「언어연구」, 제25권 제3호(2008), 1-21

한국어 색채 용어의 색채와의 관계 및 계층 구조*

노창화 김동성 채희락
(한국외국어대학교)

Roh, Chang-Hwa. Kim, Dong-Sung and Chae, Hee-Rahk. 2008. Relations between Korean Color Terms and Colors and a Classification of Color Terms. *Linguistic Research* 25(3), 1-21. The structures of color terms in current ontologies are more flat than hierarchical. However, we argue that color terms should be structured hierarchically. We have conducted psycholinguistic experiments on the mapping between Korean color terms and colors. The physically continuous property of colors makes it impossible for us to refer to "every color." This property exerted a negative influence in proving a reasonable account of colors. We found out that some color terms are closer than others and hence form a group. Such grouping induces a hierarchical clustering of color terms. **(Hankuk Univ. of Foreign Studies)**

Keywords Korean Color Terms, Color Ontologies, Kappa Statistics, Hierarchical Cluster Analysis, Psycholinguistic Experiments

1. 도입

본 연구에서는 한국어 색채 온톨로지 구축의 토대를 위해서 한국어 색채 용어와 색채의 관계를 심리언어학적 실험을 통해서 분석하였다. 기존의 어휘망이나 온톨로지인 SIMPLE Ontology (Lenci 2000, 2001), WordNet (Miller 1998), Art and Architecture Thesaurus에서 색채 영역을 살펴 보면, 여러 색채 용어들이 계층적 구조 없이 동등한 자매 관계를 갖고 있다. 그러나 색채 용어와 색채간의 심리적 연관 관계를 실험한 Berlin and Kay (1969), Berlin and Berlin (1975), MacLaury (2001), Payne (2006)에서는 색채 용어가 단순히 자매 관계로 맺어진 구조가 아니라는 것을 보여 준다. 따라서 본 연구에서는 심리언어학적 실험을 통해서 색채 용어와 색채간의 연관 관계가 어떠한지를 실험하고 이를 통해서 색채 용어간의 관계를 규명하고자 한다.

이 연구에서 진행한 실험의 절차는 다음과 같다. 연세 한국어 사전과 1,000만 어절 세종코퍼스를 통해서 기본 색채 어휘를 추출하고, 이를 활용하여 색채 용어와 색채간의 연관성에

* 여러 가지 건설적인 논평을 해 주신 익명의 심사위원들에게 감사드린다. 논평 내용 중 논문에 반영되지 못한 부분은 차후 연구 과제로 남겨 두려고 한다. 저자들은 한국외국어대학교 언어인지과학과에서 수행하고 있는 2단계 BK21 사업에 참여하고 있다. 채희락은 2008년도 한국외국어대학교 교내 학술연구비의 지원을 받아 본 논문을 완성하였다.

대한 심리언어학적 실험을 하였다. 실험 결과의 통계적 분석을 위해서 카파 통계(Kappa Statistics)와 계층적 군집 분석(Hierarchical Cluster Analysis)을 활용하였다. 카파 통계는 집단 내의 일치도를 검증하는 기법인데, 분석 결과 색채 용어와 색채간에 신빙성 있는 연관성이 있음이 밝혀졌다. 또한 색채 용어가 나타내는 색채들간의 유사성 정도에 차이가 있다면 색채 용어가 몇 개의 군집(cluster)으로 나눠질 것이므로 계층적 군집 분석을 통해서 이를 살펴 보았다.

색채는 물리적으로 연속적인 속성을 가지고 있기 때문에 "모든" 색채가 관련 용어를 가질 수는 없다. 따라서 색채 용어와 색채와의 관계를 살펴 보기 위해서는 각 색채 용어가 나타내는 전형적인 색채를 중심으로 관련 색채간의 가깝고 먼 관계를 구별할 필요가 있다. 이와 같이 선택된 전형적인 색채의 가깝고 먼 관계로 색채 용어의 계층 구조를 나타낼 수 있게 된다. 기존의 어휘망이나 온톨로지에서 자매 관계만을 가정한 것은 색채 용어들의 색채 속성에 의해 생긴 가깝고 먼 관계를 수렴하지 않은 것이다. 색채들의 속성을 수렴한다면 색채 용어들간의 가깝고 먼 관계가 계층적으로 나타나게 된다.

2절에서는 기존 온톨로지들의 색채 영역을 살펴 보고 색채 용어 관련 심리언어학적 기존 연구물도 살펴 본다. 그리고 3절에서는 색채 어휘 추출 결과와 본 연구에서 실시한 실험 및 실험 절차를 소개한다. 4절에서는 실험 결과를 소개하고 결과에 대한 통계 분석을 제시한다.

2. 선행 연구

기존 온톨로지나 어휘망의 색채 영역을 살펴 보면, 모든 색채 용어가 '색채(color)' 라는 상위어 아래에 색채 용어들이 서로 자매 관계를 맺는 구조를 가지고 있다. 다음의 [그림 1, 2, 3]은 각각 SIMPLE Ontology, WordNet, Art and Architecture Thesaurus의 색채 구조이다.[1)]

1) 한 익명의 심사위원이 지적한 바와 같이, CoreNet, KoreLex, U-Win 등의 기존 한국어 어휘망 / 온톨로지에서 색채 용어가 어떻게 처리되고 있는지도 살펴 보아야 한다. CoreNet에서는 색채 용어들이 자매 관계를 이루고 있으며, KoreLex에는 색채 용어가 포함되어 있지 않다. U-Win에서는 색채 용어가 계층적 구조를 이루고 있으나 몇 가지 문제가 있는 것으로 보인다. 각 용어의 부모가 '색'이 되기도 하고 '빛깔'이 되기도 하며, 녹색과 초록색이 같은 색임에도 불구하고 붉은색, 노란색 등과 동등한 자매 관계를 이루는 구조를 가지고 있다.

[그림 1] SIMPLE Ontology (Lenci 2000, 2001)

[그림 2] WordNet (Miller 1998)

[그림 3] Art and Architecture Thesaurus

위에서 볼 수 있듯이 red, green, blue 등 모든 색채 용어가 자매 관계로만 나타난다. 그렇지만 Berlin and Kay (1969), Berlin and Berlin (1975), MacLaury (2001), Payne (2006) 등의 심리학적인 연구에서는 색채 용어들이 자매 관계의 구조만으로 설명될 수 없는 특성을 가지고 있음을 보여 준다. 이와 같은 속성은 색채 용어간의 유사성 정도에 따른 군집성으로 드러난다.

대표적인 심리학적 연구인 Berlin and Kay (1969)에서는 모국어 화자들에게 각 색채 용어가 나타내는 모든 "색채편(color chips)"을 고르고 그 중에서 가장 전형적인 색채편을 고르게 하는 실험을 했다. 이 실험에서 한 색채편에 여러 색채 용어가 중복되거나 한 색채 용어에 여러 색채 편이 선택되기도 했다. 반면 색채의 물리적 속성에서 거리가 먼 색채편을 나타내는 색채 용어들은 차별성이 컸다. 이는 색채 용어와 색채와의 관계가 단순히 평면적인 관계를 나타내는 것이 아니라 유사한 속성간에 군집되는 관계가 존재한다는 것을 반영하고 있다.

색채간의 유사한 속성에 의한 색채 용어의 군집 관계를 물리적 "색채 스펙트럼(color spectrum)"에서의 색채편의 분포로 살펴 보자. 예를 들어, Berlin and Kay (1969: 124)에서 제시된 한국어 색채 용어 '똥색'과 '노랗다', '갈색'이 나타내는 색채편 중 일부의 색채편은 공통적이다. 반면, 다른 색채 용어 '청색'과는 스펙트럼 내에서 상대적으로 먼 거리에 위치하고 있기 때문에 공유하는 색채편이 없다. 이 실험 결과는 [그림 1, 2, 3]의 온톨로지 / 어휘망에서 제시한 자매 관계로는 색채 용어의 가깝고 먼 관계를 포착할 수 없음을 보여 준다.[2)]

2) 한 익명의 심사자가 Berlin and Kay (1969), Payne (2006) 등의 연구를 좀 더 자세히 소개할 것을 요구했다. 그렇지만 이들은 기본적으로 색채와 관련된 심리 실험을 다루고 있으며 온톨로지와는 직접적인 관계가 없다. 이들은 색채 스펙트럼 내에서의 색채 용어들의 일반적인 분포를 살펴 보고 있다. 본 논문에서 이들을 언급한 이유는 이들이 색채 용어들이 계층적 관계를 가지고 있음을 암묵적으로 보여 주고 있기 때문이다. 그러나 이런 계층적 관계가 직접적으로 언급되지는 않았다.

앞에서 제시한 기존의 온톨로지와 어휘망과 같이 색채 용어들의 관계를 자매 관계라고 가정할 경우, 두 가지 문제점이 발생한다. 첫째, 색채 용어들간의 유사성과 차이점을 포착하기 어렵다. 즉, 각 색채 용어에 해당하는 색채의 속성이 충실히 반영되지 않는다는 것이다. 둘째, 색채 용어들이 일부 색채편을 공유하는 현상을 설명할 수가 없다. 예들 들어, 앞 단락에서 제시한 예에서 '똥색'과 '노랗다', '갈색'이 일부 색채편을 공유하기 때문에 그 만큼 유사하지만 이런 유사성을 설명할 수 없다. 본 연구에서는 이와 같은 기존 온톨로지의 색채 용어 구조의 문제점을 극복하기 위해 심리언어학적 실험을 실시하고 이를 통해서 색채간의 관계성을 살펴보게 될 것이다. 그 결과를 바탕으로 기존의 색채 온톨로지들을 재구성해야 할 필요성이 있음을 보이고자 한다.

3. 실험

3.1 색채 어휘 추출

본 연구에서는 연세 한국어 사전과 1,000만 어절 세종코퍼스에서 빈도 6만 이내의 어휘 중에서 색채 명사를 추출했는데, 선별과 추출의 절차는 다음과 같다. 사전에 등재된 단어 중 파생 접미사 '－색'과 '－빛'이 모두 존재하는 경우 '－색'이 들어가는 단어를 선택했다. 그 이유는 전자가 후자보다 빈도수가 높기 때문이다. 예를 들면, '보라색'과 '보라빛'이 모두 사전에 등재되어 있는데, 이 경우에 '보라색'만을 택했다. 이런 기준으로 연세 한국어 사전에서는 총 58개의 색채 명사를 추출했고 (부록 1 참조), 동일한 방식으로 세종코퍼스에서 총 65개의 색채 명사를 추출했다 (부록 2 참조). 이 두 그룹의 명사 중 겹치는 것은 총 49개이다.[3] 이 중 빈도수가 높은 순으로 살펴 보면 녹색 (221),[4] 흰색 (202), 회색 (171) 순이며, 빈도수가 낮은 색채 어휘들은 쑥색 (5), 진분홍 (5), 진홍 (5), 황토색 (5), 가지색 (4) 등이 있다.[5] 이 중 '국방색'처럼 특정 사물의 색채를 나타내거나 기타 불분명한 색채를 나타내는 단어 8개를 제외하고[6] 총 41개의 단어로 연구 대상 색채 용어 목록을 만들었다.

---

3) 겹치는 49개의 색채 명사는 다음과 같다: 갈색, 감색, 감청색, 검은색, 고동색, 곤색, 국방색, 금색, 노랑, 녹색, 누렁, 다홍, 물색, 미색, 밤색, 백색, 베이지색, 보라색, 분홍색, 붉은색, 빨강색, 선홍색, 연두색, 연분홍, 연초록, 오렌지색, 은색, 자색, 자주색, 적갈색, 주홍색, 주황색, 진보라, 진분홍, 진홍색, 청록색, 청색, 초록색, 파랑 / 파란색, 푸른색, 핑크색, 하늘색, 황갈색, 황금색, 황색, 회갈색, 회색, 흑갈색, 흰색 / 하양 / 하얀색.

4) ( ) 안의 숫자는 출현 빈도를 나타낸다.

5) 김영선/이만영 (1998)에서는 색채 어휘를 채도가 높은 것과 낮은 것으로 나누어 빈도를 분석했다. 채도가 높은 색채 어휘는 다음과 같다: 녹색, 하늘색, 연두색, 파란색 (파랑), 청록색, 노란색 (노랑), 보라색, 초록색, 풀색, 청색, 흰색, 회색, 남색, 황토색, 군청색, 고동색, 푸른색, 바다색, 국방색, 카키색, 검은색 (검정), 겨자색, 갈색, 상아색, 옥색, 기타 (사용 빈도 순서임). 채도가 낮은 색채 어휘는 다음과 같다: 녹색, 하늘색, 연두색, 보라색, 청록색, 노란색 (노랑), 파란색 (파랑), 회색, 남색, 풀색, 초록색, 검은색 (검정), 청색, 황토색, 군청색, 고동색, 카키색, 국방색, 푸른색, 갈색, 살색, 겨자색, 흰색, 옥색, 바다색, 기타 (사용 빈도 순서임).

(1) 갈색, 감색, 감청색, 검은색, 고동색, 곤색, 노랑, 녹색, 다홍, 물색, 미색, 베이지색, 보라색, 분홍색, 붉은색, 빨강색, 선홍색, 연두색, 연분홍, 연초록, 오렌지색, 은색, 자색, 자주색, 적갈색, 주홍색, 주황색, 진보라, 진분홍, 진홍색, 청록색, 청색, 초록색, 파랑 / 파란색, 푸른색, 핑크색, 하늘색, 황갈색, 황금색, 황색, 회갈색, 회색, 흑갈색, 흰색 / 하양 / 하얀색.

본 연구에서는 “한국 표준 색채 규격(KSA 0011)”에서 제시된 색채 용어를 사용하지 않고 사전과 코퍼스를 추출해 활용하였다. 표준 색채 규격에서 제시된 용어가 있음에도 불구하고 이를 사용하지 않은 이유는 다음과 같다. 무엇보다 표준 색채 규격 용어는 색채 전문가를 위한 것이므로 비전문가인 일반인들이 일상적으로 쓰는 색채 어휘와 거리감이 있는 어휘도 함께 섞여 있는 문제가 있다. 또한 색채 용어의 사용 빈도가 제시되어 있지 않기 때문에 그 용어들이 일상적으로 어느 정도로 활용되고 있는지도 알 수 없는 문제점이 있다. 이런 이유들로 표준 색채 규격에 제시된 색채 용어를 상위 온톨로지 구축을 위한 대상 용어로 보기 어렵다고 판단하였다. 그래서 사용 빈도를 고려한 색채 용어를 다시 추출하게 되었다.[7]

3.2 실험 및 실험 절차

실험에 참여한 피험자는 학부 및 대학원생 11명이며, 모두 정상 시력과 정상 색감을 가지고 있음을 실험 전에 검증하였다. 실험에 사용된 색채 용어는 앞 장에서 기술한 연세 한국어 사전과 국립국어원의 빈도 6만 이상 단어의 중복되는 색채 용어 41개이다 (앞 절의 (1) 참조).

“색판(board of colors)”은 “먼셀 색채 모음집(Musell color glossary book)”에서 뽑은 33개의 색채편을 사용하였다. 이 중 30개는 유채색이고 3개는 무채색이다. 유채색 색채편들은 먼셀 공간에서의 기본 10 색상에서 색상별로 3개씩 뽑아서 30개를 만들었다.[8] [표 1]의 빨강 색상을 예로 들면, 명도와 채도별로 나타낼 때 동그라미로 표시된 세 부분이다. 무채색 3개는 명도 1, 5, 10으로 무채색군 중 가장 구별이 뚜렷한 색채편들이다. 선정된 색채편 33개를 담고 있는 색판은 A와 B로 나누었는데, 색판 A는 가로로 색상 1∼10 순서로 배열하고 세로로 각 색상에서 명도와 채도를 고려했을 때 대비가 가장 분명하게 드러나는 1번, 3번, 5번과 무채색의 순서로 제시했다. 색판 B는 세로로 색채군 3번, 1번, 10번, 무채색의 순서로 제시하고 가로로는 기본 색상 10∼1 순서로 배열하여 피험자가 색채의 순서를 학습할 수 있는 요인을 제거했다.

---

6) ‘국방색, 금색, 물색, 은색, 황금색’은 특정 사물의 색채를 나타내고 ‘누렁, 백색, 자색’은 어떤 색을 나타내는지 불분명하여 대상에서 제외했다.

7) 그렇지만 표준 색채 규격 용어와 이만영/김영선 (1997)이 제안한 기본색 이름 후보, 우리나라 문교부 고시 20 색상에 대응하는 색채 어휘, 한국 공업규격에서 선정한 기본 색채 어휘를 포함하고 있는 김영선/이만영 (1998)의 색채 어휘 등을 본 연구에서 추출한 어휘와 빈도를 대조하는 데에 참고하였다.

8) 먼셀 공간에서의 기본 10 색상은 1 (빨강), 2 (주황), 3 (노랑), 4 (연두), 5 (초록), 6 (청록), 7 (파랑), 8 (남색), 9 (보라), 10 (자주)이다.

[표 1] 채도와 명도에 따른 '빨강'색의 분포

예비 실험으로 2명의 피험자를 대상으로 테스트한 후 설문지의 내용을 수정했으며, 색채 용어의 순서를 가나다순으로 재배열해서 필요시 쉽게 찾아 볼 수 있도록 했다. 실험은 [유형 A]와 [유형 B]로 나누었는데, [유형 A]는 색채 용어를 제시하고 색판에서 색채편을 고르게 하는 실험이며, [유형 B]는 색판에서 색채편들을 제시하고 그 색채편에 적합한 색채 용어를 고르게 하는 실험이었다. 전체 실험의 절차는 먼저 색맹 테스트를 실시했다. [유형 A] 설문지를 나눠 준 뒤 설문지에 기재된 설문 답변 방법을 읽어 준 후 설문을 작성하게 했다. 설문 작성이 끝나고 5분의 휴식을 두고 다시 [유형 B] 실험을 실시했다. 실험에 참가한 11명의 피험자가 작성한 실험지 중 두 명이 작성한 것은 제외했는데, 하나는 '없다'라는 답변이 설문지 문항의 40%를 넘으며 다른 하나는 집단 내 다른 피험자의 실험 결과에 너무 상이하게 작성되어 질문의 내용을 잘 이해하지 못한 것으로 판단되었다.

4. 결과 분석

4.1 실험 결과

[유형 A] 실험은 색채 용어를 제시하고 색판에서 색채편을 선택하는 실험으로, 색채 용어와 색채의 일치하는 답변 비율은 아래 [표 2]와 같다.

[표 2] [유형 A] 실험 결과의 색채 용어와 색채의 일치 비율

일치 비율	색채 용어 (색판 A의 색채편 번호)[9]	비고
100%	노랑 / 노란색 (3), 보라(색) (9), 빨강(색) (1), 회색 (32)	
88.9%	녹색 (5), 분홍(색) (20), 붉은(색) (1), 오렌지색 (2), 주황(색) (2), 진보라 (10), 파랑 / 파란색 (8)	
77.8%	곤색 (28), 연두(색) (4), 진분홍 (10), 초록(색) (5), 흰색 / 하양 / 하얀색 (31)	
66.7%	갈색 (22), 검은색 (33), 선홍(색) (10), 적갈색 (21), 핑크(색) (20), 하늘색 (18)	
50% 이하	감색, 감청색, 고동색, 다홍(색), 미색, 밤색, 베이지색, 연분홍, 연초록, 자주(색), 주홍, 진홍(색), 청록(색), 청색, 푸른색, 황갈색, 황색, 회갈색, 흑갈색	

모든 피험자들이 '노랑 / 노란색, 보라(색), 빨강(색), 회색'의 경우는 동일한 색채편을 골랐으며 나머지 용어의 경우에는 모두가 일치하지는 않았다. 일치하는 정도가 50% 이하인 경우도 있었다.

[유형 B] 실험은 색채편을 주고 색채 용어를 선택하게 하는 실험으로, 색채 용어와 색채의 일치하는 답변 비율은 [표 3]과 같다.

[표 3] [유형 B] 실험 결과의 색채와 색채 용어의 일치 비율

일치 비율	색판 A의 색채편 번호[10] (색채 용어)	비고
100%	13 (노랑 / 노란색), 3 (노랑 / 노란색), 32 (회색), 31 (흰색 / 하양 / 하얀색)	'노랑'에 2개의 색채편이 선택됨
88.9%	20 (분홍(색)), 1 (빨강(색)), 33 (검은색)	
77.8%	18 (하늘색), 9 (보라(색)), 29 (진보라)	
66.7%	5 (초록(색)), 4 (연두(색)), 27 (청록(색))	
50% 이하	생략	

'노랑 / 노란색, 회색, 흰색 / 하양 / 하얀색'의 경우는 색채편과 용어가 모두 일치했다. 특히 '노랑 / 노란색'은 두 개의 색채 편에 선택되었으며 이에 대한 피험자간 일치도도 100%였다.

위 실험 결과에서 색채 용어와 색채의 일치도가 높은 것과 일치도가 낮은 것은 어떤 차이가 있을지 생각해 보기로 하자. 일반적으로 일치도가 높은 것은 해당 용어가 "현저성(saliency)"이 높은 색채를 나타내고 일치도가 낮은 것은 현저성이 떨어지는 색채를 나타낼 것이다.

9) 예를 들어, "노랑 / 노란색 (3)"은 노랑 / 노란색에 색채편 3번이 선택되었음을 표시한다.

10) 실제로는 색판 B를 사용해 실험했으나 [유형 A] 실험 결과와의 일관성을 주기 위해 색판 A의 번호로 기술하겠다. 색판 A와 색판 B는 같은 개수의 색채편을 가지고 있으나 그 배열순서만 다르다.

[그림 4] 한 색채 용어에 여러 색채가 대응된 예

[그림 5] 한 색채에 여러 용어들이 대응된 예

[그림 4]는 '자주색'이라는 용어에 세 개의 색채편이 배당된 것을 나타내고 [그림 5]는 하나의 색채편에 여러 개의 용어가 배당된 것을 나타낸다. 이런 관계를 가지고 있는 용어들은 현저성과 사용 빈도가 낮은 특징이 있다.

'푸른색, 청록색, 청색' 등의 경우에는 일치도가 낮을 뿐만 아니라 한 색채 용어에 여러 색채편이 연결되었다. 이는 관련 색채의 현저성이 낮은 이유도 있겠지만 한국어 고유의 색채 분류 체계에서 기인한 바도 있다고 생각된다. 전통적으로 '푸르다, 파랗다' 등의 용어가 '초록'과 '파랑'의 두 색채를 모두 포괄하여 지칭하는 용어이기 때문일 것이다.

4.2 통계 분석

앞에서 언급한 바와 같이 실험의 결과는 카파 통계와 계층적 군집 분석을 활용해서 분석하였다. 카파 통계는 피험자 상호간의 일치도(inter-rater agreements)에 대한 신뢰성 정도를 나타내기 위한 통계적 장치이다 (Careltta 1996). 카파 통계를 산출하기 위한 다음 수식에서, P(A)는 관찰되는 일치이고 P(E)는 우연한 일치이다.

(2) $\kappa = \dfrac{P(A) - P(E)}{1 - P(E)}$

연산의 결과로 얻어진 수치는 신뢰도의 정도에 따라서 다음의 수치적 해석이 가능하다 (Landis and Koch 1977).

(3)

κ 값	해석
0.00	거의 일치하지 않음 (Poor)
0.21∼0.40	약간 일치함 (Fair Reliability)
0.41∼0.60	보통 정도의 일치를 보임 (Moderate Reliability)
0.61∼0.80	실질적으로 일치함 (Substantial Reliability)
0.81	거의 완벽하게 일치함 (Perfect Reliability)

[유형 A] 실험의 전체적 신뢰도를 나타내는 카파 통계 값은 0.34로 다소간 일치한다고 볼 수 있는 정도인데, 이는 실험 결과가 통계적으로 어느 정도의 신뢰성이 있음을 나타낸다. 설문지에 제시한 41개의 색채 용어 중 [유형 A] 실험 결과에서 색채와의 개별적 일치도를 나타내는 "Si 값"이 0.5가 넘는 색채어는 15개였다.[11] 이 중 주어진 색채 용어에 색채가 100% 일치하는, 즉 Si 값이 1인 색채 용어와 색채는 노랑 / 노란색 (5Y 8 / 12), 보라색 (5P 3 / 11), 빨강(색) (5R 4 / 16), 회색 (N5)이다. 이들의 공통점은 무채색인 회색을 제외하고는 모두 순색이라는 것이다.[12]

11) "Si 값"은 카파 통계 값을 산출하는 과정에서 나오는 계산 값인데, 각 "실험 항목(object)"이 전체 피험자들의 실험 결과에 대비하여 어느 정도로 일치하는지를 보여 주는 수치이다. [유형 A] 실험에서 실험 항목인 색채 용어의 Si 값이 0.5가 넘는 것은 다음과 같다: 갈색 (5YR 4 / 7), 곤색 (5PB 3 / 8), 노랑 / 노란색 (5Y 8 / 12), 녹색 (5G 5 / 11), 보라색 (5P 3 / 11), 분홍색 (5RP 7 / 8), 빨강색 (5R 4 / 16), 선홍색 (5RP 4.5 / 13), 연두색 (5GY 7 / 11), 오렌지, 주황색 (5YR 6.5 / 14), 진보라색 (5P 2.5 / 7), 파랑 / 파란색 (5PB 4 / 12), 회색 (N5).

12) 먼셀 공간의 10 가지 기본색은 순색 R (빨강), Y (노랑), G (초록), B (파랑), P (보라)와 순색들이 혼합된 RY (주황), YG (연두), BG (청록), PB (남색), RP (자주)로 이루어져 있다.

[표 4] [유형 A] 실험의 실험 항목 중 Si가 0.5가 넘는 색채 용어

그 외에 Si 값이 0.5 이하인 색채 용어는 다음과 같다: 감색(0.07), 감청색(0.19), 검정(0.33), 고동색(0.14), 다홍색(0.19), 미색(0.29), 밤색(0.25), 베이지색(0.40), 연분홍(0.36), 연초록(0.18), 자주(0.21), 적갈(0.27), 주홍(0.14), 진분홍(0.40), 진홍(0.07), 청록(0.11), 청색(0.19), 초록색(0.48), 푸른색(0.07), 핑크색(0.44), 하늘색(0.20), 황갈색(0.18), 황색(0.11), 회갈색(0.18), 흑갈색(0.19), 흰색/하양(0.47). 여기서 특이한 사항 두 가지는 무채색을 3개밖에 제시하지 않았음에도 불구하고 회색만 Si 값이 1이었다는 것과 유채색 중 먼셀 공간의 순색에 속하는 R (빨강), Y (노랑), P (보라)만이 Si 값이 1이며 같은 순색에 속하는 B와 G, 즉 파랑과 초록의 Si가 0.7을 넘지 않는다는 것이다. 이것은 앞의 실험 결과 분석에서 지적한 바와 같이 한국어에서 B, G가 현저성이 떨어지는 특성이 반영된 것으로 보인다.

[유형 B] 실험의 카파 통계값은 0.49로 이전 [유형 A] 실험의 통계 결과보다 더 신뢰성이 높았다. [유형 B] 실험 결과에서 색채편 중 Si 값이 0.5가 넘는 것은 총 11개인데, 색채순으로 정리하면 다음과 같다: [5GY 8.5/9] (연두색), [5P 2.5/7] (진보라색), [5P 3/11] (보라색), [5P 7/8] ('모르겠다'로 표시), [5R 4/16] (빨강색), [5RP 7/8] (분홍), [5Y 8/12]/[5Y 9/19] (노랑), N1.5 (검정), N5 (회색), N9.5 (흰색).

[표 5] [유형 B] 실험 항목 중 Si가 0.5가 넘는 색채편

이 중 Si 값이 1인 것은 [5Y 8/12], [5Y 9/9], N5, N9.5로 먼셀 공간의 순색인 노랑과 무채색인 회색과 검정이다. 또한 [유형 B]에서는 순색 중 B와 G의 색채 용어와 색채간의 일치도가 0.5 이하로 나타났는데, 이 현상도 B와 G의 색채가 동일한 용어로 지칭된 한국어적 특성 때문인 것으로 보인다.

[유형 A]와 [유형 B]의 두 실험에서 Si가 모두 1인 색채 용어와 색채들은 "노랑/노란색－[5Y 8/12)]" 쌍과 "회색－N5" 쌍이다. 그리고 '노랑/노란색'과 관련 색채편과의 일치도가 가장 높았다. 그 외에 Si가 공통적으로 0.5 이상인 것 중 색채 용어와 색채의 관련 일치도가 높은 것 순으로 나열하면, "[5P 2.5/7]－진보라색," "[5P 3/11]－보라색," "[5R 4/16]－빨강(색)," "[5RP 7/8]－분홍"이 있다.

두 실험에서 색채 용어와 색채의 일치 관계 유형은 일대일 관계와 일대다 관계로 나누어진다. 일대일 관계는 일정 비율 이상의 다수에 의해 하나의 색채 용어에 하나의 색채가 대응되는 경우나 그 반대의 경우로 볼 수 있다.[13] 예를 들면, 색채 용어 '빨강(색)'은 색채판의 적색 계열 색채편 중 하나의 색채편과 연결되어 일대일 관계를 나타냈다. 반면에 색채 용어 '노랑 / 노란색'은 다수에 의해 두 개의 색채편과 연결되어 일대다 관계를 나타냈다.

아래 [그림 7]은 카파 통계에서 Si 값이 0.6 이상인 색채 용어 14개와 색채를 대상으로 색채 용어들이 어떤 구조를 갖는지, 계층적 군집 분석 방법을 이용한 통계적 분석을 통해 "덴드로그램(Dendrogram)"으로 나타낸 것이다.

13) 본 실험에서는 피험자들간의 일치도가 60% 이상인 것만 통계 분석의 대상으로 삼았다. 일치도가 60% 미만인 것은 모국어 화자들에게 색채 용어와 색채간의 관련성이 잘 정립되어 있지 않다고 판단하여 제외했다.

[그림 7] 덴드로그램을 통한 분석

‘연두1’과 ‘연두2’ 및 ‘노랑1’과 ‘노랑2’는 [유형 A]와 [유형 B]의 실험 결과에서 한 색채 용어가 2개의 색채에 관련되는 경우를 표시한 것이다. 위의 색채 용어 계층 구조를 보면, 크게 보아 오렌지색부터 녹색까지의 부류와 곤색부터 분홍색까지의 부류로 나누어진다. 전자는 (녹색을 제외하고는) 노란색 계통으로 볼 수 있고 후자는 보라색과 빨강색 계통으로 볼 수 있을 것 같다. 색채들간의 친소 관계가 이웃하는 색채들에 따라 상대적으로 정해지는 것이지만, ‘녹색’과 ‘파랑색’이 서로 다른 부류에 속한다는 것은 예상 밖의 결과라고 할 수 있다. 위에서도 언급했듯이, 전통적으로 우리말에서는 ‘푸르다, 파랗다’ 등의 용어가 초록과 파랑의 두 색채를 모두 포괄하는 것으로 사용되어 심리적으로는 가깝게 느껴지기 때문이다. 그렇지만, 초록과 파랑을 지칭하는 용어가 분리되어 있고 비교 대상 용어가 위의 14개인 색채 용어 체계에서는 ‘녹색’과 ‘파랑색’이 물리적 속성에 따라 서로 다른 부류에 속함을 보여 주고 있다.

위에서 ‘노랑’과 ‘연두’는 각각 2개의 색채편과 연결되어 있음을 보았다. 이와 같은 관계는 색채 용어와 색채가 일대다 관계임을 보인다. 이 특성은 어떤 단어가 다의성을 보이는 현상과 유사하다.[14] 그 외의 다른 색채 용어들은 색채와 일대일 대응 관계를 보였다. 일대다와 일대일 관계는 색채 용어가 물리적 색채 스펙트럼상에 분포하는 여러 색과 관계가 있거나 하나의

14) 한 익명의 심사자가 지적했듯이, 본 연구의 결과가 실제로 어떻게 (언어 중립적) 온톨로지 구축에 반영될 수 있을지에 대해서는 추가적인 연구가 필요하다. 첫째, 예를 들어, ‘노랑’과 ‘연두’를 일반 다의어와 동일하게 취급할 수 있을지, 다른 어휘들과의 의미 관계를 어떻게 포착할 수 있을지 등이 문제가 된다. 둘째, 본 연구에서 활용한 방법론의 색채 용어 계층 구조는 언어권/문화권별로 다를 수 있기 때문에 관련 온톨로지를 어떤 방식으로 언어 중립적으로 만들 수 있을지가 과제로 남는다.

색과 관계가 있음을 보여 준다. 한 색채 용어와 일대다 관계를 가진 두 개의 색채편들은 물리적 색채 스펙트럼 내에서의 거리가 더 가깝다. 그리고 일대일 관계의 용어들도 그 거리의 멀고 가까움을 관련 색채의 속성을 통해 알 수 있게 된다. 그러므로 기존의 색채 온톨로지가 자매 구조를 보이고 있는 것은 색채 용어와 연결된 색채의 가깝고 먼 관계가 수렴되지 않았기 때문으로 설명할 수 있다. 색채 용어들에 연결된 색채의 값(색채의 색상, 명도, 채도 값)들이 반영되지 않으면 색채 용어들은 모두 동등한 위치로 나열된다. 그러나 색채 용어에 연결된 색채의 값이 수렴될 경우, 색채의 값으로 인해 색채 용어간에 필연적으로 계층 구조가 형성된다. 따라서 자매 관계로 표시된 색채 온톨로지는 각 색채 용어와 관계를 가지는 색채의 속성을 수렴하지 않은 것으로 볼 수밖에 없다.

5. 결론

본고에서는 기존의 색채 온톨로지가 자매 관계로 표현되어 있는 것에 대한 문제점을 지적하였고, 자매 관계가 아닌 계층 구조임을 보여 주었다. 이러한 주장을 뒷받침하기 위해 색채 용어와 색채에 대한 심리언어학적 실험을 실시했다. 실험의 결과 색채 용어와 색채와의 일치도 정도에 따라 현저성의 관계가 정비례함을 보였다. 다음으로 카파 통계 분석 결과를 통해 색채 용어와 색채의 일치 관계 유형을 일대일 관계와 일대다 관계로 나누어 살펴 보았다. 마지막으로 색채 용어와 색채의 계층 구조 분석에서 나타난 색채 용어들 간의 관계가 색채 용어에 연결된 색채의 속성에 따라 가깝고 먼 관계로 나타남을 설명했으며 이것으로 인해 기존의 색채 온톨로지와는 다르게 색채가 자매 구조가 아닌 계층적 구조를 지님을 보았다.

본 연구에서 우리는 일반인 대상의 한국어 색채 온톨로지 구축을 위한 토대를 마련하였다. 그러나 언어 중립적이어야 하는 온톨로지의 본질을 살리기 위해서는 한국어 이외의 다른 언어 색채 용어들도 다루어야 한다. 그 언어들에 대해서도 기본 색채 용어를 추출해서 통계적 방식으로 처리하여 색채 용어와 색채와의 관계를 정립해야 한다. 그리고 그 결과를 본 연구의 결과와 비교해 보아야 언어 중립적인 색채 온톨로지를 구축할 수 있다.

참고문헌

김영선, 이만영. 1998. “한국어 기본 색이름과－NBS 기초유목 색 이름의 색채 공간 비교”. 한국색채학회 동계 학술대회 논문집.

문은배. 2002.「색채의 활용」. 국제.

이만영, 김영선. 1997. “한국인의 기본 색이름에 대응하는 색상에 관한 연구”.「인지과학」8(4).

Berlin, Brent and Elois Ann Berlin. 1975. Aguaruna Color Categories. *American Ethnologist* 2.
Berlin, Brent and Paul Kay. 1969 / 1991. *Basic Color Terms: Their Universality and Evolution*. University of California Press.
Carletta, Jean. 1996. Assessing Agreement on Classification Task: The Kappa Statistic. Computational Linguistics 22(2).
Landis, J. R. and G, Koch. 1977. *The Measurement of Observer Agreement for Categorical Data*. Biometrics.
Lenci, Alessandro. 2000. SIMPLE Work Package 2: Linguistic Specifications-Deliverable D2.1.
Lenci, Alessandro. 2001. Building an Ontology for the Lexicon: Semantic Types and Word Meaning. in *Ontology-based Interpretation of Noun Phrases.* Per Anker Jensen and Peter Skadhauge. eds. 130-120. University of Southern Denmark.
MacLaury, Robert E. 2001. Color terms. *Language Typology and Language Universals* Vol. 2. HSK.
Miller, G. 1998. *WordNet: An Electronic Lexical Database*. MIT Press.
Payne, D. L. 2006. Color Terms. *Encyclopedia of Language & Linguistics*. 2nd edition. Elsevier.

참고 자료

산업자원부 기술표준원. 2005. 물체색의 색이름 KS A 0011: 2005. 한국표준협회.
연세 한국어 사전 (1998/2007). 두산동아.
표준 국어 대사전. http://www.korean.go.kr/06_new/index.jsp
Art & Architecture Thesaurus. http://www.getty.edu/research/conducting_research/vocabularies/aat

노창화 · 김동성 · 채희락
경기도 용인시 처인구 모현면 왕산리 산 89
한국외국어대학교 언어인지과학과
449-791
E-mail: {rayr, dsk202, hrchae}@hufs.ac.kr

접수일자 : 2008.11.19
수정일자 : 2008.12.16
게재일자 : 2008.12.19

부록 1 연세 한국어 사전에서 추출한 색채 명사 (58개)

구릿빛, 다홍, 분홍빛, 빨강, 선홍색, 연분홍색, 적색, 진분홍, 진홍빛, 핑크색, 핏빛, 고동색, 밤색, 적갈색, 주홍색, 주황색, 홍색, 갈색, 금색, 노랑, 누렁, 다갈색, 미색, 베이지색, 암갈색, 황갈색, 황금색, 황색, 흑갈색, 연두색, 국방색, 녹색, 진초록, 청록색, 감색, 곤색, 남색, 물색, 비색, 암청색, 옥색, 청색, 청회색, 파랑색, 하늘색, 자색, 자줏빛, 보라색, 검정색, 까망, 깜장, 백색, 연회색, 유백색, 은색, 하양, 회색, 흑색.

부록 2 국립 국어원 빈도 6만에서 추출한 색채 명사 (65개)

가지색, 갈색, 감색, 감청색, 검은색, 고동색, 곤색, 국방색, 금색, 남색, 노란색 / 노랑색, 녹색, 녹황색, 누렁, 다갈색, 다홍색, 물색, 미색, 밤색, 백색, 베이지색, 보라색, 분홍색, 붉은색, 비색, 빨강색, 선홍색, 순백색, 쑥색, 암갈색, 연두색, 연분홍, 연초록, 오렌지색, 옥색, 와인색, 은백색, 자색, 자주색, 적갈색, 적색, 주홍, 주황색, 진보라, 진분홍, 진홍색, 청록색, 청색, 초록색, 파란색 / 파랑, 푸른색, 핑크색, 하늘색, 하얀색 / 하양 / 흰색, 호박색, 홍색, 황갈색, 황금색, 황색, 황토색, 회갈색, 회백색, 회색, 흑갈색, 흑색.

부록 3

Type A 설문

◎ 설문에 응해 주셔서 감사합니다. 이 설문의 정보 중 개인 신상정보는 유출되지 않으며, 피험자의 유형을 구분하기 위해 묻는 내용입니다.

성별: 출생년도: 전공(소속):

◎ 다음의 각 색채 용어에 해당하는 색을 골라 번호를 쓰세요. 하나의 색채 용어가 두 개 이상의 색을 나타낼 수도 있습니다. 용어에 해당하는 색이 없으면 × 표시를 하고 그 용어를 모를 때는 Y표시를 해 주세요.

1 갈색:	26진보라:
2 감색:	27진분홍:
3 감청색:	28진홍(색):
4 검은색:	29청록(색):
5 고동색:	30청색:
6 곤색:	31초록(색):
7 노랑 / 노란색:	32파랑 / 파란색:
8 녹색:	33푸른색:
9 다홍(색):	34핑크(색):
10미색:	35하늘색:
11밤색:	36황갈색:
12베이지색:	37황색:
13보라(색):	38회갈색:
14분홍(색):	39회색:
15붉은색:	40흑갈색:
16빨강(색):	41흰색 / 하양 / 하얀색:
17선홍(색):	
18연두(색):	
19연분홍:	
20연초록:	
21오렌지색:	
22자주(색):	
23적갈색:	
24주홍:	
25주황(색):	

Type A **실제 실험에서는 색채편 안에 먼셀의 색채값은 포함되어 있지 않다.**

1	2	3	4	5	6	7	8	9	10
5R 4/16	5YR 6.5/14	5Y 8/12	5GY 7/11	5G 5/11	5BG 4/9	5B 4/9	5PB 4/12	5P 3/11	5RP 4.5/13
11	12	13	14	15	16	17	18	19	20
5R 7/8	5YR 8/16	5Y 9/9	5GY 8.5/9	5G 7.5/9	5BG 7/8	5B 7/7	5PB 7/7	5P 7/8	5RP 7/8
21	22	23	24	25	26	27	28	29	30
5R 4/8	5YR 4/7	5Y 4/5	5GY 4/6	5G 3.5/7	5BG 3/8	5B 3/8	5PB 3/8	5P 2.5/7	5RP 3/8
31	32	33							
N 9.5	N 5	N 1.5							

Type B 설문

◎ 설문에 응해 주셔서 감사합니다. 이 설문의 정보 중 개인 신상정보는 유출되지 않으며, 피험자의 유형을 구분하기 위해 묻는 내용입니다.

성별: 출생년도: 전공(소속):

◎ 아래 숫자는 <첨부 1: 색판>에 주어진 색의 번호입니다. 해당하는 색에 어울리는 색채 용어를 <첨부2: 용어목록>에서 골라 숫자 옆에 써 주세요. 하나의 색에 두 개 이상의 색채 용어가 배정될 수도 있습니다. 용어에 해당하는 색이 없으면 × 표시를 해주세요.

1	26
2	27
3	28
4	29
5	30
6	31
7	32
8	33
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	

Type B <첨부1: 색판>

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33

Type B <첨부 2: 용어 목록>

1 갈색:	26진보라:
2 감색:	27진분홍:
3 감청색:	28진홍(색):
4 검은색:	29청록(색):
5 고동색:	30청색:
6 곤색:	31초록(색):
7 노랑 / 노란색:	32파랑 / 파란색:
8 녹색:	33푸른색:
9 다홍(색):	34핑크(색):
10미색:	35하늘색:
11밤색:	36황갈색:
12베이지색:	37황색:
13보라(색):	38회갈색:
14분홍(색):	39회색:
15붉은색:	40흑갈색:
16빨강(색):	41흰색 / 하양 / 하얀색:
17선홍(색):	
18연두(색):	
19연분홍:	
20연초록:	
21오렌지색:	
22자주(색):	
23적갈색:	
24주홍:	
25주황(색):	