

<자료>

한국어 단어의 연상 빈도 및 심상가 조사

박 태 진
전남대학교 심리학과

단어 당 100명의 대학생을 대상으로 665개 한국어 명사 단어의 연상빈도와 심상가를 조사하여 기준을 작성하였다. 연상빈도는 다중 반응 자유연상과 연속적 연상 방법을 사용하여 조사하였고 심상가는 7점 척도상에 평정하도록 하는 방법을 사용하였다. 조사 결과를 바탕으로 심상가와 연상빈도 표를 작성하였다.

주제어: 연상 빈도, 심상가, 한국어 명사 단어

이 논문은 한국학술진흥재단의 지원에 의하여 연구되었음 (KRF-2002-074-HS1002).
본 연구의 자료수집에 도움을 준 김태호군, 김미라와 박정경양에게 감사를 표합니다.
교신저자: 박 태 진, (500-757) 광주시 북구 용봉동, 전남대학교 심리학과
E-mail: tpark@chonnam.ac.kr

인간의 인지기능에 관한 인지심리학 및 인지과학 연구에서 사용되어온 주요 실험자료로서 어휘자료를 들 수 있는데, 영어를 비롯한 여러 언어권마다 어휘 수준에서의 다양한 언어적 및 인지적 특성을 조사 분석하고 이러한 특성이 인지과제 수행에 미치는 효과를 실험적으로 밝혀왔다. 언어 정보처리 및 지식표상의 형성과 활용에 영향을 미치는 것으로 알려진 대표적인 어휘 수준의 특성으로서 사용빈도(frequency), 심상성(imagery), 연상(association) 등을 들 수 있는데, 사용빈도와 심상성은 어휘 자체의 고유한 특성인 반면, 연상은 어휘들간의 관계성에 관한 특성이 다.

연상은 한 단어로부터 다른 단어나 생각을 떠올리는 경향성으로서 단어간에 이미 존재하는 관계성을 나타내는데¹⁾, 이에 관한 경험적 연구는 Francis Galton(1879-1880)과 Hermann Ebbinghaus(1885)에서 비롯되었다(Kling과 Riggs(1971)에서 재인용). 연상 측정방법은 크게 생성 측정법과 관계성 측정법으로 나눌 수 있는데, 전자는 특정 자극에 의해 유발되거나 이 자극에 대한 반응으로서 생성된 반응의 수를 측정하며, 후자는 두 개 자극간의 연합적 관계성을 측정한다. 그밖에 연상 측정방법은 단일 대 다중 반응 연상과 자유 대 통제 연상으로 구분되는데, 전자는 요구하는 연상 반응의 수가 한 개인이 아니면 여러 개인가에 따라 구별되는 것이고 후자는 연상 반응의 유형에 제약을 두지 않는가 아니면 제약을 두는가에, 단어에 대해 유의어 또는 반의어를 연상해내도록 요구하는 경우에 따라 구별되는 것이

다. 다중 반응 연상은 다시 연쇄적 연상(continuous association)과 연속적 연상(continued association)으로 구분되는데, 전자는 원 자극 뿐만 아니라 참가자가 이미 연상해낸 반응에 대해서도 다시 연상해낼 수 있는데 반해, 후자는 원 자극을 반복적으로 제시하고서 원 자극에 대해서만 연상해내도록 한다.

연상의 정도는 측정방법에 따라 연상가나 연상 빈도로 표시된다. 생성 측정법을 처음 도입한 Glaze(1928)는 이 방법을 사용하여 무의미철자의 연상 정도를 측정하였는데, 실험참가자로 하여금 무의미철자에 대해 의미 있는 것이 연상되는지 여부를 판단하도록 요구하였다. 그는 특정 무의미철자에 대해 연상됨을 보고한 참가자의 비율을 연상가(association value: av)로 표시하였다. Noble(1952)은 생성 측정법을 사용하되 2 음절 항목에 대해 연속적 연상을 하도록 요구하고서(60초 동안) 각 항목에 대한 평균 연상 반응 수를 의미충만성(meaningfulness: m)으로 표시하였다. 즉 연상가나 의미충만성은 특정 자극에 의해 유발된 언어적 반응의 단순한 수에 관심을 두고서, 특정 반응을 보고한 사람의 수나 유발된 반응의 수에 따라 자극들을 비교한다. 반면 관계성 측정법은 자극들간의 연합을 기술하는데 사용되는데, 이 방법은 Galton(1879-1880)에서 비롯되었다(Kling과 Riggs(1971)에서 재인용). 관계성 측정법에서는 특정 자극에 대해 연상된 반응들 각각의 연상 빈도(association frequency)를 표시하는데, 이는 연합의 구조적 특성 및 연합적 관계 등을 잘 드러내준다. 흔히 연상 빈도가 높은 반응은 낮은 반응에 비해 연상을 유발하는 자극과 높은 연상 강도를 갖는다고 간주된다. 연상 빈도는 기억과 지식표상 연구에서 흔히 고려되는 변인으로서, 연상 강도의 조작을 통해 재인기억과 회상기억의 차이를 규명하거나 연상 강도의 차이를 고려한

1) 'association'은 '연상' 또는 '연합'으로 번역되는데, 연합은 연상을 포함하는 보다 광의의 개념이라 할 수 있다. 연상은 한 단어로부터 다른 단어 또는 생각을 떠올리는 경향성을 뜻하며, 연합은 사상(事象)이나 사교들간의 관련성 또는 연결성을 뜻한다.

의미기억 모델링이 시도되어 왔다(Kling과 Riggs, 1971).

심상성은 기억 수행에 큰 영향을 미치는 어휘 특성으로서 어휘가 얼마나 명확한 심상을 일으킬 수 있는가 하는 정도를 뜻한다. 심상성의 정도는 심상가로 표현되는데(Paivio, 1971), 이는 통상 실험참가자가 7점 척도상에 평정한 값으로 기술된다. 여기서 심상이 빨리 그리고 명확하게 떠오르면 6 또는 7로, 심상이 떠오르지 않거나 느리게 떠오르면 1 또는 2로 평정될 것이다. 심상성이 높은 어휘일수록 통상 기억 수행이 우수한 것으로 밝혀졌다 (Paivio, 1965). 심상성과 관련된 개념으로서 구체성(concreteness)이 있는데, 이는 단어가 감각경험을 직접적으로 가리키는 정도를 뜻한다. 심상성과 구체성은 일반적으로 높은 상관 경향을 보이지만, 양자가 분리될 수 있을 때에는 심상성이 구체성보다 더 좋은 예언변인으로 알려져 있다 (Paivio, 1965, 1968).

연상 빈도와 심상가는 인지기능에 관한 여러 연구에서 변인의 조작이나 통제 목적으로 흔히 사용되는 기초 자료이다. 앞서 살펴본 바와 같이 영어의 경우 다양하고 방대한 연상 빈도와 심상가 자료가 조사 발표되었지만, 한국어의 경우에는 이러한 기초 자료가 매우 빈약하다. 한국어의 경우 이남표와 이수원(1970)에 의해 아동의 연상 빈도표가 제작 발표된 바 있지만, 그 이후에는 심상가나 연상가 조사 결과가 발표된 바 없다. 따라서 심상가와 연상가와 관련된 연구를 수행하고자 하는 연구자는 관련 자료의 부재로 인해 많은 어려움을 겪어왔다. 본 연구에서는 심상가와 연상가를 조사 측정하여 기준을 작성함으로써 인간 인지기능 연구에 유용한 도구를 제공하는데 그 목적이 있다.

방 법

조사 참가자 참가자들은 전남대학교 심리학개론 수강 학생들로서 특정 학년이나 단과대학, 그리고 학과에 치우치지 않았다. 400명(연령, 만 18세부터 27세; 성별, 남 161명, 여 239명)의 반응 결과가 최종 분석에 사용되었는데, 심상가 조사에 100명, 연상 빈도 조사에 300명의 반응 결과가 포함되었다. 각 단어 당 100명의 반응 결과가 최종 분석에 사용되었는데, 심상가 조사에 참여한 참가자는 심상가 조사 대상 단어 전체에 반응하였고 연상가 조사에 참여한 참가자는 3개 집단으로 나뉘어 각각 연상 빈도 조사 대상 단어 전체의 1/3에 대해 반응하였다.

조사 대상 단어 연세대학교 언어정보개발연구원(1998)의 ‘현대 한국어의 어휘빈도’ 보고서를 바탕으로 1 내지 3개 글자로 구성된 명사 단어 665개를 선정하였다. 선정 기준으로서 (1) 저빈도부터 고빈도에 이르기까지 고루 표집 (2) 일상적으로 사용되는 단어 위주로 선정 (3)구체어에서 추상어에 이르기까지 고루 표집 (4)복합명사는 제외 등을 적용하였다.

설문지와 조사 절차 665개 단어들을 가지고 설문지를 작성하였는데 설문지는 지시문과 평정해야 할 단어들로 구성되었다. 설문지에서 단어들은 가나다 순으로 배치되었는데, 의미상 서로 관련된 단어들(예, ‘교사’와 ‘교수’, ‘군대’와 ‘군악’ 등)은 연달아 또는 근접하여 제시되지 않도록 40개 단어들 이상의 간격을 두고 배치되었다.

연상 빈도 조사 설문지는 각각 상이한 단어들이 실린 3개 설문지들로 구성되었는데 각 설문지는 가나다 순에 따라 구분되었고, 각 연상 빈도 조사 설문지마다 220여개의 단어가 인쇄되었

다. 3개 설문지 각각마다 상이한 100명의 참가자가 반응하였다. 각 단어는 오른쪽에 괄호와 함께 8번 반복되어 인쇄되었으며, 참가자는 괄호 안에 연상되는 단어를 쓰도록 요구받았다. 조사는 30내지 50명 정도의 집단으로 실시되었으며, 각 단어에 대해 15초 동안만 연상어들을 쓰도록 제한을 가했는데, 실험자가 15초마다 마감시간을 알려주고서 다음 단어로 넘어갈 것을 지시하였다.

연상 빈도 조사에 사용된 지시문의 주요 내용은 다음과 같은데, 연쇄적 연상이 아니라 연속적 연상을 할 것을 강조하였다: “단어들로부터 연상되는 단어를 단어 옆의 괄호 안에 써 주시기 바랍니다. 각 단어마다 15초의 시간이 주어지는데, 이 시간 동안 최대한 많은 연상어를 써 주시기 바랍니다. 이때 주의할 사항은, 단어A로부터 단어B를 연상해내고, 다시 자신이 연상해 낸 단어B로부터 단어C를 연상해내는 식으로 해서는 안됩니다. 항상 단어A로부터만 단어를 연상해내시기 바랍니다.”

심상가 조사 설문지에는 66개 단어를 모두 실었으며 100명의 참가자가 이 단어 모두에 대해 반응하도록 요구받았는데, 각 단어의 오른쪽 칸에 1점(심상가가 매우 낮음)부터 7점(심상가가 매우 높음)까지의 7개 숫자 가운데 하나를 쓰도록 하였다. 심상가 설문지는 엑셀 파일로 작성되었으며 참가자들은 이 파일을 웹 상에서 다운로드 받아 응답한 뒤 다시 메일로 반송하였다. 참가자들은 파일을 다운로드 받기 전에 심상가 평정 지시를 집단적으로 받았다. 심상가 평정 지시에서 심상형성 속도를 강조하였는데 이는 Paivio 등(1968)과 Flieller와 Tournois(1994)에 따른 것이다 특히 후자의 연구에 따르면 사람들은 심상가를 평정할 때 심상의 명료성이나 정교성보다 심상형성 속도를 압도적으로 선호하는데, 심상형성 속도에 따른 평정 결과는 심상의 명료성이나 정

교성에 따른 평정 결과와 동일하다.

심상가 조사에 사용된 지시문의 주요 내용은 다음과 같다: “우리가 사용하는 단어들은 이미지를 형성할 수 있는 정도에 있어 서로 차이가 있습니다. 여러분이 할 일은 각 단어에 대해 정신적인 이미지를 얼마나 쉽게 형성할 수 있는지를 평가하는 일입니다. 예를 들어 '나비' 단어에 대해서는 이미지를 즉각적으로 형성할 수 있고 그 이미지가 명료합니다. 반면 '평화' 단어에 대해서는 이미지 형성이 더디게 이루어지며 그 이미지가 모호합니다. 각 단어의 심상가를 평가할 때 주로 얼마나 빨리 이미지를 형성할 수 있는지에 따라 평가해주시요. 심상가가 매우 낮다는 "1", 상당히 낮다는 "2", 약간 낮다는 "3", 중간이라는 "4", 약간 높다는 "5", 상당히 높다는 "6", 매우 높다는 "7"로 평가해 주시기 바랍니다. 앞서의 예에서 '나비'는 7에, '평화'는 1에 가까울 것입니다. 각 단어의 심상가를 평가하여 해당 숫자를 단어 오른쪽 칸에 써주시기 바랍니다.”

자료 처리 및 기준 작성

심상가: 7점 척도상에서 이루어진 참가자들의 응답 결과를 가지고서 단어마다 평균치를 구한 결과를 부록에서 심상가 점수란에 표기하였다.

연상 빈도: 단어별로 각 연상어에 대한 응답자 수(100명당) 즉 연상빈도를 세어서 부록에 기술하였는데, 연상 빈도가 10 이상인 연상어만 표시하였다. 다음 경우들은 동일한 연상어로 간주하였다: 유의어(예, '레스토랑'과 '식당', '노랑'과 '노란색', '칼'과 '검', '해산물'과 '해물'과 '수산물' 등), 품사는 다르지만 의미가 유사하거나 동일한 단어(예, '제거'와 '없애다', '기쁘다'와 '기쁨', '끈기'와 '찰진', '흐름'과 '흐르다'와 흐르는; '시다'와 '시큼'과 '세콤', '푸르른'과 '푸른'과 '푸르름' 등), 맥락상 의미가 동일한 단어(예, '직

업'에 대해 '선생님'과 '교사' 등), 제시된 단어를 포함한 연상어와 포함하지 않은 연상어가 동일한 경우('깎이'에 대해 '깎이머리'와 '머리', '그릇'에 대해 '밥'과 '밥그릇' 등), 단어는 다르지만 동일한 사물을 지칭한 경우('전봉준'과 '독두장군' 등), 동일한 어간을 공유하는 연상어('고도에 대해 '비행'과 '비행기', '강'에 대해 '고기'와 '말고기', '국회'에 대해 '싸움'과 '말싸움' 등).

결과 및 논의

심상가와 연상빈도에 관해 기본적인 몇 가지 기술통계치를 구함으로써 그 특성을 파악하고자 하였다. 심상가의 평균은 4.86(7점 만점), 표준편차는 1.31, 범위는 1.76부터 6.68이었다. 심상가와 연상빈도간의 관계를 알아보기 위해 Pearson 적률상관계수(r)를 구한 결과 심상가와 전체 연상빈도(제시어에 대해 연상빈도 10 이상인 연상어들의 연상빈도의 합) 간에는 $r=0.41(p<.001)$, 심상가와 연상단어 수(제시어에 대해 연상빈도 10 이상인 연상어들의 수) 간에는 $r=0.35(p<.001)$ 의 상관관이 있었는데, 이는 심상가와 연상빈도간에 상당 정도의 상관관계가 존재한다는 것을 시사한다. 또한 전체 연상빈도와 연상단어 수간의 상관관은 매우 컸다($r=0.75, p<.001$). 한편 연상어의 평균 연상빈도는 제1연상어부터 제10연상어까지 순서대로 46.2, 30.6, 23.1, 18.3, 15.4, 13.8, 12.5, 11.5, 11.0, 10.5 이었는데 이는 본 연구에서 사용된 제시어들 전반에 대해 제1연상어의 경우 약 50%에 가까운 참가자들이 동일한 단어를 연상해 냈음을 의미한다.

연상 빈도 결과에서 드러난 특성 가운데 어휘 자체의 의미론적 특성과 무관한 특성을 기술하면 다음과 같다. 첫째, 조사 대상자(대학생)의 특성과 관련된 연상어를 들 수 있다(예, '게임'에

대해 '컴퓨터', '스타크레프트', '오락실', '보드카페', '피시방'; '신문'에 대해 '한겨레'; '깎이'에 대해 '깎이머리'(중고등학생에서 유행하는 헤어스타일) 등). 둘째, 시대적 특성을 반영하는 연상어를 들 수 있다(예, '사진'에 대해 '디카'; '전화'에 대해 '휴대폰(핸드폰)'; '전쟁'에 대해 '이라크'와 '미국'; '대통령'에 대해 '노무현' 등). 셋째, 대중매체의 영향을 반영하는 연상어를 들 수 있다(예, '비'나 '거미'에 대해 '가수(비)와 '거미'라는 예명을 가진 가수가 있음), '나라'에 대해 '강나라'(가수 이름); '살인'에 대해 '살인의 추억'(영화 제목); '도박'에 대해 '올인'(TV드라마 제목); '씨름'에 대해 '강호동'(연예인으로 전업한 씨름 선수); 옥상에 대해 '옥탑방'; 언어에 대해 '언어아가씨'(TV 드라마 제목); '연어'에 대해 '노래'(대학생에게 인기있는 노래에서 '흐르는 강물을 거꾸로 거슬러 오르는 연어들..'이라는 가사 부분이 있음); '반지'에 대해 '반지의 제왕'(영화 제목); '보라'에 대해 '보라들이'(텔레토비)라는 TV드라마의 인물) 등). 넷째, 응답자의 지역적 특성을 반영하는 연상어를 들 수 있다(예, '소주'에 대해 '잎새주'; '민주'에 대해 '5.18'; '대통령'에 대해 '김대중' 등). 다섯째, 단어를 오인한 연상어를 들 수 있다(예, '도교'에 대해 '일본(도교)를 '도교'로 오인한 듯 함) 등). 이러한 특성들은 응답자의 특성이나 시대문화적 배경에 따라 달라질 가능성이 크므로 실험 자료로 이용하는 데 유의해야 할 것이다.

참고문헌

연세대학교 언어정보개발연구원 (1998). 현대 한국어의 어휘빈도. 연세대학교 언어정보개발연구원 보고서 CLID-WP-98-02-28.

- 이남표, 이수원 (1970). 아동의 연상 빈도표 제작. 한국 심리학회지, 1, 149-151.
- Deese, J. (1959). Influence of inter-item associative strength upon immediate free recall. *Psychological Reports*, 5, 305-312.
- Flieller A., & Tournois, J. (1994). Imagery value, subjective and objective frequency, date of entry into the language, and degree of polysemy in a sample of 998 French words. *International Journal of Psychology*, 29, 471-509.
- Glaze, J. A. (1928). The association value of nonsense syllables. *Journal of Genetic Psychology*, 35, 255-267.
- Kling, J. W., & Riggs, L. A. (1971). *Woodworth & Schlosberg's Experimental psychology* (3rd. Ed.). NY: Holt, Rinehart and Winston, Inc.
- Noble, C. E. (1952). The role of stimulus meaning (m) in serial verbal learning. *Journal of Experimental Psychology*, 43, 437-446.
- Paivio, A. (1965). Abstractness, imagery, and meaningfulness in paired-associate learning. *Journal of Verbal learning and Verbal Behavior*, 4, 32-38.
- Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart & Winston.
- Paivio, A., Yuille, J. C. & Madigan, S. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology Monograph Supplement*, 76, 1-25.

1 차원고집수 : 2004. 5. 19
최종게재결정 : 2004. 6. 24



Investigation of Association Frequency and Imagery Value of Korean Words

Taejin Park

Department of Psychology, Chonnam National University

Association frequency and imagery value of 665 Korean noun words were investigated with 100 university students per word. Association frequencies were obtained by multiple-response free association and continued association, and imagery values were obtained by rating on 7-point scale. Tables of word frequency and imagery value were made.

Keywords: association frequency, imagery value, Korean noun word

부 록

한국어 명사 단어 665개의 심상가 및 연상 빈도 표
(심상가 범위는 1점부터 7점; 연상 빈도는 100명당 연상 빈도임)

KCS I

KCS I

KCS I

КСІ

KCS I

КСІ

K S I

КСІ

KCS I

КСІ

KCS I

KCS I

KCS I

K C I

K S I

K C I

KCS I

KCS I