

우리글 명사 어절에서의 단어 빈도 효과: 안구운동 추적 연구

고 성 룡^{1)2)*} 홍 효 진¹⁾ 윤 소 정²⁾ 조 병 환¹⁾

¹⁾서울대학교 심리학과

²⁾서울대학교 협동과정 인지과학 전공

본 연구에서는 안구운동을 추적하여 우리글 명사 어절에서의 단어 빈도 효과를 알아보았다. 실험 1에서는 문장 읽기에서 표적 단어의 빈도 효과를 알아보았다. 실험 결과, 고빈도 표적단어의 첫 고정 시간(단어에 처음 위치하여 머무른 시간), 주시 시간(단어를 읽기 시작한 순간부터 단어를 벗어나기 전까지의 고정 시간의 합)이 저빈도 표적 단어의 시간들보다 짧았으며, 단일 고정 시간(단어에 단 한번 고정된 경우의 고정시간)도 유의할 정도로 짧았다. 실험 2에서는 실험 1에서 쓰인 표적단어와 관형격 조사(‘~의’)가 결합된 명사 어절의 읽기 과정에서 단어 빈도 효과가 나타나는지를 알아보았다. 실험 1의 결과와 마찬가지로, 고빈도 단어로 구성된 명사 어절의 첫 고정 시간, 단일 고정 시간, 및 주시 시간이 저빈도 단어로 구성된 명사 어절보다 짧았다. 특히, 첫 고정 시간과 단일 고정 시간에서 명사의 빈도에 따른 차이가 나타난 것은 명사 어절이 명사와 조사로 분리되어 처리된다는 민승기와 이광오(2005, 2007)의 가설을 지지한 것으로 해석되었다.

주제어 : 한국어, 명사 어절, 단어 빈도, 안구 운동

† 교신저자 : 고성룡, 서울대학교 심리학과, (151-742) 서울시 관악구 신림동 산 51
E-mail: koh@snu.ac.kr

글 읽기에서 눈은 물이 흐르듯 유연하게 움직이는 것이 아니라 빠른 도약(saccade)과 움직임이 미미한 고정(fixations)을 번갈아 한다(고성룡, 윤낙영, 2007; 이춘길 2004). 이런 고정과 도약의 연쇄 속에서 독자는 입력되는 단어들을 인식하여 문장의 구조를 형성하고 동시에 앞선 문장들과의 의미 관계를 파악하여 글을 이해한다. 이해 과정에서 가장 기본적인 것이 단어를 인식하는 것이라 할 수 있다. 이 단어 인식 과정에는 단어 빈도, 단어 길이, 주관적 친숙도, 단어 유사성, 맥락 등의 여러 변인이 영향을 미치며(Balota, Yap, & Cortese, 2006 참조), 이 중 단어 빈도가 가장 중요한 변인의 하나라고 알려져 있다. 본 연구에서는 우리글 읽기에서 안구운동 추적 실험으로 거의 확인된 적이 없는 단어 빈도 효과를 확인하며, 더 나아가 어절 처리에서 단어 빈도 효과를 확인하여 최근의 모형(민승기, 이광오, 2005, 2007)을 검증해 보고자 한다.

단어 빈도 효과는 단일 단어에서의 음독(naming)이나 어휘 판단 과제, 또는 문장 읽기에서 안구운동 추적으로 연구되어 왔는데, 고빈도 단어가 저빈도 단어보다 빨리 읽혔고(이광오, 1993; Balota, & Chumbley, 1984), 단어로 빨리 판단되었다(박태진, 2003; 권효원, 김선경, 이해원, 2006; Forster, & Chambers, 1973). 또한, 저빈도 단어들에서 발음이 불규칙한 단어들은 규칙적인 단어들에 비해 늦게 읽혔으나 고빈도 단어들에서는 차이가 없었다(이광오, 1996; Seidenberg, Waters, Barnes, & Tanenhaus, 1984).

자연스러운 글 읽기 상황에서도 단어 빈도는 눈의 움직임에 영향을 미쳤다(Rayner, 1998 참조). 여러 연구에서 고빈도 단어들이 저빈도

단어들보다 처리가 쉽다고 보고 되었는데, 대체로 눈이 단어에 착지한 이후 그 단어를 벗어나기 전까지 시간의 총합으로 정의되는 주시 시간(gaze duration)과 단어의 첫 고정 시간(first fixation duration)에 의해 측정되었다. 고빈도 단어의 주시 시간이 저빈도 단어의 주시 시간보다 짧았으며(Henderson, & Ferreira, 1990; Inhoff, & Rayner, 1986; Just, & Carpenter, 1980; Kliegl, Nuthmann, & Engbert, 2006; Rayner, Ashby, Pollatsek, & Reichle, 2004; Rayner, & Duffy, 1986; Rayner, Sereno, & Raney, 1996; Vitu, McConkie, Kerr, & O'Regan, 2001), 대체로 고빈도 단어의 첫 고정시간도 저빈도 단어의 첫 고정시간보다 짧았다(Rayner, & Duffy, 1986).

글을 읽는 동안 단어의 빈도는 고정하고 있는 단어뿐 아니라 다음 단어의 정보 처리에도 영향을 미쳤다. 이는 가상의 선을 경계로 설정해서 눈이 선을 넘을 때 자극을 변화시키는 경계선(boundary) 기법으로 주변화 미리보기(parafoveal preview) 효과를 살펴본 실험에서 확인되었다(Rayner, 1998 참조). 구체적으로, 표적 앞 단어가 고빈도 단어인 경우에는 주변화에 표적과 유사한 자극이 제시되었을 때가 상이한 자극이 제시되었을 때보다 표적 단어의 읽기가 빨랐으나, 저빈도 단어인 경우에는 차이가 없었다(Henderson, & Ferreira, 1990). 또한, 단어의 빈도 효과는 고정하고 있는 단어에서 끝나는 것이 아니라 다음 단어로 넘어가기도 한다는 보고도 있었다(Rayner, & Duffy, 1986; Kliegl 등, 2006). Rayner와 Duffy(1986)는 표적 단어가 저빈도 단어일 때 표적 다음 단어의 읽기 시간이 길어지는 일출(spillover, 溢出) 효과

과를 보고하였다.

빈도 효과가 글 읽기 방향으로만 영향을 미치는 것이 아니라 반대 방향으로도 영향을 미친다는 보고도 있다(Kliegl, Nuthmann, & Engbert, 2006; Pynte, & Kennedy, 2005). 코퍼스를 분석한 Pynte와 Kennedy(2005)는 고정하고 있는 단어의 길이가 짧은 경우에 고정 시간은 다음 단어가 저빈도일 때가 고빈도일 때보다 길다는 결과를 보였다. 이 효과는 안구 운동 모형에서 단어 처리 방식과 연관된 문제이기 때문에 관심을 끌었다. 하지만, 코퍼스 분석에서만 확인되었고 실험에서는 아직 보고 되지 않은 상태이기 때문에 추후 연구가 필요하다고 하겠다.

단어 빈도 효과는 단일 형태소 단어에 집중되어 연구되어 왔으나 근래에는 다형태소 단어 처리 방식에 관심을 가지면서 그에 대한 연구도 행해지고 있다. 합성, 파생, 굴절 등의 방식으로 만들어지는 다형태소 단어의 처리에 대한 가설은 단어가 형태소 단위로 분해 되어 이루어진다는 형태소 분해 가설(Taft, & Forster, 1976), 분해 없이 단어 단위로 처리된다는 전체 목록 가설(Girardo, & Grainger, 2000), 두 표상과 방식이 혼용되어 있다는 혼합 가설(Stemberger, & MacWhinney, 1986; Stanners, Neiser, Heron, & Hall, 1979; Marslen-Wilson, Tyler, Waksler, & Older, 1994) 등으로 나누어 볼 수 있다(조명한 등, 2003 참조). 다형태소 단어의 처리 방식을 연구하는 한 가지 방법은 위에서 살펴본 단어 빈도를 구성 형태소와 다형태소 단어 수준에서 조작하여 음독이나 어휘 판단을 시키거나 글 읽기에서 안구 운동의 양상을 살피는 것이다. 최근에 Niswander-

Klement와 Pollatsek(2006)은 접두사를 가진 복합어의 안구 운동 추적 연구에서 단어 길이에 따라 어근 빈도 효과와 단어 빈도 효과가 달리 나타나는 것을 보였다. 7.5자 이상의 긴 단어에서는 어근 빈도가 강하게 나타났으나 짧은 단어에서는 반대로 단어 전체 빈도 효과가 강하게 나타났다. 이들은 짧은 단어들에서 단어 빈도 효과가 첫 고정 시간에서 나타나고 긴 단어의 어근 빈도 효과가 주시 시간에서 나타난 실험 결과는 전체 단어 직접 경로와 형태소 분해 경로가 병행하는 이중 경로 모형으로 설명된다고 주장했다.

한편, 영어와 달리 우리글에서 띄어쓰기 단위는 어절이다. 예컨대, ‘학교가’라는 어절은 ‘학교’라는 명사와 ‘가’라는 주격 조사로 구성된다. 그래서 우리글 읽기에서 기본적인 과정은 주어진 자소를 모아 쓴 글자들에서 단일 단어나 단어와 조사로 구성된 어절을 파악하는 것이라 할 수 있다. 우리글 어절 처리 연구는 어절이 다형태소로 구성된다는 점에서 위에서 살펴본 영어권의 다형태소 단어 연구에 영향을 받아 구성 형태소 단어의 특성을 조작하여 연구되어 왔다. 최근에, 민승기와 이광오(2005)는 주어진 글자열이 우리말에서 어절이 될 수 있는지를 판단하는 어절 판단 과제를 사용하여 위에서 언급한 영어 복합어 연구에서처럼 명사의 빈도와 어절 전체 빈도의 효과를 살펴보았는데, 참가자 무선 분석에서 명사 빈도와 어절 전체의 빈도의 주 효과를 얻었다. 민승기와 이광오(2005)는 이 결과를 두 단계 모형으로 설명했는데, 첫 번째 단계는 어절이 명사와 조사로 분리되어 어휘집에 접근하는 과정이고 두 번째 단계는 맥락과 어

절의 적절성 판단 단계이다. 이 모형에 의하면, 단어 빈도 효과는 어절이 분리되어 어휘로 접근하는 과정에서 나타난 것이다. 민승기와 이광오(2007)는 두 단계 모형에 첫 단계 이후 어휘집에서 명사와 조사를 확인하는 단계를 첨가한 세 단계 모형으로 구체화했다. 민승기와 이광오의 모형에서 핵심은 어절이 초기에 명사와 조사로 분리되어 처리된다는 것인데, 이 가설이 옳바르다면, 분리 과정을 반영하리라 여겨지는 단어 빈도 효과가 명사 어절의 첫 고정 시간과 단일 고정 시간에 나타나리라 예측된다.

본 연구의 실험 1에서 먼저 자연스런 문장 읽기에서 안구운동 추적으로 단어 빈도 효과를 살펴보았다. 실험 2에서는 민승기와 이광오(2005, 2007) 모형의 예측을 검증하기 위해 실험 1에서 사용된 명사에 관형격 조사 ‘의’를 첨가하여 구성된 명사 어절에서 단어 빈도 효과가 나타나는지를 확인하고자 했다.

실험 1

안구 운동 추적 연구에서 단어 빈도 효과는 여러 영어권 연구에서 보고 되었고, 우리글 연구에서도 고성룡과 윤낙영(2007)이 대학생들의 문장읽기 자료를 사후 분석하여 고빈도 단어들의 주시시간이 저빈도 단어의 주시 시간보다 짧다는 결과를 보고했다. 실험 1에서는 통제된 안구 운동 추적 실험을 통해 우리글 읽기에서 단어 빈도 효과를 확인하고자 했다. 구체적으로, 동일한 문장 틀에서 고빈도 명사 단어가 표적일 때와 저빈도 명사가 표적일 때에 표적 어절의 읽기 시간과 평균 고정 횟수

를 비교하여 단어 빈도 효과를 확인하고자 했다.

방 법

참가자 서울대학교 재학 중이며 심리학 개론을 수강하는 16명의 학생이 참가했다. 참가자들은 눈에 특별한 이상이 없었으며 시력 혹은 교정시력이 정상이었다.

도구 안구 운동은 EyeLink II(<http://www.eyelinkinfo.com>)로 추적되었다. EyeLink II는 동공을 추적하는 비디오기반 안구 운동 추적 장치로서 시간적 해상도(초당 500번 표집)가 높고 최대 공간 해상도가 또한 0.01도로 좋은 편이었다. 본 실험에서는 눈 밑에서 동공을 찍는 카메라를 이용하여 오른쪽 눈만을 추적하였다. 머리의 움직임은 모니터의 네 모서리에 부착된 적외선 표지에서 입력 받은 값을 바탕으로 보정하였다. 각 자극 문장은 삼성 모니터에서 1024 X 768모드에서 20포인트 바탕 서체로 제시되었으며 한 자의 크기는 시각 약 0.83도였다.

자극 실험에 사용된 문장은 4200만 어절의 연세 말뭉치를 대상으로 한 현대 한국어의 어휘 빈도를 참고하여 36개의 고빈도와 저빈도의 단어 쌍을 찾고 각 쌍에 알맞게 구성되었다(부록1. 참조). 이때 표적단어는 모두 3음절 단어였으며 빈도가 1000 이상(백만당 232 이상)은 고빈도 단어로, 100 미만인 단어는 저빈도로 분류되었다. 고빈도의 평균은 2499(1011-11493) 이었고 저빈도의 평균 26(7-99)이

었다. 문장들은 부록 1에 있는 것처럼 평이한 문장들이었다. 문장의 길이는 평균 7.3(6-9)어절이었으며 한 어절의 길이는 평균 3.1(1-6)자였다. 문장들은 피험자별로 역균등화하여 제시하였다.

6개의 연습 시행을 마치고 잠시 후에 본 연구의 36개 문장을 읽었고 12개의 이해 질문에 답했다. 실험은 총 20분 정도 소요되었다.

결과 및 논의

절차 실험은 개별적으로 시행되었다. 참가자는 지시문을 읽고, 안구 운동 장치에 대해 간략히 소개받았다. 그리고 나서 안구 운동 추적 장치가 머리에 착용되고 눈의 위치가 정위(calibration)¹⁾되었다. 정위 시 참가자는 화면에 제시된 9개의 점에 시선을 고정하도록 지시를 받았으며, 정위는 제시된 점의 위치와 계산된 눈의 위치의 차가 0.5도 이내일 때 유효한 값으로 받아들여졌다. 약 3분 정도의 정위와 확인 절차 후에 연습 시행이 시작되었다. 매 시행 초에는 화면 중앙에 한 점이 제시되어 눈 미끄러짐이 측정되었고 실험자는 참가자가 읽을 준비가 되면 시작 버튼을 눌렀다. 그러면 화면에 문장이 제시되고 읽기가 시작되었다. 참여자가 문장을 다 읽어 버튼을 누르면 한 시행이 끝나고 다음 시행이 시작되었다. 매 시행마다 눈의 위치가 정확하지 않다고 판단되면 정위가 다시 행해졌다. 참여자는 이렇게

실험 스크립트 작성 오류로 두 문장은 자료 처리에서 제외되었으며, 또한 고정 시간이 50ms 미만이거나 1000ms를 넘는 고정들도 제외되었다. 고빈도와 저빈도 조건에서 표적 전 어절, 표적 단어/어절, 표적 다음 어절의 첫 고정 시간, 어절을 보기 시작해서 떠나기 전까지의 고정 시간의 총합인 주시 시간 및 어절/단어에 단 한번 고정된 경우의 고정시간인 단일 고정 시간의 평균이 표 1에 제시되어 있다.

표적 전 단어 분석 표적 전 단어에서는 첫 고정, 주시 시간 및 단일 고정에서 차이가 없었다.

표적 다음 단어 분석 표적 단어 다음 단어의 항목 분석에서는 첫 고정 시간이 저빈도 단어 조건에서 고빈도 단어 조건보다 조금 긴 성향

표 1. 표적 단어 빈도에 따른 표적 전 단어, 표적 단어, 및 표적 다음 단어의 첫 고정, 주시 및 단일 고정 시간의 평균(ms) 및 표준편차 (괄호 안에)

빈도	첫고정시간			주시시간			단일고정시간		
	표적 전	표적	표적 다음	표적 전	표적	표적 다음	표적 전	표적	표적 다음
고빈도	210(42)	220(45)	224(35)	256(74)	247(59)	278(62)	215(44)	223(47)	234(46)
저빈도	214(48)	236(40)	240(55)	248(67)	280(73)	296(99)	225(56)	238(46)	244(59)

1) 일반적으로 보정이라 함.

이 있었다(순서대로 $t_1(15) = 1.535, p = .146, t_2(33) = 1.735, p = .092$). 이는 Duffy와 Rayner (1986)가 영어 연구에서 보인 일출 효과를 우리글에서도 찾은 것이라 할 수 있으나 이는 좀 더 연구가 필요하다.

표적 단어 분석 첫 고정 시간과 주시 시간은 고빈도 단어보다 저빈도 단어에서 길었다(순서대로, $t_1(15) = 2.846, p < .05, t_2(33) = 2.371, p < .05; t_1(15) = 3.654, p < .01; t_2(33) = 3.278, p < .01$). 또한 단일 고정 시간도 고빈도 단어에서보다 저빈도 단어에서 유의할 정도로 길었다($t_1(15) = 1.872, p = .081; t_2(33) = 1.909, p = .065$).

이 결과는 영어권의 안구운동 연구에서 누누이 확인된 단어 빈도 효과를 우리글에서 확인한 것이며, 우리글에서도 고성룡과 윤낙영 (2007)이 사후 분석에서 찾은 것을 통제된 실험으로 확인한 것이다.

또한, 표적에 착지하고 나서 표적을 벗어나기 전의 고정 횟수도 저빈도 표적 단어(1.14번)가 고빈도 표적 단어(1.02번)보다 많았다 ($t_1(15) = 3.450, p < .01, t_2(33) = 2.212, p < .05$). 이 결과는 저빈도 단어를 읽을 때가 동일한 길이의 고빈도 단어를 읽을 때보다 재고정이 많았다는 것을 보여주는 것이다.

요약하면, 실험 1에서 표적 단어의 빈도가 표적 단어의 읽기 시간, 평균 고정 횟수에 영향을 미친다는 것이 확인되었다.

실 험 2

실험 2에서는 실험 1에서 사용된 명사에 조

사를 첨가하여 새롭게 표적 어절을 구성하고, 민승기와 이광오(2005)가 명사 어절 연구에서 확인한 어절 내 명사 빈도 효과를 자연스러운 글 읽기 상황에서 알아보고자 했다. 다른 요소의 개입을 최대한 통제하기 위해 실험 2에 사용된 조사는 관형격 조사 ‘의’로 한정하였다.

방 법

참가자 서울대학교 재학 중이며 심리학 개론을 수강하는 14명의 학생이 참가했다. 참가자들은 눈에 특별한 이상이 없었으며 교정시력이 정상이었다.

도구 실험 1과 동일

자극 기본적으로 실험 1의 표적 단어에 관형격 조사 ‘의’를 첨가하여 만들었다. 이러한 변경으로 인해 문장의 의미가 부자연스러운 문장들은 표적 다음의 부분을 바꾸었다 (부록 II-1. 참조). 따라서 표적 명사 어절은 4자, 문장 길이는 평균 6.9(5-9) 어절, 한 어절은 평균 3.2(1-6)자였다.

절차 실험 1과 동일

결과 및 논의

실험 1과 동일하게 고정 시간이 50ms 미만이거나 1000ms를 넘는 고정들은 자료 처리에서 제외되었다. 고빈도와 저빈도 조건에서 표적 전 어절, 표적 어절, 표적 다음 어절의 첫

표 2. 저빈도와 고빈도 단어에서 표적 전 어절, 표적 어절, 및 표적 다음 어절의 첫 고정, 주시 및 단일 고정 시간의 평균(ms) 및 표준편차 (괄호 안에)

빈도	첫고정시간			주시시간			단일고정시간		
	표적 전	표적	표적 다음	표적 전	표적	표적 다음	표적 전	표적	표적 다음
고빈도	218(47)	215(52)	236(44)	244(53)	256(64)	286(61)	221(49)	221(54)	234(41)
저빈도	213(42)	240(60)	239(43)	244(59)	316(71)	320(71)	218(43)	253(68)	253(42)

고정 시간, 주시 시간, 및 단일 고정 시간의 평균이 표 2에 제시되어 있다.

표적 전 어절 분석 실험 1과 마찬가지로 표적 전 단어 어절에서는 첫 고정 시간, 주시 시간 및 단일 고정 시간에서 차이가 없었다.

표적 다음 어절 분석 실험 1에서는 첫 고정 시간이 긴 경향이 있었는데 비해, 실험 2에서는 주시 시간이 참가자 분석에서 고빈도 조건에서보다 저빈도 조건에서 길었으며 항목 분석에서는 차이가 없었다($t_1(13) = 4.411, p < .01, t_2(35) = 1.623, p = .114$). 이 결과도 실험 1에서처럼 일출 효과로 해석할 수 있을 것 같다.

표적 어절 분석 결과는 실험 1과 거의 동일했다. 첫 고정 시간, 주시 시간, 및 단일 고정 시간은 고빈도 단어 조건보다 저빈도 단어 조건에서 길었다(순서대로, $t_1(13) = 2.569, p < .05, t_2(35) = 3.079, p < .01; t_1(13) = 4.412, p < .01; t_2(35) = 4.149, p < .001; t_1(13) = 2.777, p < .05, t_2(35) = 2.837, p < .01$). 이 결과는 명사와 관형격 조사로 구성된 명사 어절에서도 어절을 구성하는 단어의 빈도 효과가 나타난다는 것을 보여준다. 특히 4자의 어절

의 첫 고정 시간 및 단일 고정 시간에서 단어 빈도 효과가 나타난 것은 단어 빈도가 초기에 작용하고 있는 것을 시사한다.

또한, 실험 1의 결과처럼, 표적에 착지하고 나서 표적을 벗어나기 전의 고정 횟수도 저빈도 표적 어절(1.40번)이 고빈도 표적 어절(1.17번)보다 많았다($t_1(13) = 3.667, p < .01, t_2(35) = 3.319, p < .01$). 이는 저빈도 단어의 어절에서 고빈도 단어의 어절보다 재고정이 많음을 보여주는 것이다.

본 연구에서는 표적 명사 어절의 어절 빈도를 사전에 통제하지 않았기 때문에 고빈도 단어가 사용된 명사 어절과 저빈도 단어가 사용된 명사 어절의 읽기 시간 및 평균 고정 횟수에서 관찰한 차이가 어절 빈도의 효과일 가능성을 배제하지 못한다. 그래서 어절 빈도가 영향을 미쳤을 가능성을 간접적으로 살펴보기 위해 김홍규와 강범모(2000)를 참조하여 표적 명사 어절의 어절 빈도를 사후 분석하였다(부록II-2. 참조).

김홍규와 강범모(2000)의 자료에 등장하지 않은 어절을 제외하였을 때(저빈도 어절 조건은 30개, 고빈도 어절 조건은 4개가 제외되었다), 저빈도 명사 어절의 어절 빈도는 150만 어절 당 평균 1.3(1~3)(100만 어절 당 평균 0.9), 고빈도 명사 어절의 어절 빈도는 150만

어절 당 평균 16.2(1~159)(100만 어절 당 10.8)이었으며, 관형격 조사 ‘의’의 빈도는 150만 어절당 82611(100만 어절 당 55074)이었다.²⁾ 저빈도 단어의 어절 빈도가 고빈도 단어의 어절 빈도보다 낮지만 명사 빈도의 차이와 비교하여 어절 빈도의 차이가 뚜렷하게 크지 않고, 또 두 조건의 어절 빈도가 전반적으로 낮기 때문에 본 실험 결과에서 관찰한 조건 간 차이를 명사 빈도가 아닌 어절 빈도에서 나왔을 것 같지는 않다. 물론 추후의 연구가 필요하다.

요약하면, 실험 2에서 중요한 결과는 명사와 관형격 조사로 구성된 어절에서 단어 빈도 효과가 첫 고정 시간에서 나타난 것이며 이 결과는 초기에 어절이 분리되어 처리된다는 민승기와 이광오의 가설을 지지하는 것이라 할 수 있다.

종합논의

실험 1과 2를 종합해 보면, 실험 1에서 고빈도 단어와 저빈도 단어의 첫고정 시간 평균 차(236ms vs. 220ms, 16ms), 단일 고정 시간 평균 차(238 vs. 223, 15ms), 주시 시간의 평균 차(280ms vs. 247ms, 47ms)와 비교했을 때, 실험 2에서 고빈도 단어와 저빈도 단어의 명사어절

2) 김홍규와 강범모(2000)에 등장하지 않은 어절의 빈도를 0으로 계산하였을 때, 저빈도 명사 어절의 어절 빈도는 150만 어절 당 평균 0.2 (0~3) (100만 어절 당 평균 0.1), 고빈도 명사 어절의 어절 빈도는 150만 어절 당 평균 14.4 (0~159) (100만 어절 당 9.6)이었으며, 관형격 조사 ‘의’의 빈도는 150만 어절당 82611 (100만 어절 당 55074)이었다.

에서의 첫 고정 시간 평균 차(240 ms vs. 215ms, 25ms), 단일 고정 시간 평균 차(253ms vs. 221ms, 32ms), 주시 시간의 평균 차(316ms vs. 256ms, 60ms)가 전혀 줄지 않았다. 이 결과는 단어 빈도 효과가 명사 어절 처리에서 그대로 나타나고 있음을 보인 것이다. 더 나아가 첫 고정 시간과 단일 고정 시간에서 차이가 나타나고 있는 결과는 단어 빈도의 영향이 어절 처리의 초기에 나타나고 있다고 해석할 수 있다. 그래서 이 결과는 민승기와 이광오(2005, 2007)의 명사 어절의 초기 분리 가설을 지지한다.

그러나 민승기와 이광오(2005, 2007)의 명사 어절 처리 가설이 공고히 지지받기 위해서는 명사와 다양한 조사로 구성된 어절에서 초기에 명사와 조사의 분리를 지지하는 결과가 관측되어야 할 것이다. 왜냐하면 다양한 조사로 구성된 어절 처리에서 명사와 조사의 분리 처리를 지지하는 결과는 명사와 조사 각각의 빈도뿐만 아니라 처리 방식에 달려있기 때문이다. 예컨대, 명사와 조사가 순차적으로 처리된다면, 어떤 어절에서도 명사 빈도 효과가 관측될 것이나, 명사와 조사가 병렬적으로 처리된다면, 명사와 조사의 상대적인 빈도 효과가 관측될 것이기 때문이다. 민승기와 이광오(2007)는 이런 점을 인식하고 명사 어절이 명사와 조사로 분리 되어 개별적으로 확인된다는 가설을 제안했다. 이 가설에서도 명사가 처리되고 나서 조사가 처리되는지 아니면 명사와 조사가 병렬적으로 처리되는지 명시적으로 언급되지는 않지만, 아마도 명사와 조사가 병렬적으로 처리된다고 여기는 것 같다. 이는 추후의 연구에 의해 검토되어야 할 것이다.

본 연구의 제한점은 명사 어절 전체의 빈도를 고려하지 못한 점이다. 명사 어절이 명사와 조사라는 두 단어로 구성된다는 우리의 직관에서 보면, 민승기와 이광오(2005, 2007)의 주장대로 명사 어절 빈도 효과가 후기 단계에 나타나는 것이 타당할 수 있다. 다만, 다른 과제 상황에서 제안된 어절 적합성 검증 단계가 자연스러운 글 읽기 상황에서 어떤 형태로 드러날지 예측하기는 쉽지 않다. 따라서 후속 연구에서는 명사 빈도, 조사 빈도, 어절 빈도를 고려하여 문장을 구성하고 각각의 빈도 효과가 여러 가지 안구 운동 측정 지표에서 어떻게 나타나는지를 살펴보는 것이 필요하다.

결론적으로, 본 연구는 위에서 언급한 제한점이 있으나 우리글 명사 어절에서 명사 빈도 효과를 자연스러운 글 읽기 상황에서 안구 운동 추적으로 확인하였고, 또한 이 결과가 민승기와 이광오(2005, 2007)의 어절 처리 모형과 잘 들어맞는다는 것을 보였다는 점에서 그 의의를 찾을 수 있다.

참고문헌

고성룡, 윤낙영 (2007). 우리 문장 읽기에서 안구 운동의 특성: 어절 길이, 단어 빈도 및 착지점 관련 효과. *인지과학*, 18(4), 325-350.

권효원, 김선경, 이혜원 (2006). 한글단어재인에서 단어빈도와 의미점화효과의 관계. *한국심리학회지: 실험*, 18(3), 203-220.

김홍규, 강범모 (2000). 한국어 형태소 및 어휘 사용빈도의 분석. 서울: 고려대학교 민족문화연구원.

민승기, 이광오 (2005). 한국어 명사어절 처리에서의 단어 빈도 효과. *한국실험심리학회 겨울학술대회 발표 논문집*, 78-83.

민승기, 이광오 (2007). 한국어 명사 어절 처리의 세 단계 모형. *한국실험심리학회 겨울학술대회 발표 논문집*, 1-7.

박태진 (2003). 한국어 단어의 주관적 빈도 추정치 및 단어 재인에 미치는 빈도 효과. *한국심리학회지: 실험*, 5(2), 349-366

연세대학교 언어정보개발연구원 (1998). *현대 한국어의 어휘빈도*. 연세대학교.

이광오 (1993). 한글 단어인지과정에서 표기법이 심성어휘집의 구조와 검색에 미치는 영향. *한국심리학회지: 실험 및 인지*, 5, 26-39.

이광오 (1996). 한글 글자열의 음독과 음운규칙. *한국심리학회지: 실험 및 인지*, 8(1), 1-24.

이춘길 (2004). *한글을 읽는 눈의 움직임*. 서울: 서울대학교 출판부.

조명환, 이정모, 김정오, 신현정, 이광오, 도경수, 이양, 이현진, 김영진, 김소영, 고성룡, 정혜선 (2003). *언어심리학*. 서울: 학지사.

Balota, D., & Chumbly, J. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, 10(3), 340-357.

Balota, D., Yap, M. J. & Cortese, M. J. (2006). Visual word recognition: The journey from features to meaning. In Traxler, J. M., &

- Gernsbacher, M.(Eds.) *Handbook of Psycholinguistics*. 285-376. London: ELSEVIER.
- Forster, K., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, 12, 627-635.
- Graudo, H. & Grainger, J. (2000). Effects of prime word frequency and cumulative root frequency in masked morphological priming. *Language and Cognitive Processes*, 15(4/5), 421-444.
- Henderson, J., & Ferreira, F. (1990). Effects foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(3). 417-429.
- Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics*, 40, 431-439.
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87, 329-354.
- Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology: General*, 135(1), 12-35.
- Marslen-Wilson, W., Tyler, L., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review*, 101, 3-33.
- Masson, E. J. M., & Loftus, R. G. (2003). Using confidence interval for graphically based data interpretation. *Canadian Journal of Experimental Psychology*, 57(3), 203-220
- Niswander-Klement, E., & Pollatsek, A. (2006). The effects of root frequency, word frequency, and length on the processing of prefixed English words during reading. *Memory and Cognition*, 34(3), 685-702.
- Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye movement control. *Cognitive Psychology*, 52, 1-56.
- Pynte, J. & Kennedy, A. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research*, 45(2), 153-168.
- Pynte, J., & Kennedy, A. (2006). An influence over eye movements in reading exerted from beyond the level of the word: Evidence from reading English and French. *Vision Research*, 46, 3786-3801.
- Rayner, K. (1998) Eye Movements in Reading and Information Processing: 20 Years of Research. *Psychological Bulletin*, 124(3), 372-422
- Rayner, K., Ashby, J., Pollatsek, A. & Reichle, E. (2004) The Effects of Frequency and Predictability on Eye Fixations in Reading: Implications for the E-Z Reader Model. *Journal of Experimental Psychology: Human Perception and Performance*, 30(4), 720-732
- Rayner, K., Sereno, S. C. & Raney G. E. (1996)

- Eye Movement Control in Reading: A comparison of Two Types of Models. *Journal of Experimental Psychology: Human Perception and Performance*, 22(5), 1188-1200
- Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, 14, 191-201.
- Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior*, 23, 383-404.
- Stanners, R. F., Neiser, J. J., Heron, W. P., & Hall, R. (1979). Memory representation for morphologically related words. *Journal of Verbal Learning and Verbal Behavior*, 18, 399-412.
- Stemberger, J. P. & MacWhinney, B. (1986) Form-oriented inflectional errors in language processing. *Cognitive Psychology*, 18(3), 329-354
- Taft, M. & Forster, K. I. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior*, 15(6), 607-620
- Vitu, F., McConkie, G. W., Kerr, P. & O'Regan, K. (2001). Eye movements in reading isolated words: evidence for strong biases towards the center of the screen. *Vision Research*, 41(25-26), 3513-3533

1 차원고접수 : 2008. 1. 18

최종게재결정 : 2008. 3. 25

The frequency effect in Korean noun eojeols: An eye-tracking study

Sungryong Koh¹⁾ Hyojin Hong¹⁾ Sojeong Yoon²⁾ Pyeongwhan Cho¹⁾

¹⁾Department of Psychology, Seoul National University

²⁾Interdisciplinary Program in Cognitive Science, Seoul National University

This study investigated the frequency effect in the noun eojeols(noun cluster) in Korean using eye-tracking. In Experiment 1, high-frequency nouns were read faster than low-frequency nouns in sentence context. In Experiment 2, the target eojeol consisted of the nouns in Experiment 1 with the genitive case. Similar to Experiment 1, a noun eojeol with a high-frequency noun was read faster than that with a low-frequency noun. We discussed that these results support Min and Lee's hypothesis (2005, 2007), which claims that a noun eojeol in Korean would be initially analyzed into a noun and a bound morpheme like a case marker, in that the first fixation duration and single fixation duration were shorter in an eojeol with a high-frequency noun than in that with a low-frequency noun.

Key words : Korean reading, eojeol(noun cluster), word frequency, eye movement

부록 I³⁾

실험 1의 자극 문장

1. 새로 임명된 사장은 임원진(13)/ 종업원(1011) 월급을 삭감하라고 지시했다.
2. 강도사건을 수사하기 위해 몽타쥬(15)/ 목소리(8248) 분석이 진행되고 있었다.
3. 친구는 나에게 자기 만동서(7)/ 마누라(1825) 나이도 모른다고 대답했다.
4. 오늘 신문의 일면에는 식용유(29)/ 아파트(2994) 가격급등에 관한 뉴스가 실렸습니다.
5. 놀러온 사촌동생이 갑자기 큰스님(65)/ 선생님(11493) 성함을 물어봐서 당황했다.
6. 연구소에서는 최근 들어 황경막(7)/ 어린이(3682) 관련 연구를 하고 있다.
7. 발렌타인 데이에 받은 콩자반(7)/ 라디오(1266) 선물은 매우 달갑지 않았다.
8. 자치단체가 갑작스럽게 요구한 분향소(9)/ 사무실(2759) 이전은 결국 불가능하다고 판단했다.
9. 회의가 언제 열릴지는 아흐레(9)/ 일요일(1439) 저녁이 지나야 알게 된다.
10. 연수에게 선반에 있는 금속기(7)/ 자동차(3488) 광고지를 건네 달라고 말했다.
11. 논문을 읽기 위해서는 대출자(10)/ 전문가(1143) 수준의 지식이 필요하다고 한다.
12. 목욕탕에서 놀던 초등학생이 호흡기(71)/ 영당이(1143) 주위에 큰 화상을 입었다고 한다.
13. 운동장을 달리다가 문득 어깨뼈(15)/ 손바닥(1816) 위가 아파오는 것을 느꼈다.
14. 그녀는 걷는데 지쳐서 대공원(75)/바닷가(1154) 한가운데에 털썩 주저앉아버렸다.
15. 삼억 원의 기부금이 합창대(12)/ 아버님(1470) 이름으로 고아원에 전달되었다.
16. 호텔의 연회장을 빌려서 동문회(9)/ 신문사(1011) 사람들만 출입이 가능하도록 조치했다.
17. 휘황찬란한 빛으로 가득한 갈대숲(12)/골짜기(1349) 너머로 그가 살고 있는 집이 보였다.
18. 수미가 공차기를 하다가 포목점(7)/ 화장실(1078) 유리창을 깨는 실수를 저질렀다.
19. 다른 행사가 있어서 회갑연(8)/ 위원회(3095) 참가는 힘들 것으로 예상된다.
20. 회장의 오랜 희망은 영화제(47)/ 올림픽(1052) 개회식에서 노래를 부르는 것이었다.
21. 우리 주택가의 단점은 물탱크(29)/ 쓰레기(1096) 청소하는 날이 비정기적인 데 있다.
22. 녀석은 자신의 이름이 유럽인(72)/ 계집애(1660) 같다고 말하며 빙긋 웃었다.
23. 조카가 드라마를 보면서 대만인(7)/ 일본인(1714) 흉내를 내는 것에 배꼽 빠지게 웃었다.
24. 어제 보여준 모습은 왈가닥(11)/ 아저씨(5884) 같은 평소의 행동과는 전혀 다른 행동이었다.
25. 모든 사건은 룸메이트의 고자질(53)/ 거짓말(2413) 때문에 시작된 것이었다.
26. 비싸게 팔린 그림은 귀빈실(15)/ 경찰서(1265) 안에 보관 중이라고 한다.
27. 학창시절 나의 별명은 비계살(7)/ 호기심(1198) 덩어리였다는 사실이 떠올랐다.
28. 레스토랑의 선택기준은 맛과 음식값(18)/ 분위기(4150) 등에서 우선순위를 정해서 고른다.
29. 이번 일정에서 오로지 여교사(34)/ 대통령(4928) 위주로만 짜여진 스케줄은 많은 불만을 낳았다.
30. 들이닥친 경찰을 피해 부장실(10)/ 화장실(1078) 바닥에 흩어진 서류들을 재빨리 불태웠다.

3) 괄호 안의 수치는 현대 한국어의 어휘빈도(말뭉치 4200만 어절)를 참조한 단어의 출현 빈도 수
 저빈도 명사 빈도(N=36) 범위 7~99, 평균 25.69, 표준편차 24.694
 고빈도 명사 빈도(N=36) 범위 1011~11493, 평균 2499.0, 표준편차 2233.423

31. 초창기에 주목받은 주제는 기술직(28)/ 노동자(4181) 처우 개선에 관한 것이었다.
32. 평균 90점을 받아오면 종달새(56)/ 강아지(1118) 인형을 선물로 주겠다고 하셨다.
33. 사람을 구하려다 부상당한 여중생(22)/ 공무원(2769) 이야기가 인터넷에서 화제가 되고 있다.
34. 지구 밖을 여행하는 비행선(8)/ 비행기(1591) 제작이 세계 최초로 시도되었다.
35. 운전과 차량 정비, 무전기(99)/ 컴퓨터(1136) 수리까지 군대에서 온갖 것을 배워왔다고 자랑했다.
36. 군인이 직업인 형은 시가전(22)/ 쿠데타(1267) 진압의 경험이 무려 13회나 된다고 말했다.

부록 II-1

실험 2의 자극 문장

1. 새로 임명된 사장은 임원진(13)/ 종업원(1011)의 월급을 삭감하라고 지시했다.
2. 강도사건을 수사하기 위해 몽타쥬(15)/ 목소리(8248)의 분석이 진행되고 있었다.
3. 친구는 나에게 자기 만동서(7)/ 마누라(1825)의 나이도 모른다고 대답했다.
4. 오늘 신문의 일면에는 식용유(29)/ 아파트(2994)의 가격급등에 관한 뉴스가 실렸습니다.
5. 놀러운 사촌동생이 갑자기 큰스님(65)/ 선생님(11493)의 성함을 물어봐서 당황했다.
6. 연구소에서는 최근 들어 횡경막(7)/ 어린이(3682)의 회귀질병을 연구하고 있다.
7. 벨런타인 데이에 받은 콩자반(7)/ 라디오(1266)의 의미가 궁금하다.
8. 자치단체가 갑작스럽게 요구한 분향소(9)/ 사무실(2759)의 이전은 결국 불가능하다.
9. 회의가 언제 열릴지는 아흐레(9)/ 일요일(1439)의 선거후에 결정된다.
10. 연수에게 선반에 있는 금속기(7)/ 자동차(3488)의 광고지를 건네 달라고 말했다.
11. 논문을 읽기 위해서는 대출자(10)/ 전문가(1143)의 도움이 필요하다고 한다.
12. 목욕탕에서 놀던 초등학생이 호흡기(71)/ 영당이(1143)의 주위에 큰 화상을 입었다.
13. 운동장을 달리다가 문득 어깨뼈(15)/ 손바닥(1816)의 윗부분이 아파오는 것을 느꼈다.
14. 그녀는 걷는데 지쳐서 대공원(75)/바닷가(1154)의 한가운데에 털썩 주저앉아버렸다.
15. 삼억 원의 기부금이 합창대(12)/ 아버지(1470)의 이름으로 교아원에 전달되었다.
16. 호텔의 연회장을 빌려서 동문회(9)/ 신문사(1011)의 간부들만 출입이 가능하다.
17. 휘황찬란한 빛으로 가득한 갈대숲(12)/골짜기(1349)의 구석에서 그가 살고 있다.
18. 수미가 공차기를 하다가 포목점(7)/ 화장실(1078)의 유리창을 깨는 실수를 저질렀다.
19. 다른 행사가 있어서 회갑연(8)/ 위원회(3095)의 참가는 힘들 것으로 예상된다.
20. 회정이의 오랜 희망은 영화제(47)/ 올림픽(1052)의 개회식에서 노래를 부르는 것이었다.
21. 우리 주택가의 단점은 물탱크(29)/ 쓰레기(1096)의 관리가 매우 불편한 것이다.
22. 녀석은 자신의 이름이 유럽인(72)/ 계집애(1660)의 이름 같다고 말하며 빙긋 웃었다.
23. 조카가 드라마를 보면서 대만인(7)/ 일본인(1714)의 흉내를 냈다.
24. 어제 보여준 모습은 왈가닥(11)/ 아저씨(5884)의 평소 행동과는 전혀 다른 행동이었다.
25. 모든 사건은 룸메이트의 고자질(53)/ 거짓말(2413)의 반복 때문에 시작된 것이었다.
26. 비싸게 팔린 그림은 귀빈실(15)/ 경찰서(1265)의 안에 보관 중이라고 한다.
27. 학창시절 나의 별명은 비계살(7)/ 호기심(1198)의 다이너마이트였다.
28. 레스토랑의 선택기준은 맛과 음식값(18)/ 분위기(4150)의 조건들에서 순위를 정해서 고른다.
29. 이번 일정에서 오로지 여교사(34)/ 대통령(4928)의 편의 위주로만 짜여진 스케줄이 많다.
30. 들이닥친 경찰을 피해 부장실(10)/ 화장실(1078)의 바닥에 흩어진 서류들을 불태웠다.
31. 초창기에 주목받은 주제는 기술직(28)/ 노동자(4181)의 처우 개선에 관한 것이었다.
32. 평균 90점을 받아오면 종달새(56)/ 강아지(1118) 인형을 선물로 주겠다고 하셨다.
33. 사람을 구하려다 부상당한 여중생(22)/ 공무원(2769)의 이야기가 화제가 되고 있다.

34. 지구 밖을 여행하는 비행선(8)/ 비행기(1591)의 제작이 세계 최초로 시도되었다.
35. 운전과 차량 정비, 무전기(99)/ 컴퓨터(1136)의 수리까지 군대에서 온갖 것을 배워왔다.
36. 군인이 직업인 형은 시가전(22)/ 쿠데타(1267)의 진압 경험이 무려 13회나 된다고 말했다.

부록 II-2⁴⁾

실험 2의 자극 문장의 어절빈도

임원진의(-1)/ 종업원의(6)	회갑연의(-1)/ 위원회의(19)
몽타쥬의(-1)/ 목소리의(6)	영화제의(-1)/ 올림픽의(6)
만동서의(-1)/ 마누라의(3)	물탱크의(-1)/ 쓰레기의(2)
식용유의(-1)/ 아파트의(6)	유럽인의(3)/ 계집애의(6)
큰스님의(-1)/ 선생님의(47)	대만인의(-1)/ 일본인의(8)
황경막의(-1)/ 어린이의(31)	왈가닥의(-1)/ 아저씨의(23)
콩자반의(-1)/ 라디오의(7)	고자질의(-1)/ 거짓말의(-1)
분향소의(-1)/ 사무실의(8)	귀빈실의(-1)/ 경찰서의(1)
아호레의(1)/ 일요일의(-1)	비계살의(-1)/ 호기심의(-1)
금속기의(1)/ 자동차의(29)	음식값의(-1)/ 분위기의(8)
대졸자의(-1)/ 전문가의(4)	여교사의(-1)/ 대통령의(159)
호흡기의(-1)/ 엉덩이의(1)	부장실의(-1)/ 화장실의(1)
어깨뼈의(-1)/ 손바닥의(2)	기술직의(-1)/ 노동자의(37)
대공원의(-1)/ 바닷가의(7)	종달새의(1)/ 강아지의(-1)
합창대의(-1)/ 아버님의(2)	여중생의(-1)/ 공무원의(18)
동문회의(1)/ 신문사의(3)	비행선의(1)/ 비행기의(14)
갈대숲의(-1)/ 골짜기의(2)	무전기의(-1)/ 컴퓨터의(49)
포목점의(-1)/ 화장실의(1)	시가전의(-1)/ 쿠데타의(1)

4) 김홍규와 강범모(2000) 말뭉치(150만 어절)

저빈도 어절 빈도(N=6), 범위 1~3, 평균 1.33, 표준편차 .816, 100만 어절 당 0.89

고빈도 어절 빈도(N=32), 범위 1~159, 평균 16.16, 29.317, 100만 어절 당 10.77

관형격 조사('의') 빈도: 82611, 100만 어절 당 55074

김홍규와 강범모(2000)에 등장하지 않은 어절의 빈도는 괄호 안에 -1로 표시하였다.