘의미한다’,
‘뜻한다’, ‘가리키는 말이다’ 등의 서술어와 ‘일명’, ‘이른바’, ‘소위’ 등의
관형어가 있다(이기황, 2006 참조). 이들 ‘도입어’들이 항상 신어와 사용
되는 것은 아니므로 절대적이지는 않지만, 신어들을 골라내는 보조 도구
로 사용할 수 있다.

‘기러기 아빠’의 사용 빈도가 급격히 증가하는 2002년에 이르면 노래
제목으로서의 쓰임은 찾아보기 힘들어진다. 이때부터는 사회 현상으로서
의 ‘기러기 아빠’에 관한 기사들이 본격적으로 나타나기 시작하고 자연스
럽게 기사당 평균 빈도수도 크게 늘어난다. (2다)는 2002년 ‘기러기 아빠’
용례 가운데 하나이다. 2002년이 용례들을 자세히 살펴보면 사회 현상으
로서의 기술이 주로 이루어진 이유로 ‘늘어나다’, ‘늘다’, ‘흔하다’, ‘신조
어가 등장하다’와 같은 서술어와의 사용이 눈에 띄며, 그러한 상태에 놓
인 사람들의 상황을 설명하는 문맥에서 인성 명사로 쓰이는 예도 많이 발
견된다. (2라)가 그러한 예이다.

이상에서 볼 때, ‘기러기 아빠’는 2001년경에 신조어의 자격을 얻은 다
음 2002년에 본격적으로 정착하기 시작한 것으로 파악된다. 이와 같은 사
실은 ‘기러기 아빠’가 어떤 문맥에서 어떤 낱말들과 어울려 사용되었는가

6) 이 노래의 가사에는 ‘기러기 아빠’라는 표현이 쓰이지 않았다.

를 관찰함으로써 극명히 드러난다.

최근에는 ‘기러기 아빠’ 현상을 사회에서 기정 현실로 받아들이면서 ‘기러기 아빠 스토리’, ‘기러기 아빠 카페’, ‘기러기 아빠 재테크’ 등의 복합 표현이 등장하였다.⁷⁾

3.2. ‘오노스럽다’의 용례 분석

이번에는 최근까지 사용되기는 하지만 ‘기러기 아빠’와는 사용 양상이 사뭇 다른 ‘오노스럽다’의 용례를 살펴보자. 이 낱말의 뜻을 2002년 신어 조사에서는 “염치가 없어 뻔뻔스럽고 마음씨가 좋지 않다. 2002년 솔트레이크 시티 동계 올림픽에서 부정하게 금메달을 딴 미국의 아폴로 안톤 오노 선수를 비꼬는 뜻이 담겨 있다”라고 풀이하고 있다.

먼저 이 낱말의 연도별 사용 빈도수 통계를 보이면 다음 표와 같다.⁸⁾

통 계	2000	2001	2002	2003	2004	2005	2006
기 사 수	0	0	20	6	1	3	0
빈 도 수	0	0	39	6	1	3	0
기사당 평균 빈도	0	0	1.95	1	1	1	0

앞서 살펴본 널리 쓰이게 된 신어인 ‘기러기 아빠’와는 대조적으로 ‘오노스럽다’는 이 낱말이 만들어진 2002년에 집중적으로 사용되었고 그 이후로는 점점 사용 빈도가 줄어들었으며 2006년 기사에서는 전혀 용례가 발견되지 않았다. 이 통계에 따르면 이 낱말은 2002년 생성 당시에만 주로 사용되었고 지금은 거의 사용되지 않는 낱말인 것처럼 보인다. 그러나 인터넷 검색 사이트인 구글(<http://www.google.co.kr>)을 이용하여 검색한 결과 아직도 제법 사용되고 있음을 확인할 수 있었다. 그러나 역시 ‘오노스럽다’는 ‘기러기 아빠’와는 달리 언론 기사보다는 비격식적인 상황에서 주

7) ‘기러기 아빠’와 어울려 쓰이는 낱말들을 조사하는 과정에서 외로이 지나다 홀로 죽음을 맞은 ‘기러기 아빠’들에 관한 최근 기사들 또한 적지 않게 확인되었다.

8) ‘오노스럽다’의 용례 수집을 위해서 검색어로는 ‘오노스럽’, ‘오노스러’, ‘오노스런’을 지정하였다.

로 사용되는 시사성이 강한 낱말로 점차 사용 빈도가 줄고 있음은 부인할 수 없다. 언론 자료에서의 이 낱말의 대표적인 용례를 보이면 다음과 같다.

- (3) 가. 오노스럽다: 정정당당하지 못한 행동을 하는 사람을 묘사하는 속어 (유사어:리자준스럽다. 용례: 정치인들은 늘 오노스러운 짓만 하고 다닌다).[시사저널 2002. 03. 07.]
- 나. 요즘 유행하는 신조어를 빌려 말하자면 참으로 ‘오노스러운’ 행위라고밖에 할 수 없습니다.[동아일보 2002. 03. 15.]
- 다. 그런데 우리 사회에서 ‘오노스런’ 인물들이 가장 많은 곳을 꼽으라면 단연 정치권이요, 그들이 한껏 재주를 발휘할 계절이라면 선거철이 아닐까.[경향신문 2002. 03. 22.]
- 라. 한때 유행한 ‘오노스럽다(치사하다)’는 표현처럼.[동아일보 2005. 10. 18.]

(3가)는 ‘오노스럽다’의 2002년 초기 용례 가운데 하나로 이 신어의 의미와 함께 유사어와 용례까지 제시하고 있다. 이러한 용례는 앞서 언급한 도입어의 사용과는 달리 신어나 유행어의 등장을 명시적으로 보이고 있다. 이보다는 간접적이지만 역시 신어의 등장을 알리는 문맥을 가진 용례는 (3나)의 것이다. 언론 기사는 사회에서 유행하는 신어의 등장을 보고하는 역할을 할 경우가 많아 (3가, 나)와 비슷한 형식으로 신어를 사용하는 경우가 자주 발견된다.

일단 ‘오노스럽다’가 신어로 받아들여진 이후에는 (3다)와 같이 사용되기 시작하였다. 그러나 이 낱말은 특정한 사건과 인물에 깊은 연관을 지닌 풍자성이 강한 표현으로 신문 기사에서 적극적으로 사용되지는 못했고, 사용되더라도 따옴표를 사용하여 특별히 이 낱말을 선택한 저자의 의도를 드러내는 경우가 많았다. 최근 용례에서는 (3라)에서처럼 쓰여서 이 낱말의 유행이 철 지난 일임을 보이고 있다.

요약하면, ‘오노스럽다’는 시사성과 풍자성이 강한 유행어 성격을 지닌 신어로서 뚜렷한 용법상의 특이성을 보이지 않는다. 다만 이 낱말의 유행어적 특성을 반영하는 기재들이 사용된 용례들은 다수 관찰되었다.

3.3. ‘꽃히다’의 용례 분석

『표준국어대사전』은 ‘꽃히다’의 의미를 ‘꽃다’의 피동사로 풀이하고 있는데, ‘꽃다’의 의미는 “1. 쓰러지거나 빠지지 아니하게 박아 세우거나 끼우다. 2. 내던져서 거꾸로 박히게 하다.”로 주어져 있다. 『연세한국어사전』에는 ‘꽃히다’의 의미를 “1. (날카롭거나 뾰족한 것이) 움직이지 않게 박히다. 2. (일정한 곳에) 끼이거나 붙여지다. 3. (핀이나 비너 등이) 빠지지 않게 어디에 찔려 있거나 끼이다. 4. (시선이 어디에) 머무르다.”로 좀 더 자세히 풀이하고 있다. 한편 ‘꽃히다’가 비격식적으로 다음과 같이 쓰이는 것이 자주 관찰된다.

- (4) 가. 이것저것 검색하다가 또 디카에 필이 꽃혀 버렸다.[구글]
나. 저도 언니랑 같은 두 곡에 화~악 꽃혔거든요.[구글]
다. 전 지금 강인 군과 이특 군에게 꽃혀 있삼.[구글]
라. 그 대사가 저는 꽃혔습니다.[구글]

위의 용례에서 보이는 ‘꽃히다’의 의미는 ‘반하다’와 비슷한 것으로 보인다. (4가)는 ‘꽃히다’의 용례라기보다는 굳은 표현인 ‘필이 꽃히다’의 용례이다. 언론정보원 데이터베이스의 1990년 이후 전체 자료에서 검색된 ‘필이 꽃히다’의 최초 용례는 2002년의 다음 용례이다.

- (5) 미국에서 선배의 소개로 알게 된 꼭 감독의 권유로 ‘그냥’ 출연했다가 “필이 꽃혔다.”라고 한다.[서울신문 2002. 05. 08.]

이 용례는 한 배우의 인터뷰 기사에서 나온 것으로 ‘필이 꽃히다’라는 표현은 이 기사가 작성되기 전에 이미 사용되고 있었을 것이다. ‘필이 꽃히다’ 역시 비격식적으로 쓰이는 표현이므로 언론 기사에서 용례를 찾는 것은 쉽지 않은 일이다. 그럼에도 불구하고 이 표현이 쓰이기 시작한 것은 적어도 2000년 이후일 것이라고 조심스럽게 짐작해 볼 수 있다.

‘꽃히다’가 (4가)~(4다)와 같이 새 뜻을 부여받아 신어의 지위를 갖게

된 것은 이 낱말과 어울려 쓰이는 낱말들의 변화로 포착된다. 즉, 『연세한국어사전』의 1~3의 뜻풀이에서 보이는 것과 같이 ‘꽃히다’와 어울려 쓰이는 ‘이/가’ 명사항은 ‘날카롭거나 뾰족한 것’, 더 구체적으로는 ‘핀이나 비녀 등’이다. 이 동사는 두 자리 서술어이므로 ‘에’ 명사항도 필수적으로 요구하는데, 이는 ‘일정한 곳’으로 대표되는 장소를 나타내는 말이 주로 쓰인다. 그런데 같은 사전의 뜻풀이 4에서 보인 것과 같이 뜻이 변하면 어울려 쓰이는 낱말들에 변화가 생긴다. 결국 이 변화가 계속 진행되어 (4나)에서는 ‘이/가’ 명사항은 사람(말하는 사람), ‘에’ 명사항은 추상적 개체(‘꼭’), (4다)에서는 ‘이/가’, ‘에’ 명사항 모두 사람(‘저’, ‘강인 군과 이 특 군’)이 쓰이고 있다. (4라)는 (4다)와 비슷한 쓰임인데 ‘에’ 명사항이 ‘이/가’ 명사항으로 바뀌어 이중 주어문의 형식을 가지는 점이 다르다.

‘꽃히다’가 어떠한 낱말들과 어울려 쓰는지 알아보기 위해 언론연구원 데이터베이스의 2004년 이후 기사에서 ‘꽃히다’의 용례를 수집한 후, 21세기 세종 계획에서 배포하는 지능형 형태소 분석기로 형태소 분석을 하여 ‘꽃히다’와 가장 가깝게 위치한 ‘이/가’ 명사항과 ‘에’ 명사항을 추출하였다.9) 조사 결과의 일부를 표로 보이면 다음과 같다.

순위	‘이/가’ 명사항	‘에’ 명사항	빈도
1	X	그물/NNG+에/JKB	46
2	X	네트/NNG+에/JKB	44
3	X	스탠드/NNG+에/JKB	34
4	X	구석/NNG+에/JKB	32
5	X	콜네트/NNG+에/JKB	26
6	X	상단/NNG+에/JKB	22
7	X	코트/NNG+에/JKB	20
8	X	림/NNG+에/JKB	19
9	X	가슴/NNG+에/JKB	18
10	펼/NNG+이/JKS	X	18

9) 정확히 밝히자면 ‘꽃히다’의 모든 용례를 추출한 것이 아니고 ‘꽃혔’이라는 활용형의 일부 형태를 검색한 것이다. 격조사가 쓰이지 않기도 하고 문장 성분이 생략되기도 하기 때문에 기계적인 방법만으로는 완전한 정보 습득이 불가능하다.

11	X	관중석/NNG+에/JKB	12
12	44/SN+개/NNB+가/JKS	스트라이크/NNG+에/JKB	9
13	X	오른쪽/NNG+에/JKB	9
14	시선/NNG+이/JKS	X	8
15	X	모서리/NNG+에/JKB	6
16	X	왼쪽/NNG+에/JKB	5
17	백/NNG+어택/NNG+이/JKS	코트/NNG+에/JKB	5
18	X	바스켓/NNG+에/JKB	4
19	화살/NNG+이/JKS	과녁/NNG+에/JKB	4
20	X	하단/NNG+에/JKB	3

위의 표에서 볼 수 있는 것과 같이 ‘에’ 명사항으로는 주로 장소나 위치, 방향을 나타내는 구체 명사가 쓰이고 있다. 구체 명사도 신문의 스포츠란에서 주로 볼 수 있는 것들로 대부분 한정되어 있다. 주목할 만한 것은 위에서 살펴본 굳은 표현 ‘필이 꽃히다’의 용례가 상당히 많이 발견되었다는 것이다. 표에 나타나지 않은 ‘에’ 명사항들 가운데 전형적인 장소 및 구체 명사가 아닌 경우에는 ‘뇌리’, ‘가슴’, ‘감성’, ‘칫눈’, ‘노래’, ‘아름다움’ 등이 있다. 그러나 ‘이/가’ 명사구에 인성 명사가 쓰인 예는 이 자료에서 찾을 수 없었다. 이는 자료를 부분적으로 추출한 탓도 있지만, 새 뜻을 지닌 신어로서의 ‘꽃히다’는 앞서도 언급했듯이 주로 비격식적 상황에서 쓰이므로 언론 자료에서는 그 용례를 발견하기 힘들 것으로 판단된다.

4. 새로운 신어 조사 방법의 제안

글쓴이는 이 글에서 신어를 조사, 연구함에 있어서 신어의 목록과 조사 당시의 용례와 빈도, 그리고 과거에 조사된 신어가 조사 당시에도 사용되는가를 점검하여 보고하는 것에서 한 걸음 더 나아가 신어의 사용 양상을 사용 빈도와 용법을 연도별로 추적하는 일을 시도해 보았다. 모든 신어의 사용 양상을 완전하게 파악할 수 있는 방법은 없지만, 대략 다음과 같은

방법을 통하여 신어의 생성과 사멸, 그리고 용법의 변화를 부분적으로나마 살펴볼 수 있을 것으로 기대한다.

4.1. 감시 말뭉치의 구축

‘감시 말뭉치’는 외국의 사전편찬학 용어 ‘monitor corpus’를 우리말로 옮긴 것이다(Sinclair, 1987). 감시 말뭉치란 시간의 흐름에 따라 지속적으로 내용이 추가되는 말뭉치로서 한 언어 체계의 시간적 변화를 추적하는데 이용된다.

이 글에서 이용한 언론연구원의 데이터베이스는 일종의 감시 말뭉치로 간주할 수 있다. 그러나 이 데이터베이스는 언론 자료만을 담고 있기 때문에 신어의 조사에 충분하지 못하다. 신어는 언론 현장에서 생성되기도 하지만, 다른 삶의 현장에서 더 많이 생성될 가능성이 크다. 특히, 기존의 낱말이 새 뜻을 가지게 되는 경우는 더욱 그러할 가능성이 크다. 물론 그러한 신어의등장이 언론을 통해 보도 및 보고되기도 하나 이를 통한 신어 용례의 수집은 매우 제한될 수밖에 없다. 그러므로 감시 말뭉치는 언론 자료 이외에 다양한 정보원을 포함해야 한다. 특히 유용한 것은 비격식적 언어 사용 예를 많이 포함하고 있는 인터넷 커뮤니티 게시판 및 블로그 게시물, 채팅 로그, 전자 우편 등의 자료이다. 이러한 자료들은 언론 자료와 달리 검색 사이트에서 적극적으로 보관되지 않으므로 일정한 주기로 포털 등의 사이트에서 원자료를 수집하여 감시 말뭉치에 포함시켜야 한다.

물론 감시 말뭉치를 물리적으로 하나의 덩어리로, 즉 오프라인으로 존재하도록 구성하는 것도 가능하지만, 아래의 작업들을 잘 설계하면 오프라인과 온라인이 결합된 형태의 감시 말뭉치를 이용하는 것도 가능할 것이다.

4.2. 새로 형성된 신어의 수집

없던 형태의 낱말이 새로 만들어진 경우의 처리이다. 이러한 신어는 자동적인 형태소 분석과 기존 단어의 목록(사전)을 이용하여 기계적으로 골라낼 수 있다. 경우에 따라서는 잘 구축된 어절 사전을 이용하는 것이 훨씬

썬 더 간편하고 유리할 수도 있다. 실제 작업에서는 띄어쓰기, 문장 부호 사용 등의 다양한 문제를 해결해야 한다.

일단 신어 후보 목록이 작성되면 이들의 용례를 추출하여 검증하는 단계를 반드시 거쳐야 한다. 위에서 잠깐 언급한 도입어를 이용하는 것도 보조적으로 이용할 수 있는 방법 가운데 하나이다.

4.3. 새 뜻을 가지는 신어의 수집

이러한 형태의 신어는 수집하기가 매우 어렵다. 완벽하지는 않아도 시도할 수 있는 방법은 이 글에서 부분적으로 보인 것과 같이 일정 낱말들과 어울려 쓰이는 낱말들의 변화를 추적하는 것이다. 모든 낱말에 대해 이와 같은 작업을 하기가 현실적으로 불가능하므로 품사별로 일정 빈도 이상이 되는 낱말들만 대상으로 작업을 해야 할 것이다.

이 작업도 실제 현장에서는 여러 가지 문제와 부딪는다. 가장 먼저 겪게 되는 문제는 자료의 희소성 문제이다. 앞서도 보였지만, 특히 대상 낱말이 용언인 경우 어울려 쓰이는 명사항에 붙는 격조사가 중요한 정보를 제공하는데, 실제 문장에서는 격조사가 쓰이지 않는 경우가 많고, 때로는 명사항이 생략되기도 하므로 자료의 희소성 문제가 발생하는 것이다. 이에 대한 뾰족한 대응책은 없으며 충분한 크기의 말뭉치를 구축하여 이용할 수밖에 없다.

5. 맺음말

이 글에서는 신어가 정착해 가는 과정을 용법 및 함께 어울려 쓰이는 낱말들의 변화를 통해 관찰해 보고자 하였다. 여러 가지 제약으로 소규모의 작업 결과만 제시하였으나, 이 결과만으로도 신어의 정착 과정에서 발생하는 용법의 변화를 추적하는 일의 중요성은 충분히 보였다고 판단된다.

또한 이 글에서 시도한 작업의 확장판인 새로운 신어 조사 방법을 제안하였다. 결코 간단한 일은 아니지만 이를 통하여 보다 정밀한 신어 조사

결과를 실시하는 것이 가능할 것이며, 여러 분야에서 유용히 쓰일 수 있는 자료를 도출할 수 있을 것이다.

| 참고 문헌 |

김한샘(2005), “2005년 신어”, 국립국어원.

남기심(1983), ‘새말[新語]의 생성과 사멸’, 이기문 외, “한국 어문의 체문 제”, 일지사. pp. 192~228.

박용찬(2002), “2002년 신어”, 국립국어연구원.

이기황(2006), ‘말뭉치에 나타난 새말의 문맥 특성’, “2006년 대학언어학회-한국언어학회 가을 공동 학술대회 발표 논문집”, pp. 280~289.

Sinclair, J. M.(1987), *Looking Up*, London and Glasgow: Collins ELT.

| 부록 |

언론연구원 데이터베이스 검색을 위한 파이썬 프로그램

```
#!/usr/bin/env python
# -*- coding: cp949 -*-

import urllib, re, sys

def get_num_page(text):
    """
    Get the number of pages for the search result.
    """

    p1 = text.find('[전체 ')
    p2 = text.find('건 중 (')
```

```

num_res_str = text[p1 + len('[전체 '):p2].replace(', ', '')

try:
    num_res = int(num_res_str)
except:
    sys.stderr.write('Error: %s\n' % num_res_str)
    return -1

num_page = num_res / 25

if num_res % 25 > 0:
    num_page += 1

return num_page

def get_doc_ids(text):
    """
    Get the document IDs from the text.
    """

    ids = re.findall(r'popupViewW(W'(.*)W',W'W',W'')', text)

    return ids

def process_one_page(text, keyword):
    """
    Process one page of the search results.
    """

    base_url = W
    'http://www.kinds.or.kr/main/search/popupcontent.php?docid='

    doc_ids = get_doc_ids(text)

    for doc_id in doc_ids:
        url = base_url + doc_id
        site = urllib.urlopen(url)
        article = site.read()
        output_file_name = keyword + '.' + doc_id + '.html'
        output_file = open(output_file_name, 'w')
        output_file.write(article)
        output_file.close()

    return

def get_web_text(keyword):
    """
    Get the web text for the search terms.
    """

```

```

url1
'http://www.kinds.or.kr/main/search/searchresult.php?keyword='
url2 = ₩
'&range=content&peroid=assign&startdate=20000101&' + ₩
enddate=20061125'
url3 = '&cpgroup=00&sorttype=docsdate_y'

url = url1 + keyword + url2 + url3
site = urllib.urlopen(url)
text = site.read()
site.close()

num_page = get_num_page(text)

if num_page == -1:
    sys.exit(1)

# Process page 1
process_one_page(text, keyword)

if num_page > 1:
    # Process the rest of the result pages.
    base = 'http://www.kinds.or.kr/main/search/searchresult.php'

    for page_num in range(2, num_page + 1):
        # Load the next page.
        print 'page ', page_num
        post_query = [
            ('keyword', keyword),
            ('range', 'content'),
            ('peroid', 'assign'),
            ('startdate', '20000101'),
            ('enddate', '20061125'),
            ('cpgroup', '00'),
            ('sorttype', 'docsdate_y'),
            ('listnum', '25'),
            ('oldquery', '(title:(%s) OR content:(%s))' ₩
            % (keyword, keyword)),
            ('oldkeyword', keyword),
            ('query', '((title:(%s) OR content:(%s)) ' + ₩
            'AND docdate:[20000101-20061125]) AND ' + ₩
            'NOT cpname:(01100901) AND NOT ' + ₩
            'cpname:(03100051)' % (keyword, keyword)),
            ('page', page_num)
        ]
        data = urllib.urlencode(post_query, doseq=True)
        site = urllib.urlopen(base, data)
        text = site.read()
        site.close()

        process_one_page(text, keyword)

```

```
if __name__ == '__main__':  
    if len(sys.argv) < 2:  
        sys.stderr.write('Need a keyword.\n')  
        sys.exit(1)  
    get_web_text(sys.argv[1])
```